Premier Reference Source

New Approaches to Data Analytics and Internet of Things Through Digital Twin

P. Karthikeyan, Polinpapilinho F. Katina, and S. P. Anandaraj

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

New Approaches to Data Analytics and Internet of Things Through Digital Twin

P. Karthikeyan
National Chung Cheng University, Chiayi, Taiwan

Polinpapilinho F. Katina
University of South Carolina Upstate, USA

S.P. Anandaraj
Presidency University, India

A volume in the Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) Book Series

Published in the United States of America by
IGI Global
Engineering Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Karthikeyan, P., 1981- editor. | Katina, Polinpapilinho F., editor. | Anandaraj, S. P., 1982- editor.
Title: New approaches to data analytics and internet of things through digital twin / P. Karthikeyan, Polinpapilinho F Katina, and S.P. Anandaraj, editors.
Description: Hershey, PA : Engineering Science Reference, an imprint of IGI Global, [2023] | Includes bibliographical references and index. | Summary: "This book investigates that though many data analytics tools have been developed in the past few years, their usage in the field of Digital Twin warrants new approaches considering many aspects including unified data representation, zero-day attack detection, data sharing across threat detection systems, real time analysis, sampling and dimensionality reduction, resource-constrained data processing, and time series analysis for anomaly detection"-- Provided by publisher.
Identifiers: LCCN 2022018654 (print) | LCCN 2022018655 (ebook) | ISBN 9781668457221 (h/c) | ISBN 9781668457238 (s/c) | ISBN 9781668457245 (ebook)
Subjects: LCSH: Internet of things. | Digital twins (Computer simulation) | Data mining.
Classification: LCC TK5105.8857 .N49 2023 (print) | LCC TK5105.8857 (ebook) | DDC 004.67/8--dc23/eng/20220701
LC record available at https://lccn.loc.gov/2022018654
LC ebook record available at https://lccn.loc.gov/2022018655

This book is published in the IGI Global book series Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) (ISSN: 2327-3453; eISSN: 2327-3461)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) Book Series

Vijayan Sugumaran
Oakland University, USA

ISSN:2327-3453
EISSN:2327-3461

MISSION

The theory and practice of computing applications and distributed systems has emerged as one of the key areas of research driving innovations in business, engineering, and science. The fields of software engineering, systems analysis, and high performance computing offer a wide range of applications and solutions in solving computational problems for any modern organization.

The **Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) Book Series** brings together research in the areas of distributed computing, systems and software engineering, high performance computing, and service science. This collection of publications is useful for academics, researchers, and practitioners seeking the latest practices and knowledge in this field.

COVERAGE

- Metadata and Semantic Web
- Computer Networking
- Human-Computer Interaction
- Performance Modelling
- Storage Systems
- Computer Graphics
- Network Management
- Parallel Architectures
- Virtual Data Systems
- Distributed Cloud Computing

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: http://www.igi-global.com/publish/.

The Advances in Systems Analysis, Software Engineering, and High Performance Computing (ASASEHPC) Book Series (ISSN 2327-3453) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, www.igi-global.com. This series is composed of titles available for purchase individually; each title is edited to be contextually exclusive from any other title within the series. For pricing and ordering information please visit http://www.igi-global.com/book-series/advances-systems-analysis-software-engineering/73689. Postmaster: Send all address changes to above address.

Titles in this Series

For a list of additional titles in this series, please visit: http://www.igi-global.com/book-series/

Futuristic Trends for Sustainable Development and Sustainable Ecosystems
Fernando Ortiz-Rodriguez (Tamaulipas Autonomous University, Mexico) Sanju Tiwari (Tamaulipas Autonomous University, Mexico) Sailesh Iyer (Rai University, India) and José Melchor Medina-Quintero (Tamaulipas Autonomous University, Mexico)
Engineering Science Reference • © 2022 • 320pp • H/C (ISBN: 9781668442258) • US $270.00

Emerging Technologies for Innovation Management in the Software Industry
Varun Gupta (Universidad de Alcalá, Madrid, Spain) and Chetna Gupta (Jaypee Institute of Information Technology, Noida, India)
Engineering Science Reference • © 2022 • 282pp • H/C (ISBN: 9781799890591) • US $270.00

Technology Road Mapping for Quantum Computing and Engineering
Brojo Kishore Mishra (GIET University, India)
Engineering Science Reference • © 2022 • 243pp • H/C (ISBN: 9781799891833) • US $250.00

Designing User Interfaces With a Data Science Approach
Abhijit Narayanrao Banubakode (MET Institute of Computer Science, India) Ganesh Dattatray Bhutkar (Vishwakarma Institute of Technology, India) Yohannes Kurniawan (Bina Nusantara University, Indonesia) and Chhaya Santosh Gosavi (MKSSS's Cummins College of Engineering, India)
Engineering Science Reference • © 2022 • 325pp • H/C (ISBN: 9781799891215) • US $270.00

Implementation of Machine Learning Algorithms Using Control-Flow and Dataflow Paradigms
Veljko Milutinović (Indiana University, Bloomington, USA) Nenad Mitić (University of Belgrade, Serbia) Aleksandar Kartelj (University of Belgrade, Serbia) and Miloš Kotlar (University of Belgrade, Serbia)
Engineering Science Reference • © 2022 • 296pp • H/C (ISBN: 9781799883500) • US $270.00

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

701 East Chocolate Avenue, Hershey, PA 17033, USA
Tel: 717-533-8845 x100 • Fax: 717-533-8661
E-Mail: cust@igi-global.com • www.igi-global.com

Table of Contents

Detailed Table of Contents

The data generated by the big data-based clinical need analysis plays a key role in improving the consideration feature, decreasing waste and blunder, and reducing treatment expenses. The use of big data analytics (BDA) techniques for analyzing disease and predictions is discussed in this investigation. This precise survey of writing means to decide the extent of BDA in disease analysis and difficulties in treatment in the medical filed. Further, this study has discussed the comparative analysis of heart diseases, predictions using BDA techniques, predicting of breast cancer, lung cancer, and brain diseases. Digital twins will be key to delivering highly personalized treatments and interventions. Intelligent digital twins, combining data, knowledge, and algorithms (AI), are set to revolutionise medicine and public health.

The chapter aims to embed the demanding computing concepts to attain intelligent data analytics in the domain of healthcare. The targeted outcome provides the pathway to design the brainy decision support system needed to have efficient prediction with trained input patterns. The usage of IoT devices is increasing tremendously to overcome the challenges existing in handling the data related to human-relevant happenings. The volume, velocity, and variety of data are emerging newly and dominating the decision support characteristics. This scenario happens almost in all the computing fields, but more attention is expected to implement in the healthcare sector due to the existence of sensitive data. The traditional data analytics methods are deviating in the performance due to the unpredictable dynamic challenges emerging in the day-to-day operation. The efficient features of demanding computing strategies are motivated to embed together to discover crucial knowledge through intelligent data analytics.

The term internet has become more popular these days, and the whole world is connected virtually. Most people communicate through the internet. Communicating information from one to another in network without disclosing to third party is a typical task, but tracking of information is performed when certain security measures are not taken during transmission between the real-time environment and virtual environment. Creation of virtual representation for real-time process is known for the digital world. Secret data communication is performed under the digital world concept where there is a need of security. The chapter introduces a mechanism providing security using the ISDHR technique with least significant matching and dual images for secret message communication. The results show the proposed method is high enough in providing security to the secret data under transmission.

This chapter acquaints the reader to the terms and terminologies of cyber-attacks, cybersecurity, big data, data analytics, and related new age technologies, including deep learning. The types of cyber-attacks, how they become special and different within the big data analytic frameworks, a multi-layer framework for their detection, and the challenges therein are detailed next. Thereafter, an extensive review of some research works has been undertaken to provide an in-depth insight to the various cyber security detection systems using the new age technologies such as naive Bayesian networks in intrusion detection systems, deep learning in Android malware detection, and intelligent malware detection, etc. Conclusions have been drawn from these studies to establish that the emerging technologies, like artificial intelligence, machine learning, deep learning, and internet of Things, are the need of the hour to assist organizations in navigating the increasingly aggressive cyber threat landscape.

According to cognitive data analysis theory, career thoughts of students mediate the relationship between career, stress, and the career decision state. A student who opts for higher education and is new to this environment is undergoing stress. The work is focused on identifying behavioral, cognitive, and emotional features of the student, which have an impact on professional careers. So, early recognition of these features is necessary. This chapter presents an analysis by following some statistical approaches of machine learning and mathematical concepts. The study found that an increase in career and life stress is associated with a lower level of decidedness and satisfaction with career choice among genders. The results suggests that counselling is needed to aid students in their career thoughts and decision making.

The internet of things (IoT), big data analytics, artificial intelligence (AI), and cybertwin, as well as other digital technology and designed intelligence have accelerated the 4th industrial revolution known as Industry 4.0. Industry 4.0 applications must construct complicated machine representations from such fundamental pieces, which is a time-consuming, error-prone, and wasteful process that impedes machine and plant mobility. Cybertwin, a comprehensive solution for fast Industry 4.0 application creation, testing, and porting, is proposed in this study. The deployment of cybertwin with IIoT will enhance the efficiency and accuracy of real-time IIoT applications. Further, these huge mixtures of data will be analyzed by using big data analytic tools to produce intensive incident commands, and it is further deeply analyzed to discover various knowledge, which supports redesign and reengineering of the specific process. The cloud computing platform will be utilized to achieve big data analytics effectively.

Due to the blooming of Industrial 4.0 such as internet of things (IoT), cloud computing, industrial IoT (IIoT), and artificial intelligence (AI), with their innovative ideas and opportunities, the cyber attacker's modus operandi against the cyber defense triage is incredible. The genre of advanced persistent threat (APT) actors/group are equipped with sophisticated and substantial resources of tools, techniques, and procedure (TTP) at a breakneck pace. The IoT gadgets such as sensors, intelligent devices, and various rapidly emerging resources with energy, memory, and processing power are exponentially prone to multiple vulnerabilities. The nature of IIoT prompts heterogenous and rapid changes ranging the vulnerabilities from simple to complex attacks. APT menace follows the covert TTPs to target the asset of any organization like the government, military, or financial industry.

Preface

New Approaches to Data Analytics and Internet of Things Through Digital Twin provides an overview of recent research and development activities in the field of data analytics and Internet of Things (IoT) through digital twin. This book contains 11 chapters from the basic concept level to the research and application level.

Data analytics will be a must-have component of any effective digital twin solution due to the need for fast processing of the high-velocity, high-volume data from various sources to discover anomalies and attack patterns as fast as possible to limit the vulnerability of the systems and increase their resilience. Internet of Things is closely related to individuals' life and social improvement, and addition IoT has a wide application in the military, including aeronautics, military surveillance, knowledge framework, intelligent transportation, ecological examination, mechanical control, and so forth. An intelligent transportation system generates an enormous amount of data such as streets, connections, crossing point, traffic signs and other key data. The tremendous measure of data is broken down, discharged, and determined by the framework, with the goal that the street vehicles can share street data progressively. Even though many data analytics tools have been developed in the past few years, their usage in the field of digital twin warrants new approaches considering many aspects, including unified data representation, zero-day attack detection, data sharing across threat detection systems, real-time analysis, sampling and dimensionality reduction, resource-constrained data processing, and time series analysis for anomaly detection.

Digital twin improves the efficiency in cyberspace by eliminating unnecessary access to the mission-critical applications in the core cloud. Data analytics and the internet of things in the digital twin has emerged as the standard platform in recent years for securing a large pool of transaction in cyberspace. The main scope of this book is to discuss recent data analytic application which uses IoT and twin.

All the books published earlier by different authors only address the theoretical study of data analytics and IoT or do not address the practical applications/ implementation of various digital twins in various fields like medical, security, networking etc. Hence, it is decided to propose a book which not only discusses the research issues in various domains but also solves those problems with the help of the digital twin. This book also focuses on three categories of users: beginners, intermediate, and sophisticated readers and provides content accordingly. So, this is very useful for students, academicians and research scholars to explore further in their field of study. It is very much opted for readers who seek to learn from examples.

This book is a collection of 11 chapters authored by 31 academics and practitioners worldwide. The contributions in this book aim to enrich the information system discipline by providing the latest research and case studies from around the world. These are organized as follows:

Chapter 1: This chapter provides a comprehensive and unified view of Data Analytics. Data Analytics is the process of analyzing the raw data in order to draw inferences about the information at hand. Data analysis techniques are primarily used to get an insight which further facilitates the enhancement of the sector under consideration. These techniques are beneficial for optimizing a process under consideration and increasing a system's overall efficiency. These techniques also act as performance boosters as their implementation in the business models helps in the reduction of costs by a considerable amount. It is the most important for any organization as it facilitates a better decision-making approach and provides an analysis of customer trends and satisfaction, which further leads to novice and improved products and services. It also helps in the effective marketing of the products and services. Data analytics has widespread applications in various sectors. Various tools are used for carrying out data analytics jobs. All this is discussed in the proceeding chapter.

Chapter 2: Data management in the IoT context should act as a layer between the objects and devices that generate the data and the applications that access the data for analysis and services. The devices can be organized into independent subsystems or subspaces with internal hierarchical management. Depending on the level of privacy desired by the subsystem holders, the functionality and statistics collected by these subsystems will be made accessible to the IoT network. A Digital Twin is a synchronized virtual representation of real-world entities and processes. Digital Twins can represent the past, present, and predicted futures. Digital Twin Systems accelerate holistic understanding, optimal decision-making, and effective action in business.

Chapter 3: Recent advances in the Internet of Things (IoT) and sensing devices have accelerated the rise of data-intensive applications across a wide range of application domains, including online gaming, augmented/virtual reality, path navigation, autonomous driving, and video streaming, and so on. New Approaches to Data Analytics and Internet of Things Through Digital Twin considers how data analytics and the internet of things can be used successfully within the field of a digital twin, as well as the potential future directions of these technologies. Edge computation refers to the communication of IoT models with cloud models to execute analytical data. Nodes that can be accessible by network devices in order for them to interact with each other.6G is the sixth generation presently further down progress for wireless communications network technologies supportive for cellular data networks. 6G networks are determined to use more complex frequencies than 5G networks, which will empower higher data rates to stand achieved and for the 6G network to have a much superior global volume. Much lower latency levels will almost definitely be a requirement. Cyber twin architecture introduces first for 6G.

Chapter 4: The Internet of Things "IoT" has been a buzzword in recent years, the business organization has adopted the technology in an increased way. The current expansion of IoT demands intelligent models to be deployed at the edge. The Internet of Things (IoT) is escalating into diverse aspects of our lives with innovative technologies and solutions. In general, IoT devices are restricted to storage and processing power, which results in the lack of performance, reliability and privacy of IoT applications. The applications in various sectors like agriculture, health care, smart cities, smart homes, and production units are enriched by twining the IoT and cloud computing. Edge Computing is an archetype to push cloud services from the network core to the network edges. Edge Computing aims to host computing tasks as close as possible to the data sources and end-users. Edge computing and cloud computing are not mutually exclusive. Edge computing is the extension of the cloud and complements instead. The core functions of edge computing are to perform analytics at the edge, and data needs to be viewed as real-time flows. Edge computing continually processes streaming flows of data in motion. In contrast, big data analytics is focused on large quantities of data at rest.

Chapter 5: The data generated by the big data-based clinical need analysis plays a key role in improving the consideration feature, decreasing waste and blunder, and reducing treatment expenses. The use of Big Data Analytics (BDA) techniques for analyzing disease and predictions is discussed in this investigation. This detailed survey of writing means to decide the extent of BDA in disease analysis and difficulties in treatment in the medical field. Further, this study has discussed the comparative analysis of heart diseases, predictions using BDA techniques, and predicting breast cancer, lung cancer and brain diseases. Intelligent digital twins, combining data, knowledge, and algorithms (AI), are set to revolutionize medicine and public health. Digital twins will be key to delivering highly personalized treatments and interventions.

Chapter 6: This chapter provides the pathway to design the brainy decision support system needed to have efficient prediction with trained input patterns. IoT devices are increasing tremendously to overcome the challenges in handling the data related to human-relevant happenings. The volume, velocity, and variety of data are emerging and dominating the decision support characteristics. This scenario happens almost in all the computing fields. However, due to sensitive data, more attention is expected to be paid to the healthcare sector. The traditional data analytics methods are deviating in the performance due to the unpredictable dynamic challenges emerging in the day-to-day operation. The efficient features of demanding computing strategies are motivated to embed together to discover crucial knowledge through intelligent data analytics.

Chapter 7: The term internet has become more popular nowadays, and the whole world is connected virtually. Most people communicate through the internet. Communicating information from one to another in-network without disclosing it to a third party is a typical task. However, tracking information is performed when specific security measures are not taken during transmission between the real and virtual environment. The creation of virtual representation for the real-time process is known in the digital world. Secret data communication is performed under the digital world concept where security is needed. Our proposal introduces a mechanism providing security using the ISDHR technique with the least significant matching and dual images for secret message communication. Our results show that the proposed method is high enough to provide security to the confidential data under transmission.

Chapter 8: This chapter familiarizes the reader with the basic notions of cyber-attacks, cyber security, big data, data analytics and the much-acclaimed new age technologies such as AI, Machine Learning, Data Mining, IoT and Deep Learning. Insights into the types, stages, and detection of cyber-attacks within the Big Data Analytic frameworks and the challenges are discussed. Established research in cyber security, big data and deep learning are reviewed and discussed in substantial detail to establish that integrated deep learning and big data methods demonstrate far better performances to prevent, detect and respond to cyber-attacks than the traditional cyber security methods. Looking at the results, it is concluded that these technologies are not just today but will also continue to be a part of the near future of the cyber security landscape.

Chapter 9: Discuss the cognitive data analysis theory, career thoughts of students mediate the relationship between career stress and the career decision state. The work focuses on identifying the student's behavioural, cognitive and emotional features which impact professional career. A student who opts for higher education and is new to this environment is undergoing stress. So, early recognition of these features is also necessary for the student and education system. This paper presents an analysis by following some statistical approaches to machine learning and mathematical concepts. The study found that a career and life stress increase is associated with lower decidedness and satisfaction with career choices among genders. The results suggest that counselling is needed to stress students in their career thoughts and decision makings.

Chapter 10: State-of-the-art technologies such as the Internet of Things (IoT), cloud computing, big data analytics, and artificial intelligence have greatly stimulated the development of intelligent manufacturing. An essential prerequisite for smart manufacturing is cyber-physical integration, which manufacturers increasingly embrace. As the preferred means of such integration, cyber-physical systems and digital twins have gained extensive attention from researchers and practitioners in the industry. This chapter covered the various dimensions of Industrial IoT and the specifications of the most widely used sensors. The structure and deployment of cyber twins in IIoT were discussed. Covered the Bigdata analytics tools and technologies which will majorly support the IIoT and given the mechanisms which would find and evolve the hidden facts in the industrial environment. Finally, the chapter provides the various research directions and scope toward big data analytics on IIoT.

Chapter 11: This chapter discuss the Industrial 4.0 such as Internet of Things (IoT), Cloud computing, Industrial IoT (IIoT) and Artificial Intelligence (AI) etc., with their innovative ideas and opportunities, the cyber attacker's modus against a cyber defense triage is incredible. The genre of Advanced persistent threat (APT) actors / group are equipped with sophisticated and substantial resources of Tools, Techniques and procedure (TTP)at a breakneck pace from far-flung. The IoT gadgets such as sensors, intelligent devices and various rapidly emerging resources with energy, memory and processing power are exponentially prone to multiple vulnerabilities. The nature of I-IoT prompts heterogenous and rapid changes ranging the vulnerabilities to simple to complex attacks. APT menace follows the covert TTP's to target the asset of any organization like the government, military, or financial industry.

P. Karthikeyan
National Chung Cheng University, Chiayi, Taiwan

Polinpapilinho F. Katina
University of South Carolina Upstate, USA

S. P. Anandaraj
Presidency University, India

Acknowledgment

One aspect of the academic citation apparatus is acknowledging the relevance of the works of others to the topic of discussion—this we have done. The editors also wish to acknowledge different people and organizations involved in discussions of this research.

Most significantly, our gratitude goes to the following: Dr. Tyrone Toland, Informatics and Engineering Systems (University of South Carolina Upstate, Spartanburg, SC, USA); Dr. Charles B. Keating, Engineering Management and Systems Engineering (Old Dominion University, Norfolk, VA USA).

Editors are also grateful to graduate students and young researchers in the Department of Computer Science & Engineering (Jain Deemed University, India), Department of Computer Science Engineering (Presidency University, India), and Department of Informatics and Engineering Systems (University of South Carolina Upstate, USA).

We would like to thank Dr. A.M. Viswa Bharathy (GITAM University, Bengaluru, India) for guiding us in each stage of the book with his expertise and willingness to help at any time.

For help with preparing the manuscript, we are thankful to Jan Travers (IGI Global) for her encouragement throughout the project. Our sincere apologies to everyone we might have overlooked.

Chapter 1
Data Analytics:
An Overview

Anu Sayal
Jain University (Deemed), India

ABSTRACT

This chapter provides a comprehensive and unified view of data analytics. Data analytics is the process of analyzing the raw data in order to draw inferences about the information in hand. Data analysis techniques are primarily used to get an insight which further facilitates enhancement of the sector under consideration. These techniques are beneficial for optimizing a process under consideration and also for increasing the overall efficiency of a system. These techniques also act as performance boosters as their implementation in the business models help in reduction of costs by considerable amount. It is the most important for any organization as it facilitates better decision-making approaches and also provides an analysis of customer trends as well as satisfaction which further leads to improved products as well as services. It also helps in effective marketing of the products and services. Data analytics has widespread application in various sectors. Various tools are used for carrying out data analytics jobs. All this is discussed in the chapter.

INTRODUCTION: WHAT IS DATA?

Data is basically referred to as individual facts, statistics or any type of information usually in the form of numerical values. If considered in technical sense the data consists of set of values in the form of quantitative or qualitative values which may be about one or more person or object. Data and information are used in an interchangeable manner. Considering some specific cases, the data is converted into

DOI: 10.4018/978-1-6684-5722-1.ch001

information when it is being observed in some context or in any kind of post-analysis whereas in academic sense data are purely the elements of information. Data are used in various fields like finance, scientific research, business management, governance and in every other form of the human organizational activities.

Data are also referred to as the atoms of decision making, they are the smallest units related to factual information that can be utilised discussion, reasoning or calculation purpose. Data can vary from abstract ideas to concrete measurements and even statistics. Data are collected, classified, tabulated and analysed for specific purposes. They are also used for the creation of data visualizations like tables, graphs or even images. Data processing usually takes places in different stages. Data is transformed into some suitable information for making decisions once it has been analysed.

The main objective of this chapter is that it deals with the process of data analytics. The concepts related to the types of data, types of data analytics, data mining process, IoT and digital twins have been explained in a lucid manner. The techniques and the methodologies required for each type of process have been clearly explained in the chapter. In this chapter challenges in the field of data analytics have been discussed and also the solutions related to these challenges have been proposed. At the end of the chapter a comprehensive conclusion has been provided to sum up all the detailed concepts included in the chapter.

Literature review

Grubbs and Frank (1950) framed the criteria for a sample testing the outlying observations. Chakravarti et. al. (1967) prepared a handbook related to the methods in the field of applied statistics. Anscombe (1973) conducted research on the graphs related to the statistical analysis. Box et. al. (1978) compiled their work related to the data analysis, design and also model building. Draper and Smith (1981) copiled their work related to the applied regression analysis. Snedecor et. al. (1989) compiled their work in the form of statistical methods. Barnett and Lewis (1994) conducted their research on the outliers and compiled their work in the form of outliers in statistical data. Parnas and Madey (1995) made a compilation of their work for functional documents for computer systems. Berry and Linoff (2000) concentrated their research on data mining process and compiled it. Giorgini et al. (2005) conducted research on the analysis of goal-oriented form for designing the data warehousing. Supakkul and Chung (2009) directed their research to deal with the problems related to the stakeholders. Barone et al. (2010) conducted a study on the enterprise modelling for business intelligence. Hopkins and Shockley (2011) focussed their research on data analytics, big data considering the practical aspects of the related problems. Horkoff et. al. (2014) conducted their research on strategic business modelling providing the complete representation as well as the explanation of the related problems.

Before the computing devices and machines came into existence collection of data was done manually but with the advancement in technology and development of the computing machines the data was being collected by these machines and also analysed in an appropriate manner.

Categories of data

Data is further divided in two sub categories:

1. Primary data
2. Secondary data

Primary data

Primary data refers to the set of data that is collected through one's personal experience or any evidence, specifically for research purpose for the first time. It is known as the first-hand information or the raw data. The manner related to assembling this information is expensive, as this analysis is basically carried out by an agency or some external organisation so it requires investment as well as human resources. In this method it is the sole responsibility of the investigator to collect the data.

The primary modes of data collection are observation, questionnaires, surveys, telephonic interviews, personal interviews, physical testing, case studies, focus groups etc.

Advantages of Primary Data

1. Primary data basically caters the needs of the researcher and the researcher controls the data collection process.
2. Primary data has more accuracy as compared to the secondary data as there is no personal bias and it is more authentic.
3. Researcher is the sole owner of the data as it is collected through the primary research work. It is at the discretion of the researcher to make the data available to the public or to patent it or sell it.
4. It is an updated data as it is real time collected data and not obtained to older sources.
5. The data which is collected by the researcher is fully under his control as it is first hand collected data. The researcher has the full control to decide the design, methodology and the analysis techniques which have to be applied.

Disadvantages of Primary Data

1. As compared to the secondary data primary data collection is expensive so it is difficult to collect it.
2. It is laborious task as it requires more time for collection.
3. In some cases, primary data collection is not possible as it involves some complexities and commitments.

Secondary data

Secondary data refers to the data that has already been collected earlier for some other purpose. It is not related to the current research problem. It is the data collected from various sources like government publications, records of an organisation, journals, books, websites, report etc.

It is an affordable method of data collection. It saves money as well as time. But it has a disadvantage that this data has been collected for some other purpose so it cannot be utilised for carrying out research related to some other problem.

Advantages of Secondary Data

1. Secondary data is easily available in comparison to the primary data. The researcher can access different platforms through which he can obtain the secondary data.
2. It is affordable and requires very less cost for its collection.
3. The secondary data requires very little time for its collection as compared to the primary data.
4. With the help of secondary data, it is possible to carry out the longitudinal studies without any waiting time for drawing conclusions.
5. It brings out new perceptions about the primary data under consideration.

Disadvantages of Secondary Data

1. The authenticity and reliability of secondary data is not guaranteed as it is a second-hand data. The researcher has to verify the source from which the data has been collected.
2. It may be tedious for the researchers as they might have to carry out their work with the irrelevant data before they can find the actual data.
3. There is some exaggeration in the available data due to the existence of personal bias in such data.

4. The sources of the secondary data are obsolete as old data is seldom replaced by the new data.

Difference between Primary Data and Secondary Data

Primary data	Secondary data
It is the first-hand data	It is the data that has been collected earlier
It is the original data as it directly collected by the investigator	It is not the original data as it had been collected earlier for the related research work
It comprises of raw data	It is in the form of final finished data
It is suitable and more reliable as it is collected for some specific purpose	It is not reliable as it was collected earlier and may now be suitable for present research
Collection primary data requires a lot of time and money so it is expensive	Collection of secondary data requires less time and money so it is more economical
It does not require any precaution and editing as it was collected for that specific purpose	It requires precaution and editing as it was collected for some other purpose

Similarities Between Primary & Secondary Data

1. Content is same
   The content of both primary as well as the secondary data is the same as the secondary data was also a primary data when it was initially collected by the researcher.
2. Uses
   Both primary as well as secondary data find application in statistics and research. Both have the same content so they are used for carrying out similar kinds of research. Though they are collected by different methods. The method by which the data has been collected does not affect its uses, thus the data can be used to carry out similar kind of research activities.

Types of data

1. Qualitative data vs Quantitative Data
   (a) Quantitative data

Quantitative data deals with questions addressing “how much,” “how many” and “how often”. These data are expressed as numbers. Thus, it is measurable through numerical values. Statistical manipulations can be easily applied on such data. It can be represented by charts and graphs like bar graph, line graph, histogram etc.

Examples:

Scores of students in a test, weights of students, size of the shoe and room temperature.

The quantitative data is of two types: discrete and continuous.

(b) Qualitative data

Qualitative data cannot be numerically expressed and it is non-measurable. It is comprised of pictures, words, symbols etc, but not numbers. This data is known as categorical data as the evidence contained in this data can be expressed in the form of categories. It gives answers to questions like "why a particular thing happened" and "how it happened".

Examples:

Names of people, sea colours, favourite destination for holiday, ethnicity like as Indian, Asian, American etc.

2. Nominal data vs Ordinal Data
   (a) Nominal data

Nominal data refers to labelling variables without having any quantitative value. The term 'nominal' is derived from the Latin word "nomen" which refers to name.

Examples:

Gender (Male, Female), colour of hair (Black, brown, Blonde etc.), marital status (Married, Single, Widowed), ethnicity (American, Asian) etc.

These examples clearly show that no kind of basic ordering of variables is there.

(b) Ordinal data

Ordinal data refers to assigning an order to the given variables. It's the only difference between nominal and ordinal data. These data are placed in some sort of order according to their position on the scale.

Arithmetic operations cannot be applied to the ordinal numbers because these represent only sequence.

These variables are supposed to be lying in-between the quantitative and qualitative variables. Ordinal data refers to qualitative type of data having variable values in ordered form.

Examples:

The positions obtained in a competition, letter: A, B, C, and etc, rating of a company on the basis of the sales experience on a scale 1-10, economic status: low, medium and high etc.

3. Discrete vs Continuous Data

These are the bifurcations of the quantitative data. In almost all the fields the decisions depend on the type of data available i.e., discrete and continuous.

(a) Discrete data

Discrete data mainly consists of integers. It is not possible to subdivide this data into parts.

For example, students in class is an example of discrete data as it cannot be taken as a fraction or decimals.

Discrete data can assume certain values and cannot be further subdivided.

Examples:

Students attending a class, employees of a company, number of questions attempted correctly etc.

(b) Continuous data

Continuous data comprises of the values that can be taken in the form of intervals. It can have any numerical value and it is possible to measure it on a scale or in continuum.

For example, measuring height in the form of precise scales like meters, millimetres or centimetres etc.

Continuous data can be recorded for measurement of various quantities.

The continuous variables may assume any value in between two given numbers. For eg., between 60 and 70 inches, there are literally millions of possible heights: 63.052 inches, 68.9856 inches etc.

Thus, it can be generalised that the given data is continuous if the point of the measurement of given data can be reduced to half and still it makes sense in that case the given data is continuous otherwise it is discrete.

Examples:

Time required for the completion of a project, heights of students, car speed etc.

Before proceeding further, it is necessary to understand the concept of data analytics in detail. It is explained as below:

DATA ANALYTICS-INTRODUCTION

Data analytics is the process of analysing raw data to find out the trends and also answering questions. It consists of various methods with numerous goals. The data analytics is a broad term that comprehends various diverse techniques and processes which facilitate in drawing inferences from the historical data over a specific period of time. Considering the present scenario data analytics is used to define analysis in terms of huge volumes of the given data or even the data that is high-velocity that poses challenges related to data handling as well as its computation. The primary motive of data analytics is to transform the given data into actionable insights. Data analytics is highly beneficial as it helps in cost reduction, faster as well as better decision making which further helps in improving the revenue growth, launching new products as well as services and controlling the unwanted risks. The data analytics process includes components which help in achieving a variety of initiatives. A combination of these initiatives leads to a clear outcome of where one is, where one has been to and where one should go.

In general, this process starts with an analysis which is descriptive in nature. It gives a description of the trends in data in historical manner. Descriptive analytics deals with question "what actually happened?" such a process includes measurement of the indicators of traditional nature that consisting of the return on the investment (ROI). There are different indicators related with every industry. The descriptive analytics does not deal with making predictions or giving decisions. It enables summarising the data in a way that is descriptive as well as meaningful.

After descriptive analytics comes another important part is the analytics that of advanced form known as advanced analytics. This arena of the data science has an advantage as it includes advance tools for data extraction, making forecasts and discovering trends. Such tools are composed of statistics and machine learning. Advanced analytics include Machine learning technologies like neural networks, sentiment analysis, natural language processing etc. This information offers a new vision from the given data. Advanced analytics deals with the question like "what if?"

Due to availability of massive data sets, machine learning methods and economical computing power has permitted the use of such methods in various businesses. Due to the availability of big data set application of such techniques is possible. The big data analytics facilitates industries to derive meaningful inferences from various multifaceted and diverse sources of data. This is all possible due to the advancement in the economical computational power and parallel processing.

The primary motive of data analysis is the conversion of data into actionable insights. The main types of the data analytics are descriptive, diagnostic, predictive and prescriptive. There is a huge difference in their levels of complexity as well as their added values.

Steps of Data Analytics

Figure 1. Steps of Data Analytics

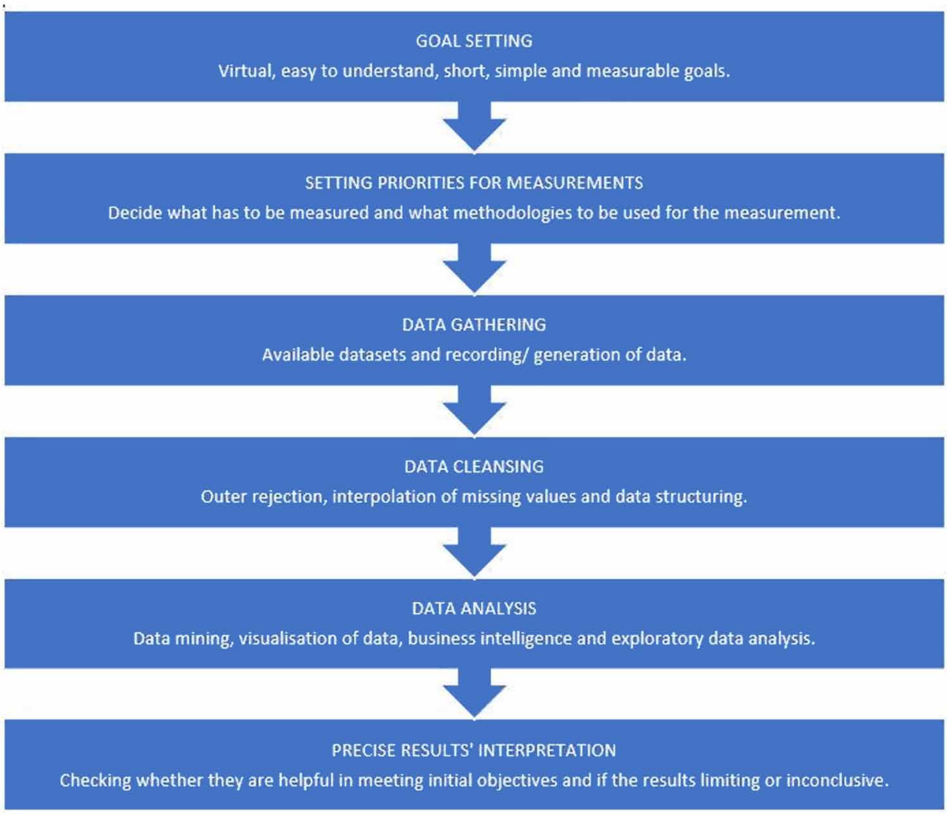


Figure 1 shows the steps involved in Data analytics.

(i) Goal setting: The business enterprises have to frame the purposes related to data analytics. These purposes may be in the form of questions. For example: If the business is facing trouble in selling their product, then the relevant questions can be:

Are the goods overpriced?

What is the difference between competitor's product and our product?

In order to find the answers to these questions the company has to gather the data related to production cost and the price of alike goods available in the market.

(ii) Setting the priorities for measurements: The business company has to determine the type of data required for obtaining an answer to the questions related to the objectives. A timeline has to be set for analysing the project. The units of measurement that are to be used have to be determined.

(iii) Data gathering: The data may be available as data sets. Generation of the data can be done by using the direct or personal interview method, indirect or may be questionnaire method, method of registration and the experimental method.

(iv) Data cleansing: Data cleansing is used to determine incomplete, incorrect, inaccurate as well as the inappropriate parts of the data. In this process the uneven data is modified, replaced or removed. The data cleansing process is described below in the figure 2:

Figure 2. Data cleansing

(v) Data analysis: It is a process of data evaluation by using analytical and logical reasoning, examination of each component of data.

Steps Involved in Data Analysis

1. Pre process: It involves data normalization or scaling. The partitioning is done to train, validate and test the available data. After this process lagged inputs are made.
2. Feature Engineering: This involves extraction of features, reduction in the dimension and also selection of the features.
3. Process: This step involves data mining, application of machine learning algorithms and formation of the AI models. Then the classification is done and regression is applied.
4. Optimization: It involves optimization of the model and also validating it. Then the parameters are tuned accordingly.
5. Performance evaluation: In this step evaluation is done the observed data. For this process evaluation matrix is taken into consideration. The accuracy of the entire process under consideration is also measured.

These processes can be explained as follows:

3. Pre-processing: This stage deals with cleaning of the data. It involves filling in the missing values, identification of the smooth or noisy data identification and removal of the outliers, removal of any inconsistencies. After this there is an integration of multiple databases or files. This process is followed by normalization or scaling of the data along with its aggregation. The data is then obtained in the reduced volume but is gives the same analytical results.
4. Feature engineering: It is a process of transformation of the raw data into a data having features that represent the problem under consideration in a better way related to the predictive models. This results in an improved accuracy on the unseen data. Transformation of the data for creation of model inputs.
5. Feature extraction: It includes reduction in the dimension by principal component analysis, non-negative matrix factorization, kernel principal component analysis, graph-based kernel principal component analysis and generalised discriminant analysis. Then there is a process of data smoothing which includes wavelet transform, kernel smoother, Laplacian smoothing and local regression.
6. Feature selection: This involves identification of features that are irrelevant or redundant. This improves model's interpretability. There are various approaches followed for feature selection, they include wrapper, which is selection through the subset space, training a model of current subset, evaluating it on the held-out data and its iteration. The next one is filter which involves using N features which are the most promising according to their ranking obtained from proxy measure i.e., from mutual information, pearson correlation coefficient, ANOVA

and chi-square. There are also embedded methods in which selecting a feature becomes a part of constructing a model. This is done by LASSO or RIDGE regression. But the process of feature engineering has its own limitations as the performance of the model decreases by addition of many correlated predictors. With the increase in the number of variables interpretability of the model decreases. The models should be generalizable to the other data. Excess of feature engineering also leads to overfitting. There is a cross connection between cross validation as well as feature engineering.

7. Model training: It deals with model construction which gives a description of a set of classes that are predetermined. Each of the tuple belongs to a class that is predefined which is determined by the class label. A training set consists of set of tuples which are used for constructing a model. The representation of the model is done by the help of decision trees, classification rules or mathematical formulae. The basic motive of the models is classification of unknown objects or future.
8. Model optimization: Tuning of the model for error reduction. Optimization of the model parameters by following different approaches. Validation of the model by K-fold cross validation and monte-carlo method.
9. Performance evaluation: This involves the process of model verification as well as the accuracy measures.
10. Result interpretation: This is a very important step. Here verification has to be made regarding the objections that may have been raised in the beginning. It also verifies if the results are inconclusive. If it is so then further research has to be conducted. Interpretation of the results of the analysis is very important for a company to be successful.

Types of Data Analytics

There are four types of data analytics:

1. Descriptive
2. Diagnostic
3. Predictive
4. Prescriptive

1. Descriptive analytics: It answers the question about what happened or what is happening. This technique summarises large sets of data in order to provide a description to stakeholders. With the development of key performance indicators (KPIs) these approaches can help in tracking success or a failure. It facilitates in summarising the large data sets into smaller more useful

nuggets of information and helps in visualizing the outcomes to provide the description of the problems and also the available opportunities. There are numerous ways of summarizing data they are arithmetic mean, frequency, range etc. descriptive analysis deals with categorization and classification of the information like that related to the revenue and expenses, sales tax, costs related with the inventory etc. This further deals with the verification of the large amounts of the data. The data is gathered from various sources. This finally helps in optimizing the entire production process which further shows that the saving potentials remain unused. Metrics like return on the investment (ROI) are used in various industries. For tracking performance in various industries there is a development of specialized metrics. The entire process consists of collection of the pertinent data, data processing, analysis of the data and data visualization. It provides an important insight into past performance.

2. Diagnostic analytics: It deals with questions related to why the things happened. Such techniques complement more basic descriptive type of the analysis. It takes findings from the descriptive analytics and investigates the primary cause behind this. Here the performance indicators are investigated further why they became better or worst. This generally happens in these three steps which are: identification of the differences in data, these can be related to unforeseen changes in a metric or a specific market, the data connected to such anomalies is collected and, in this phase, we are primarily concerned with the how the data is associated to each other. In order to find the root cause a dig deeper into the data approach is required for this purpose a drill-down approach is used. This requires breaking down of the complex data into simpler units, for example from years to months and to weeks. Diagnostic analysis basically deals with monitoring changes in the data, analysing variance, calculating the historical performance, building reasonable forecasts. The diagnostic analysis not only relies on the drill-down analysis. The process of new information in the existing data sets is known as data mining. Along with the drill-down analysis it facilitates in finding the patterns, relationships which is known as correlation between the given variables or the unusual data points which are called the anomalies in the given data set. Finding the patters helps in determining the subgroups, correlation helps in identification of the trends and anomalies helps in determination of the outliers.

3. Predictive analytics: It deals with the questions related to about what is probable to happen in future. Thus, a prediction model can be prepared and it can be continuously refined. This model is built based on the historical data for trend identification and to check if they are likely to recur and also to identify the trends and forecast the potential future outcomes based on this data. This analytics is an assessment of the outcomes related to the future and

identification of the patterns related to the forecasts. These tolls provide valuable vision to what is likely to happen in future. Such techniques consist of a variety of statistical as well as the machine learning methods like neural networks, regression as well as decision trees. A predictive analyst acts as a trusted advisor to the business leaders. For analysis in this step regression analysis is used. Correlation describes if there is any relationship between the two variables whereas regression describes the effect of one variable on the other. It helps to determine the functional relationship between the given variables in order to estimate the unknown variable to make future projections. Considering the present scenario machine learning is used to predict the outcomes. Here the main focus is on the development of the computer programs i.e., algorithms that can access data and can also use it to learn for themselves without any kind of human intervention. Here basically historical data is used for developing the algorithms in order to find the patterns and make decisions based on this by predicting the outcomes.

4. Prescriptive analysis: this is the final phase of analysis. He helps in directing problems about what should be done. It is used to turn predictions into actions by actively providing the next best action plan. By using these insights that are obtained from predictive analytics the decisions related to data may be made. In case of uncertainty the businesses can make the best decisions by using this approach. In case of this prescriptive analytics the prediction model is integrated into daily business. Prescriptive analytics techniques rely on machine learning strategies that can find patterns in large datasets. It is deployed in case of the productive environments like SAP and handed over to the decision makers.

Such data analytics provide an insight that business need to make efficient and effective decisions. When these are used in perfect combinations, they provide a good understanding of company’s needs and opportunities.

The primary work of the data analyst is concerned with working with the available data during the data analysis process. It involves dealing with set of data in numerous ways. The main stages for the data analytics process include data mining, management of data, statistical analysis and presentation of data. The importance as well as the balance linked with these steps is based on the data that is used and also the motive of the analysis.

Data Science

Data science deals with the concept of unification of statistics, data analysis, informatics as well as their related methods for understanding and analysing the phenomenon with the available data. It deals with the theories and concepts related to various fields like Mathematics, statistics, computer science, information science etc. Data science is basically an interdisciplinary field which deals with extraction of knowledge from the large data sets and application of this knowledge for obtaining solutions to problems in various domains. Data science is an essential part in a lot of industries today. Every industry has massive amounts of data that needs to be processed and manual processing this data is not possible. Data science deals with such huge data sets using modern tools and techniques for finding the unseen patterns, deriving meaningful information and making the necessary business decisions. The data that is used for analysis is derived from various sources and is available in different formats.

Data science life cycle consists of the following:

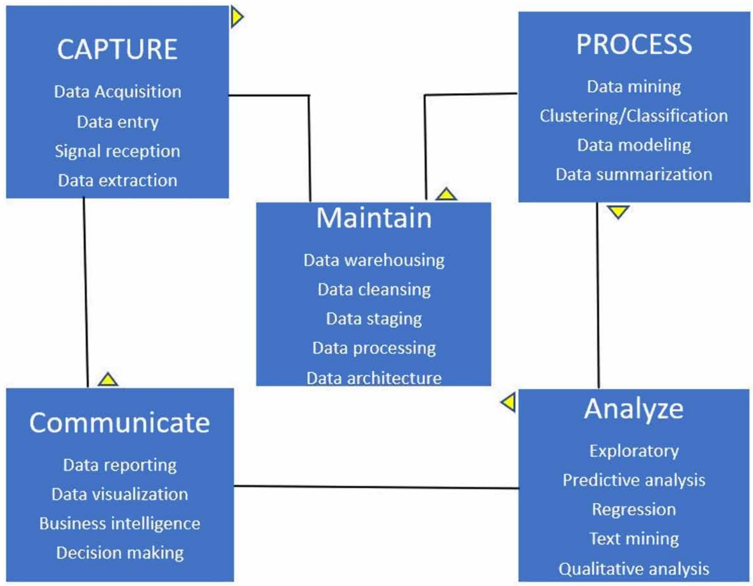
1. Capture
2. Maintain
3. Process
4. Analyse
5. Communicate

These stages can be explained as follows:

1. Capture: This stage consists of data acquisition then entry of the data in requisite format then there is signal reception and extraction of the data. It basically deals with collection of raw structured and unstructured data.
2. Maintain: It includes maintaining the data base in the data warehouse, cleansing of this data, data staging, processing of the data and finally the formation of the data architecture. This stage deals with transforming the raw data into the form so that it can be further used.
3. Process: This stage includes data mining, clustering and classification, data modelling and summarising the data. The data scientists obtain the prepared data and examine the patterns, ranges as well as the biases to analyse its utility in predictive analysis.
4. Analyse: This includes exploratory or confirmatory analysis, predictive analysis, regression analysis, text mining and qualitative analysis of the data. In this stage various analysis of data are performed.
5. Communicate: This stage deals with data reporting, data visualization, business intelligence and decision making. In this final stage the analysis of the data are represented in the form of charts, graphs as well as the reports

This process can be explained with the help of the following figure 3:

Figure 3. Data science life cycle



Data Scientist

Primarily deals with the problem under consideration, finding answers to relevant questions. The data scientist is the one who extracts meaningful insights from the available data by analysing the data. The main task of the data scientist is obtaining the solution of any business problem by following a series of steps listed below:

1. Before collecting and analysing the data, the primary task of the data scientist is determination of the problem by putting up the right questions and obtaining the relevant information.
2. After this the data scientist determines the appropriate variables to be used and also the data set related to the problem.
3. Structured as well as the unstructured data is gathered by the data scientist from various sources.
4. After the collection of the raw data the data scientist processes the raw data and converts it into a suitable format for analysing it. This process involves cleaning and validating the data for maintain the uniformity, accuracy and completeness.

5. Once the data is converted into a usable form, it is used for the analysis purpose by using the statistical model or ML algorithm for analysis and identification of trends and patterns related to the data.
6. The data scientist then carries out the process of interpretation of the data for finding out the solutions.
7. In this step the data scientist prepares the final results as well as the insights and shares it with the appropriate stakeholders and communicates the results.

Data Mining

One of the most important tasks for data analytics is data mining. It is a process of extracting data from unstructured data sources. For the discovery of knowledge, the idea of data mining has been into force since centuries and all this began with manual formulas for statistical modelling as well as regression analysis. Alan Turing in 1930 introduced the idea related to universal computing machine in which computations could be performed. This also led to the ever-expanding explosion of digital information that continues to be there. They include written texts, raw sensor data, or large complex database. The main steps in the process include extraction, transformation and loading data. This is also known as ETL. These steps facilitate in converting raw data into a format that useful and manageable. Data mining is the most time-consuming step in the data analytics pipeline.

Data is becoming an indispensable part of each facet related to business and also life. Considering the present scenario, the companies are utilizing applications related to mining and machine learning for cultivating their sales and also interpretation of financials related to investment purpose. Thus, data scientists are regarded as spine of any organization round the globe as the organizations are in a process of achieving bigger goals than before. Data mining facilitates in analysing enormous volumes of the data for discovering business intelligence which may further render support to the companies for problem solving, mitigating risks and seizing new opportunities.

There are numerous questions related to business which cannot be answered manually as they are too time consuming but can be easily answered by the process of data mining. By the use of powerful computers as well as algorithms for the execution of a variety of statistical techniques for data analysis in diverse ways, with pattern identification, trends and the finding relationships between values. These findings can be applied for the prediction of what is probable to occur in future and take necessary action to influence solutions related to business. The process of data mining is applied in many areas of research and business, together with sales and marketing, development of product, healthcare industry, and education. Data mining has an advantage over competitors as it makes it possible to learn about the consumers, developing effective strategies related to market increasing revenue and thereby decreasing the cost.

Working of Data Mining Process

The data mining plan should start by the establishment of the business question for which answers are required. If there is no clear focus on a meaningful outcome related to business then the same set of data would be repeated again and again without getting any useful information out of it. Once a clarity related to the kind of problem is sought then the next step is collection of the relevant data from numerous sources to a data warehouse where it is prepared for analysis. The rate of success in the process is dependent on the type of data being collected. If the data collected is of poor quality, then it leads to poor results and vice versa. In order to get reliable and timely results through the process of data mining one needs to follow a structured and a repeatable process that includes the following steps:

(i) Business understanding: Development of a proper understanding of the parameters related to the project. This includes primary objective of the project related to the business, the current business situation and the criteria for success.
(ii) Data understanding: Determination of the data which is required for solving the problem and also gathering data from all the sources that are available.
(iii) Data preparation: Getting the data that is available arranged for analysis. It has to be ensured that the data is suitable format for answering the business questions and also for fixing any kind of data quality problems like duplicate or missing data.
(iv) Modelling: Using algorithms for identification of patterns in data and applying these patterns to prognostic model.
(v) Evaluation: Determination of whether and how fine the results that are delivered by the assumed model facilitate in achieving the goals related to the business. This process includes a phase that is iterative in nature having the algorithm that is used is fine tuned for achieving the best results.
(vi) Deployment: Running the analysis in an appropriate manner and making the results related to the project available to the decision creators.

In entire process there is a requirement of a close collaboration between the data miners and the domain experts for understanding the significance related to the data mining outcomes focussed on the questions related to the business.

Advantages of Data Mining

Everyday there is a huge inflow of data from a variety of source into any business in multiple formats, at unprecedented speed as well as volumes. In the present era every business needs to opt for being a data driven business as the success of the business depends on quick assimilation of visions from the big data also their incorporation into the business decisions as well as processes for attainment of better results. Data mining renders support to the business for optimization of operations for future by having a clear understanding of the past as well as present and making accurate predictions about future. Data mining techniques are adopted by the sales and marketing teams for prediction about the prospects which can be converted into profitable customers. This can be done by studying the past customer demographics. It facilitates in finding the potential prospects to a specific offer being served. This helps in increasing the return on investment (ROI) by targeting such prospects who are probable to respond and can be converted to valuable customers. Data mining can be used for solving any type of business problem related to:

1. Increasing the revenue
2. Understanding the segment of customers and their preferences.
3. Acquiring new consumers
4. Improving the technique of cross-selling as well as up-selling
5. Retainment of customers and growing the loyalty
6. Increasing the return on investment by holding marketing campaigns
7. Detection and prevention of fraud
8. Identification of credit risks
9. Monitoring the operational performance

By the using this data mining process, the decisions are related to the real business intelligence instead of gut reactions or instinct. They also deliver unswerving results which keep the business ahead of any competitions. Due to the availability of large-scale processing technologies for data such as artificial intelligence and machine learning the organizations are able to automate the process and dig out terabytes of the data in just a few minutes or even hours rather than days or weeks and this enables them grow fast and innovate.

Data Mining Concepts

For attainment of best results out of data mining requires a range of techniques and tools. Some of them may be familiar one while others might be new. The basic concepts and terms related to data mining are as follows:

1. Data processes: It deals with the management of the data in proper manner.
   (i) Data cleansing and preparation: There is an in-flow of raw data from various sources in a variety of formats and quality. In order to use this data in a meaningful way it has to be converted from a raw state to a type of format that is appropriate for analysis as well as processing. It is made up of processes like identification and removal of errors, finding out the missing data and also flagging outliers.
   (ii) Data warehousing: This process deals with collection of data from variety of sources and combining it into a repository of data that is single in nature before it can be used to make decisions. Such repository is referred to as a warehouse of data. It acts as a foundation to most of the large-scale mining of data.
   (iii) Data analytics: Once the collection of data is done and cleaned it can be examined related to past trends which are further applicable to decision making related to future. This includes the process of evaluation of historical digital information for providing useful business intelligence.
   (iv) Predictive analytics: In this type of data analytics the past trends are identified for anticipating the future consequences. Predictive analytics is based on data modelling, machine learning as well as artificial intelligence to deal with designs related to big data.
2. Computer Science concepts: The basic terms related to Computer Science which describe how several programs and algorithms interrelate with the data for delivering meaningful insights are as given below:
   (i) Artificial Intelligence: With the advancement in technology, analytical activities can be performed with the automated systems which could possibly be performed by using human intelligence. Such activities include planning, reasoning, learning and solving problems. Considering the case of data mining it refers to the use of a computer program for identification of meaningful trends in data.
   (ii) Machine learning: In the past era the computers required a clear program in order to train them step by step through any of the processes with an assumption that the programmer knows about every kind of scenario that may come into existence. In the current time the programmers make the use of statistical probabilities for writing the algorithms related to machine learning that provide computers with the ability for learning as well as adaption without any kind of explicit programming.
   (iii) Natural language processing (NLP): Various treasured data sources like social media cannot be broken down into simpler fields by ease. Natural language processing is a feature related to AI that provides computer

program an ability to read and also understand a type of casual and unstructured sources of data.

(iv) Neural networks: in certain situations, Machine learning alone is not powerful enough for performing the job all alone. A neural network consists of algorithms which together provide solution to the complex problems by thinking the way humans do. Similar to a machine learning algorithm, the machine learning has the ability of adapt and learn.

3. Data mining techniques: The various techniques used by the data mining technology are as follow:

   (i) Association rule learning: It is also called market basket analysis, it deals with interesting relationship between variables of the dataset which is not immediately apparent, similar to determination of products that are typically purchased together. This can be really valuable for a planning of long term.

   (ii) Classification: It is a technique that deals with sorting out objects in dataset into diverse type of categories of targets having some sort of common features. With this the algorithm can neatly categorise data cases that are complex.

   (iii) Clustering: This helps the users to understand the structure or the natural grouping within the data. The process of segmentation of dataset into a set that consists of meaningful sub-classes known as clusters is done. This process forms a group of similar objects rather than forming groups with predetermined features.

   (iv) Decision trees: This is alternative method for categorizing the data. In this method sorting of data into relevant classes is done by putting up a series of questions.

   (v) Regression: On the basis of particular data set the technique of regression is used for the prediction of a range of numerical values related to temperature, sales, stock prices etc.

Data mining has greatly contributed to the success of businesses. These businesses select a platform that is based on the following criteria:

1. It integrates best practices related to all type of projects or the industries like healthcare organizations having different needs as compared to the e-commerce companies.
2. It helps in managing the complete data mining lifecycle, starting from exploration of data to production.
3. It aligns along with all the enterprise applications like BI system, CRM, ERP, financial systems and also other software related to the enterprises.

4. It provides flexibility and collaboration tools to the developers and data scientists for creation of applications which are innovative in nature.
5. It fulfils the demands of data scientists, IT and analysts and also catering to the reporting and visualisation requirements of the business users.

Data Analytics in Financial Reporting

1. Along with the automation process analytics facilitate understanding risk in better manner considering thousands of data points in P & L reporting.
2. Data analytics help in improving the co-operation with the external auditors for detection of patterns and trends and also identification of the process improvement that can increase the efficiency and enhance the quality of the audit.
3. They help in managing the risk across the organization in a better manner and also uncover the valuable insights within the financials.
4. In order to make the internal audit function highly efficient and effective it is important to use the automation of large parts of audit plans.

INTERNET OF THINGS (IoT)

The internet of things is basically the physical objects having sensors, processing ability, software along with other technologies that connect as well as exchange the data with other devices and systems over internet or the other communication network. These devices may include the ordinary household objects and the sophisticated industrial tools. In the 21st century IoT has become an important technology. Due to the ease with which we can connect the appliances, thermostats, cars etc, to the internet with the embedded devices continuous communication is feasible between such devices.

Due to recent advancement in various technologies IoT has come into existence. More manufacturers are finding it easy to make the best possible utilisation of IoT as the sensors are reliable and accessible. Due to the increased availability of the cloud platforms both the consumers as well as the businesses have found an easy access to the infrastructure that they need to scale up without actually managing it. With an increased advancement in machine learning and analytics and also an access to the huge amount of data stored in the cloud, it is very convenient for the business to gather faster and easier insights into everything. With the recent advancement in neural networks there is an introduction of natural language processing (NLP) to IoT devices and due to this such devices have become affordable, appealing and viable for home use.

Advantages of IoT

In manufacturing sector: It facilitates the manufacturers in getting a competitive edge by the use of monitoring of the production line. This enables the maintenance of the equipment in a proactive manner with the help of sensors as they can detect failure of any type in the equipment.

In automotive sector: IoT applications are highly advantageous to the automotive sector. There is a widespread benefit of applying IoT in the production line and also the sensors can detect any kind of failure which is there in the vehicles already on the road and this can actually make the driver alert by providing the details and also giving recommendations. All this is possible by the IoT based applications.

In logistics and transportations: There are a huge benefit of IoT on the logistics and transportation sector. With the availability of information related to the weather conditions, availability of the vehicles, driver availability etc., through the IoT sensors the trucks, ships, fleet of cars which carry the inventory can be rerouted. It is also possible to equip the inventory with sensors for temperature control track and trace monitoring. These temperature monitoring applications are highly beneficial as they send alerts whenever there is a rise or fall in the temperature that may cause harm to the product.

In retail: There are widespread advantages of IoT in the retails sector as it helps in managing the inventory, improving the customer experience, optimization of the supply chain and reduction in the operational cost.

In public sector: There is a huge benefit of IoT in public sector as well as service-related industry. In case of government owned utilities may utilise the IoT based applications for notifying the users about the interruptions in water, sewer or power services. IoT is used for collecting information related to the scarcity of the resources and also help the utilities to cope up with such situations in a speedy manner.

In healthcare: There are numerous benefits of IoT in healthcare industry. With the availability of sensors on the wheelchairs they can be easily tracked using the IoT monitoring applications so that they can be easily located. This is an efficient way for locating various hospital equipment and also ensuring proper usage of the financial accounting.

For general safety across all industries: IoT is used for making improvements in worker safety. The workers working in any kind of hazardous environment need to know in advance about the occurrence of any hazardous event which might be harmful. If they are connected to the Iot sensors-based applications they are informed about the accidents and can be rescued from them. Another application of IoT sensor-based applications is that they help the physicians to monitor the patients remotely.

Digital Twins

A digital twin is the representation of a physical object in the digital mode. It is actually a digital replica of the physical object, for example a jet engine, a building or a city etc. The digital twin technology is also used for the replication of the processes for data collection and prediction of the performance. Due to the advancement in machine learning and also the factors like big data such virtual models have gained widespread popularity for carrying out innovative processes and also improving performance.

The digital twins are built up by the specialists who are experts in data science or in the applied mathematics field. These experts develop mathematical models which stimulate the real world in the form of a digital space. The digital twin is created for receiving inputs from the sensors connected to the real world. This process gives an insight into the potential problems as well as the performance related to different processes. The digital twin also acts as a prototype before the launch of the physical version of the product. A digital twin may be a simple one or even a complicated one depending on the developer. The amount of data that is used for creating and updating it determines the precision with which a physical object is being simulated.

There is a widespread application of digital twins as it is applicable for obtaining solution to challenges across various industries. It facilitates users to determine the most appropriate solutions to problems ranging from prototype testing to product development, manufacturing as well process improvement and product life cycle extension. In all such cases the digital twins represent a problem virtually so that its solution can be framed and tested in the concerned program rather than in the real world.

There are three different types of digital twins and according to the type they are used in different times as listed below:

1. Digital Twin Prototype (DTP): Before the development of the physical product this is created.
2. Digital Twin Instance (DTI): When the manufacturing of the product is done then this is used to run various tests on the product applying different usage situations.
3. Digital Twin Aggregate (DTA): It collects the information related to DTI for determination of the capabilities of the related product, running the prognostics and also testing various parameters.

These different types are used logistics planning, development of product and also its redesigning, quality control as well as management and systematic planning. A digital twin is used for saving time and money whenever there is a requirement for testing the product, design, implementation, product monitoring or improvement.

The uses of digital twins are numerous depending on the conditions and places where they are used. They help in decreasing the defects in products and reducing the time for their launch in market by preparing a prototype before beginning with the actual manufacture of the product. It is also used for improvement of the process, proper alignment of the supply chains along with manufacturing or requirements related to maintenance of the products. This facilitates in increasing the reliability and improving the performance of the products by proper monitoring and simulation. Due to this there is a huge reduction in the unwanted accidents as well as unplanned downtime due to failure, this also reduces the maintenance cost as the failure in products can be predicted before they occur. Digital twins result improvement in the products by customizing models and ensuring the product quality with the help of performance testing.

Challenges in Data Analytics

There are numerous challenges in the data analytics process including deep learning, big data and IoT which are listed as below:

1. Collecting the meaningful data: As huge amounts of data are available it becomes very tough to draw critical insights from such data. Thus, it results in analysis of the data that is readily available and this does not add any value to the existing business. In order to overcome this issue, it is important to understand the type of data that is vital for a business. It can be done by providing a proper training to the employees and giving them the right knowledge in this domain.
2. Selection of the right tool: Numerous tools for analysing the data are already available in the market which creates a confusion about selection of the appropriate tool for analysis. Selection of the right tool is the most important in analysis process because of a wrong tool is selected it results in wastage of money, effort and time. In order to cope up with such a situation one should seek the help of a professional as they have an expertise in the field and can provide the requisite knowledge for selection of the right tool for the analysis process.
3. Consolidation of the data from multiple sources: Data is available in a scattered form through different sources. Combining all the data from different sources is the biggest challenge. It actually aggravates the chances of error thereby making the data unreliable. In order to overcome this problem, the best possible solution is creation of a centralised hub for the data. This will help the analysts

to access all the information from one source thereby reducing the chances of error.

4. Data quality: If incorrect data is collected it leads to the most dangerous outcomes. If the collected data is erroneous then it will result in an unreliable output. Therefor utmost care should be taken during the data collection process for avoiding any sort of unnecessary errors.
5. Data security: Privacy and security of the collected data sets is the most important. If the data is not properly secured it leads to hacking of the data. For overcoming this issue proper training should be provided to the staff members and also cybersecurity professionals should be hired for guarding the data.

Thus, if such challenges are overcome then the society on the whole can be highly benefitted by the process of data analytics.

CONCLUSION

Data Analytics has become an integral part of any IoT related solution. They equip the users with the requisite knowledge for taking smarter personal or business decisions. They also facilitate in generalising potential problem domains without any significant efforts of the user. IoT is primarily dependent on the power as well as the capability related to the data under consideration. Data analytics is primarily concerned about enabling the consumers to make profitable decisions. This chapter provides a complete and a comprehensive review of the related field and also throws light on the various concepts related to the data analytics process.

REFERENCES

Anscombe, F. J. (1973). Graphs in statistic analysis. *The American Statistician*.

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data. John Wiley and Sons*.

Barone, D., Yu, E., Won, J., Jiang, L., & Mylopoulos, J. (2010, November). Enterprise modeling for business intelligence. In *IFIP Working Conference on the Practice of Enterprise Modeling* (pp. 31-45). Springer. 10.1007/978-3-642-16782-9_3

Berry, M. A., & Linoff, G. S. (2000). Mastering data mining: The art and science of customer relationship management. *Industrial Management & Data Systems*, *100*(5), 245–246. doi:10.1108/imds.2000.100.5.245.2

Box, G. E., Hunter, W. H., & Hunter, S. (1978). *Statistics for experimenters* (Vol. 664). John Wiley and Sons.

Chakravarti, I. M., Laha, R. G., & Roy, J. (1967). Handbook of methods of applied statistics. Wiley Series in Probability and Mathematical Statistics (USA) eng.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). John Wiley and Sons.

Giorgini, P., Rizzi, S., & Garzetti, M. (2005, November). Goal-oriented requirement analysis for data warehouse design. In *Proceedings of the 8th ACM international workshop on Data warehousing and OLAP* (pp. 47-56). 10.1145/1097002.1097011

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, *21*(1), 27–58. doi:10.1214/aoms/1177729885

Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., & Mylopoulos, J. (2014). Strategic business modeling: Representation and reasoning. *Software & Systems Modeling*, *13*(3), 1015–1041. doi:10.100710270-012-0290-8

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 21–32.

Mitchell, T. M., & Mitchell, T. M. (1997). Article. *Machine Learning*, *1*(9).

Parnas, D. L., & Madey, J. (1995). Functional documents for computer systems. *Science of Computer Programming*, *25*(1), 41–61. doi:10.1016/0167-6423(95)96871-J

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods, eight edition*. Iowa state University press.

Supakkul, S., & Chung, L. (2009, March). Extending problem frames to deal with stakeholder problems: An agent-and goal-oriented approach. In *Proceedings of the 2009 ACM symposium on Applied Computing* (pp. 389-394). ACM.

Yu, E., Giorgini, P., Maiden, N. A., & Mylopoulos, J. (2011). *Social Modeling for Requirements Engineering: An Introduction*. Academic Press.

Chapter 2
Data Management for IoT and Digital Twin

Galiveeti Poornima
Presidency University, India

Vinay Janardhanachari
Cloud Operations, USA

Deepak S. Sakkari
Presidency Univeristy, India

ABSTRACT

The internet of things (IoT) is a dynamic and global network infrastructure in which "things"—subsystems and individual physical and virtual entities—can be identified, autonomous, and self-configurable. "Things" are expected to communicate with one another and with the environment by exchanging data generated by sensing, as well as react to events and trigger actions to control the physical world. A digital twin is a synchronised virtual representation of real-world entities and processes. Understanding the data management challenges for DT is critical to understanding the data issues. Data management is a common issue in existing systems, ranging from product design to asset management and maintenance.

1. INTRODUCTION

With the expansion of Cyber Physical System (CPS) and internet technology, and also largescale computation and advanced analytics in recent years, the notion of Digital Twin has steadily received widespread interest in smart manufacturing, Napoleone et al., 2020, Negri et al., 2020. With the expansion of Cyber Physical System (CPS)

DOI: 10.4018/978-1-6684-5722-1.ch002

and internet technology, and also largescale computation and advanced analytics in recent years, the notion of Digital Twin has steadily received widespread interest in smart manufacturing, Alam and El Saddik, 2017a. The Digital Twin virtual models use these data to update the simulated physical model in real time and transmit control orders to help physical systems optimize and make decisions. The employment of a digital twin in the design, production, operationyste, and maintenance of a complex system is common. In the product life cycle, Schleich used the Digital Twin model in the design and production phases, as well as model conception, presentation, and implementation concerns, Alam and El Saddik, 2017a. During the operation of the crane, Zhidchenko et al., 2018 created the Digital Twin model to forecast the movement of the mobile crane in real time. Glaessgen and Stargel, 2012, proposed a Digital Twin paradigm as a health management cyber system to ensure the safety and reliability of future NASA and U.S. air force vehicles. Direct process quality measurements are sometimes unavailable or infrequent when modelling physical smart manufacturing sms or products, G̈unther et al., 2016, Yun et al., 2020a. Furthermore, the applicability of currently available engineering tools to smart manufacturing with data-driven controls remains a gap. Model non-convergence can be caused by factors such as a disparity between virtual and physical manufacturing, or out-of-sync communications due to hardware latency, resulting in isolated, fragmented, and sluggish data management, Tao et al., 2018a.

2. OVERVIEW OF IoT

Information and communication technology (ICT) has complete power over our everyday routines and habits. It becomes an important part of our life-critical infrastructure, enabling the connectivity of various heterogeneous devices in various ways. Personal computers, sensing, surveillance, smart homes, entertainment, transportation, and video streaming are just a few examples. It's no secret that the Internet is a constantly evolving organism. As wireless communication trends accelerate, so does innovation in Internet connectivity and mobile broadband. Devices that communicate without relying on physical infrastructure are becoming more common, more intelligent, more powerful, more interconnected, smaller, cheaper, and easier to deploy and set up. There is a new future direction for ICT in society: the Internet of Things (IoT) (IoT). The Internet of Things (IoT), formerly known as Machine-to-Machine (M2M) connectivity, is now a hot topic in the telecommunications industry and academia. The IoT paradigm, its concepts, principles, and prospective benefits are examined in this research. Focused on the primary IoT technologies, developing protocols and wide-spread use cases.

A. Key technologies

Sensors: These are the sorts of devices that can generate electronic signals based on certain physical conditions or happenings. IoT devices feature built-in sensors to see, hear, and touch the world, converting physical data to digital. Images, temperature, motion, proximity, pressure, and so on can all be measured by sensors.

Networks: IoT devices are, in essence, networked ones. Electronic signals can be transmitted over any number of wireless connections, including WiFi, cellular, Bluetooth, near field communication (NFC), and satellite, among others.

Standards: These are the prohibitions or prescriptions for the process framework that are generally recognized and acknowledged. Devices in the Internet of Things (IoT) adhere to common technological and legal standards, ensuring network security, data protection, and other benefits.

Augmented Intelligence: A database can be described, predicted, and exploited using these cognitive techniques. Computer vision, natural language processing, speech recognition and other technologies can be used to analyze massive data in a meaningful way.

Augmented Behavior: This is the process of acting that was prescribed. Augmented behavior shows up in the form of machine-to-machine (M2M) interface and machine-to-human interface (M2H).

B. Artificial Intelligence for IoT

People will wear smart gadgets, eat smart capsules that judge the influence of medicine on the body, and live in smart dwellings. Although it might sound like something out of a science fiction novel, this is actually the subject of the majority of the research being done today. Everything will be able to communicate with one another and will be linked to the internet. It is expected that scientists from all fields will work together to produce something of great value. A "smart cyber revolution" is on the horizon for us. However, there is still dispute about whether we are on the verge of creative disaster.

For example, computers can now perform duties that were previously performed by humans, and this shift is happening during a time when many people are already having a difficult time making ends meet. Nevertheless, we have the ability, with the adoption of appropriate policies, to obtain the most favorable aspects of both automation and widespread unemployment. The function that productive labor plays will inevitably shift as a result of the inventiveness of humans. Educational opportunities will be expanded, and there will be more skilled workers available through re-skilling and up-skilling.

As we continue and implement AI models inside the wild, we would be compelled to reconsider the impact of such automation on the human condition. Although these systems provide a multitude of benefits, they also pose inherent hazards, including privacy breaches, codifying and entrenching biases, lowering accountability and impeding due process, and growing the knowledge asymmetry between data producers and data holders. The Internet of Things Core Protocol Suite is a broad and intricate network. Keeping track of unethical or security breaches will be hard. Any malfunctions or errors caused by faults in the software or hardware will have severe repercussions. Even power outages can be really inconvenient. Therefore, we could need an additional AI system on top of such an AI-enabled IoT in order to track its whereabouts at every instant. It's possible that one day we'll require a democracy comprised of such systems that are capable of preventing themselves from acting in unreasonable ways. Our lives will continue to be governed more and more by technology, and we will become completely dependent on various technological devices. No matter what the circumstances are, humans should still have superiority over all the intelligence that has been created by man. Only then will we be able to control this revolution and prevent ourselves from becoming slaves to it.

C. Blockchain for IoT

Blockchain (BC) in the Internet of Things (IoT) is a revolutionary technology that stores transactions between IoT nodes using a decentralized, distributed, public, and real-time ledger. A blockchain is a collection of blocks, each of which is connected to the ones before it. Every block contains the contents, the cryptographic hash code, as well as the hash of the preceding block. The transactions in BC serve as the fundamental units for transferring data from one node of the IoT network to another. Various physical yet intelligent objects with embedded sensors, actuators, and programming comprise the Internet of Things (IoT). They may interact with one another and with other IoT nodes. The function of the BC in the IoT is to make available a method for the processing of encrypted data recordings via the IoT nodes. Bitcoin Cash is a protected technology that may be utilized in an unrestricted and open manner. This sort of technology is necessary for the Internet of Things (IoT) because it enables safe communication between IoT nodes even when those nodes are located in a diverse environment. The transactions that took place in BC might be tracked and investigated by anyone who is authorized to connect inside the internet of things (IoT). The use of BC in IoT may contribute to an increase in the communication's level of safety.

D. Deep Learning for IoT

Internet-based sensor tools that give real-world observations and data measurements have grown as a result of the rapid development of software, hardware, and the internet. The Internet of Things (IoT) is made up of billions of smart things that can talk to each other. This makes it possible for more physical and virtual things to exist in the world. Every day, these intelligent machines generate or gather enormous amounts of data that may be used in a variety of ways. The use of analytics on such massive amounts of data is a crucial tool for gaining new information, predicting what will be known in the future, and determining how to implement control measures, all of which contribute to the Internet of Things becoming a valuable business paradigm and improving technology. IoT and mobile app initiatives have shown promising outcomes from the application of deep learning. Deep learning may be able to give state-of-the-art solutions for IoT intrusion detection thanks to its data-driven, anomaly-based technique and its ability to spot assaults that are still in the planning stages. Many academics have been drawn to the many subfields that fall under the umbrella of the Internet of Things (IoT), and methods from both DL and IoT have been investigated. Different research indicated DL as a method to manage IoT data since it was designed to handle vast volumes of varied data in near real-time. It is possible for the features, applications, and problems that DL employs to empower IoT applications to stimulate and inspire future research.

E. FoG Computing for IoT

In today's world, the Internet is connecting to a growing number of sensors and gadgets that are part of the Internet of Things (IoT). It's expected that 50 billion devices will be connected to the Internet by 2020. Technological hurdles abound in a world economy powered by Internet of Things (IoT). Since the computational and storage power of IoT devices is quite limited, it is necessary to rely on the support of expanded resources from the Cloud. This leads to the networks, which already have limited bandwidth, becoming overloaded. Servers placed in Cloud data centers that are close to main networks, but far from the edge or field locations where IoT devices receive data, are critical to the functioning of Cloud services. This causes the end-to-end latency to climb up to two orders of magnitude above the required levels for delivering the desired performance of Cloud services. This problem of latency is further worsened by the unprecedentedly massive amounts of data created in real time by the devices that are connected to the internet of things (IoT). This data could be of immense value if it were analyzed in real time, but it is not currently being done so. Data quantities created overload storage systems and analytical applications if transmitted to the Cloud. The latency in transmitting data to the Cloud and receiving analysis findings makes it impossible to meet the demands of real-world circumstances such as emergency response and health management, necessitating the development of a new paradigm known as Fog Computing. To meet the requirements of the widest possible range of Fog Computing applications in fields as diverse as health and transportation as well as smart cities and smart villages as well as smart homes, different definitions of the Fog architecture are available. The field of fog analytics is analyzed in great depth, and forthcoming models and solutions like Smart Data and Fog Engine are offered.

F. Software Defined Network (SDN) for IoT

The goal of the Internet of Things is to establish connections between physical items and the internet. Software-defined networking (SDN) enables orchestration for the administration of networks by divorcing the control plane from the data plane. The number of things that are connected is in the billions, making the administration and control of these connections a difficult issue for big dispersed networks. The SDN enables IoT networks to be more adaptable and customizable without disrupting the underlying architecture. A comparison of SDN-based IoT systems reveals developing patterns. The new era of the Internet of Things is causing a shift in the manner in which humans and machines communicate with one another. Now, people are beginning to consider beyond the connectedness of every physical thing with the internet. IoT is young and lacks programmability, agility, security, and data management to suit consumer needs. Programmability, centralized control for IoT management, and SDN integration are needed.

G. Container Technology for IoT

The variety and heterogeneity of the hardware and software components that make up the Internet of Things (IoT) distinguish it from other types of computing. The process of installing and upgrading physical components and the software that corresponds to them becomes more difficult as a result of this load. Also, these devices are constantly making data and sending it to central cloud platforms for advanced processing. Since billions of devices are expected to be connected in the coming years, this will affect how well the network works. Recently, there has been a trend toward processing these data locally, in close proximity to their sources on the network edge, in order to relieve this worry. However, edge devices typically have limited access to resources, which, in some instances, inhibits their capacity to carry out processing duties. Containers, a lightweight virtualization technique, are going to be used to implement the suggested solution. It does this by utilizing a few concepts associated with containers, most notably the orchestration concept, to enable task forwarding and resource sharing between IoT gateway devices that are part of the same cluster.

3. OVERVIEW OF DIGITAL TWIN

Everyone has come across the digitalization process at some point in their lives. Books have evolved into eBooks. MP3 records have replaced vinyl recordings. Digital twinning has a similar but slightly different theory. Artificial intelligence, the internet of things (IoT), and software analytics are all used to connect physical assets and processes in digital twin technologies. Sensors are used to collect data in physical items, devices, processes, and equipment. This data is sent to the digital world via a cloud-based technology.

A. Definition

The term DT, which has become more widely used in recent years, is frequently misunderstood to mean simply 3D depiction of the real world. Some of the definitions found in the some fo the references are listed in TABLE 1. According to the Defense Acquisition University's (DAU) glossary, DT is an "integrated multi-physics, multi-scale, probabilistic simulation of an as-built system, enabled by digital thread, that uses the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its corresponding physical twin", Wanasinghe et al., 2020.This definition is based on the digital thread concept which is defined as an "extensible, configurable and component enterprise-level analytical framework that seamlessly expedites the controlled interplay of authoritative technical data, software, information, and knowledge in the enterprise data-information-knowledge systems, based on the digital system model template, to inform decision makers throughout a system's life cycle by providing the capability to access, integrate and transform disparate data into actionable information", Grieves, 2014.

B. Framework

Digital Twin has multiple frameworks. Three essential components make up the most frequently acknowledged framework: physical space, virtual space, and connections as indicated in FIGURE 1, Grieves, 2014. The physical asset, sensors, and actuators are in the physical space, while multi-physics, multiscale, probabilistic simulation models aggregate and analyze data and run simulations to discover the best control parameters and conditions for the physical asset are in the virtual space. Data and actuators instructions (driving) are exchanged seamlessly between the physical and

Table 1. Sample definitions for "Digital Twin".

Reference	Definition
Mayani et al., 2018	"refers to the digital foot print of the physical systems in the various assets which act like a bridge between physics and digital world".
Nadhan et al., 2018	"digital copy of the physical systems and act as a connection between physics and digital world".
Poddar, 2018	"virtual and simulated model or a true replica of a physical asset".
Mohr, 2018	"an immersive data analytic technology that provides insights on human-infrastructure-machine interactions to enable execution to make contextual decisions".
Saini et al., 2018	"a virtual physics and data-based model of a system or an asset that models all the various subsystems, their properties, the interaction among them and the interactions of the system with the environment".
Sharma et al., 2018	"a virtual model of a physical asset".

virtual environments thanks to the connections between them. Physical space, virtual space, DT data fusion module, service systems, and connection/interaction between these four modules were added to the three-component DT architecture Prices, 70. In the five-component framework, the physical asset, as well as sensors and actuators, are housed in the physical space. The digital mirror for high-fidelity replication of the physical equivalent is the virtual space. Visualization services, product quality services, diagnostic services, model calibration services, algorithm services, and numerous data services are all included in the service system. The DT data fusion paradigm serves as a link between physical, virtual, and service systems. This module gathers information from sensors (in the physical world), simulation (in the virtual world), and the service system. The DT data fusion model fuses and analyses the acquired data to provide driving commands for the remaining three modules.

Deloitte recently released a technical study that outlines a DT framework with five enabling components and a six-step approach, Parrott and Warshaw, 2017, as shown in Figure 2.

C. Classification

Digital Twins can be characterized in three ways, according to IoT Analytics' research. The hierarchical level at which the Digital Twin is used (6 levels identified). The phase of the Digital Twin's lifespan in which it is utilized (6 phases identified). The Digital Twin's Application (7 most common uses identified).

Figure 1. Digital Twin Framework.

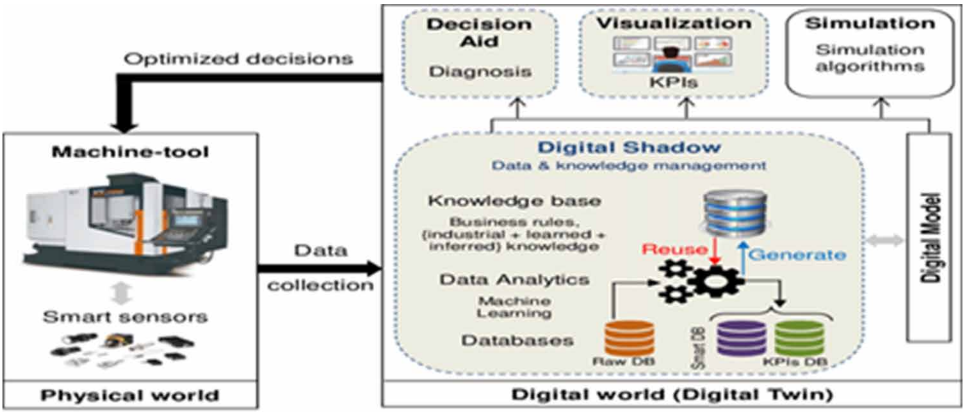
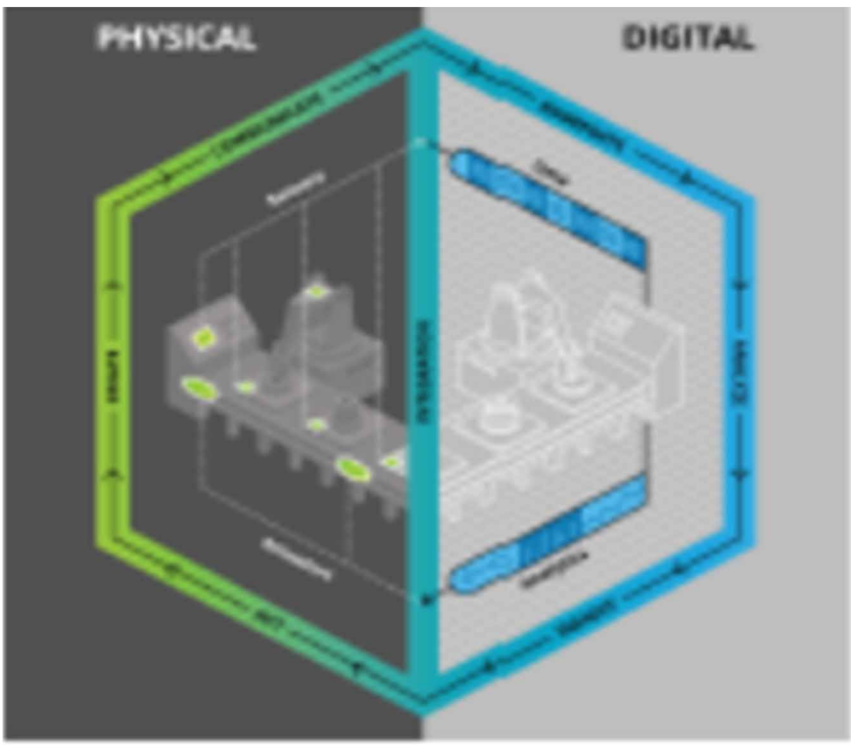


Figure 2. Manufacturing process digital twin model. Although the model was created for the manufacturing industry, it can be used to any industry.



i) Hierarchical Level:

1. **Informational**: Digital representations of information e.g., an operation manual in digital format.
2. **Component**: Individual components or elements of a real thing are represented digitally, such as a bolt or a bearing in a robotic arm.
3. **Product**: Digital representations of component/part interoperability as they interact at the product level, such as a virtual representation of a revolving robotic arm.
4. **Process**: Virtual representations of a manufacturing production line process, for example, enable the operation and maintenance of whole fleets of different goods that work together to deliver a result at a process level.
5. **System**: Digital representations of numerous processes and workflows, not just physical things, allow for system-level optimization of activities, such as a virtual depiction of an entire manufacturing system.
6. **Multi-system**: Virtual depictions of multiple systems working in tandem, such as industrial manufacturing, supply chain, traffic monitoring, communication, HR, and so on, to offer unprecedented insight, testing, and monitoring of important business KPIs in a data-driven way.

ii) Lifecycle Phases:

1. **Design**: In the design phase, requirements are obtained and one or more designs (e.g., for components, products, processes, or systems) are created that appear to achieve the desired end, such as using digital twins as the source of all information such as object attributes and parameter values on which virtual representations can be built.
2. **Build**: The construction phase employs the previously specified code requirements to create the actual software-based digital twin. This phase also includes data management, configuration provisioning, repository management, and reporting, for example, employing digital twins to virtually build and mimic prototypes instead of building more expensive physical counterparts for testing.
3. **Operate**: Actual users begin to employ online digital twins in active deployments during the operation phase. Extracting sensor data or managing devices remotely are common operating activities, for example, employing digital twins to extract real-time sensor data from a rotating robotic arm in a production line or updating device configuration over-the-air.
4. **Maintain**: To ensure its operating performance, the maintenance phase entails making changes to hardware, software, and documentation. It can include making modifications to fix issues, improve security, or meet user needs, such

as employing digital twins to perform routine maintenance chores like issuing OTA updates for system configuration or cybersecurity.

5. **Optimize**: The optimise phase necessitates the utilisation of existing capacity data and a statistical tolerancing methodology. This can be used to improve the creation of detailed design aspects, anticipate performance, and optimise operations, for example, by utilising digital twins to run a huge number of tests that generate insights about future performance and failures.
6. **Decommission**: Decommissioning is the process of removing a digital twin release from service. This is also known as system retirement or system sunsettin1. g. For example, employing digital twins to remotely decommission equipment that are no longer in use and then retiring the use of the related digital twin.

iii) Most Common Uses:

1. **Digitize**: Any digitized information
2. **Visualize**: Basic digital representation of a physical object
3. **Simulate**: Simulation model of a physical system in its environment.
4. **Emulate**: Emulation model of the physical system with real software.
5. **Extract**: Extraction model of real-time data streams, physical to virtual system.
6. **Orchestrate**: Orchestration model for virtual control/updating of physical devices.
7. **Predict**: Prediction model to predict future behaviour of the physical system.

4. FUTURE RESEARCH DIRECTIONS

The potential for further development of digital twins is vast in the future. They have been implemented in 15 different areas, including aerospace, intelligent manufacturing, intelligent cities, construction, medical, robotics, ships, vehicles, rail transportation, industrial engineering, agriculture, mining, energy, and the environment.

A. AI for Digital Twin (DT)

AI and DT technologies have expanded significantly in the last several years, with both academics and industry considering them crucial facilitators of Industry 4.0. In order to create a digital representation of a real-world object, DT relies on the infrastructure and data, the core algorithm and model, and the software and service. The foundation of DT and AI in industrial sectors is even more reliant on the systematic and comprehensive incorporation of domain-specific expertise.

Human civilization has been reshaped and the ecology and climate have been adversely affected by traditional profit-maximizing industrial technology. Sustainable development strategies are getting more attention, but they still have a long way to go before they are fully implemented. This is because all of the FESG dimensions need to be balanced. According to a research from the WBA Tooling Academy Aachen, conventional evaluation for manufacturing enterprises focusing on value added and equity ratio might reveal a much poorer performance with ESG criteria in the balance sheet. Digitalization and AI are powerful tools for sustainable thinking and actions, which are being talked about and pushed for more and more around the world. When all data from the complete value chain is continuously linked, analyzed, and connected to a digital twin or shadow, it is possible to obtain a holistic evaluation of sustainability, specifically, the consideration of ESG criteria alongside profitability (F-factor). Understanding this is important for Industry 4.0 because it enables a demand-oriented and real-time competent analytical basis, which takes into consideration FESG elements along the product lifetime. These new problems are being brought about by climate-neutral goods and manufacturing. AI-driven DTs should give manufacturing transparency and knowledge that enables a demand-oriented and real-time analytical basis, considering FESG elements along the product lifecycle.

B. BC for Digital Twin (DT)

As computer, storage, communications and networking technologies have advanced rapidly, Digital Twins may now be created (DTs). A DT is a digital version of a physical component, product, or piece of equipment. Prior to the fabrication of the actual component, a DT can be used for three-dimensional design, as well as testing, simulation, and prototyping. After a physical component has been put into service, a DT can be used to configure, monitor, diagnose, and prognosticate problems with the component. DTs are projected to acquire substantial traction in the near future and play an important role in Industry 4.0. But today’s DT-building methodologies, systems, and technologies are mostly centralised and fall short of delivering reliable data provenance, auditability and traceability. Additionally, the data associated with transactions, logs, and history are not protected from being altered or stolen in any way. To ensure that transactions, logs, and data provenance can be tracked securely and trustedly, a blockchain-based development procedure can be developed for DTs.

C. Deep Learning for Digital Twin (DT)

An increasing number of companies are using digital twin technology to help them bridge the gap between their real and virtual worlds. Grand View Research, Inc.

says that the global market for DT will grow by 38.2% per year until 2025, when it will be worth $26.07 billion. CPS, the Internet of Things, big data analytics, and cloud computing in the manufacturing industry have prepared the way for low-cost and systematic application of DT, with promising benefits on (a) product design and development, (b) machine and equipment health monitoring, and (c) product support and services. If DT were successfully implemented, it would lead to increases in production speed, scalability, transparency, collaboration, adaptability, and resilience, as well as manufacturing efficiency. Realization of smart manufacturing needs collaborative and autonomous interactions between sensing, networking, and computational resources across industrial assets, where data collected from physical systems is used to extract actionable insights and provide predictive services.

D. FoG Learning for Digital Twin (DT)

Digital Twin systems are two interconnected mirrored worlds, one real and one virtual, each mirroring the other, sharing information, and generating predictions based on analysis and simulations. Real-time Digital Twins aren't only logically accurate; they also have to adhere to strict time limits. Modern large-scale Digital Twin implementations use cloud-based infrastructures to store and analyse massive volumes of data. Moving data from the edge to the cloud slows down a Digital Twin's reaction time. Therefore, establishing Digital Twins by making use of cloud-fog architectures appears as a viable solution to bring processing capacity closer to the edge of the network, hence lowering latency and allowing for quicker reaction times. Based on a realistic development and deployment of Digital Twin software components, the distribution of Digital Twins in a fog computing configuration can minimize reaction times, satisfying real-time application requirements.

E. SDN for Digital Twin

Managing essential applications in an industrial environment necessitates the creation of a very complicated system. It is challenging to manage such networks and maintain control over all of the variables that have an influence on their operation for the entirety of their existence. Digital Twinning is one of the top ten most promising technical innovations for the next decade, utilized to enhance industrial system performance. Many Digital Twins of industrial systems exist today, but only a handful are network-ready. Development of a holistic digital twinning architecture for the IIoT is required in which the network is integrated along with the other industrial components of the system. A closed-loop network management over the whole network lifespan, from design to service, can be a primary goal of the research.

The SDN controller's primary function is to enable the establishment of connections between every single Digital Twin of an industrial system and its corresponding physical counterpart. Proposed architecture can be tested when determining the best communication mechanism to meet the real-time needs of a Flexible Production System.

CONCLUSION

Digital twins are a potential technology for digital transformation. As a result, it has to undergo constant development in order to maintain its relevance in both business and academics. As information technology matures and the global need for intelligent manufacturing rises, digital twin technology will become increasingly applicable to manufacturing processes. They serve as a bridge between the real and virtual worlds. Digital twins might also help solve societal issues, which would be a huge contribution to the discipline.

The current state of technology advancement throughout the world places certain restrictions on the options now available. However, in the past several years, a trend toward virtual reality has been set off in the industry of the internet. The notion of the Metaverse has also been postulated, and immersive virtual reality games have become the first application scenario for this technology. The idea of a digital twin was the seed that grew into what we now know as the Metaverse. The digital twin should fulfil the demands of the rapid and dynamic change of human society and economy in the next period in order to contribute to the realization of a smart society and serve as an important instrument for the process of digital transformation.

Some of the obstacles associated with the adoption of digital twins are technological in nature and can be overcome via ongoing research and development. Other obstacles are cultural, and they demand a subversion of the operational models and methods of thinking that are now in place. The notion of digital twins is continually expanding, as new sectors and use cases show. This ongoing conceptual development is a given despite the fact that there are currently insufficient real-world examples to illustrate the benefits that may be derived from employing digital twins. Because of this, we may confidently assume that digital twin technology will have a significant impact on the social sciences.

REFERENCES

Adamson, G., Wang, L., & Moore, P. (2017). Feature-based control and information framework for adaptive and distributed manufacturing in cyber physical systems. *Journal of Manufacturing Systems*, *43*, 305–315. doi:10.1016/j.jmsy.2016.12.003

Alam, K. M., & El Saddik, A. (2017a). C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access: Practical Innovations, Open Solutions*, *5*, 2050–2062. doi:10.1109/ACCESS.2017.2657006

Alam, K. M., & El Saddik, A. (2017b). C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access: Practical Innovations, Open Solutions*, *5*, 2050–2062. doi:10.1109/ACCESS.2017.2657006

Alkan, B., & Harrison, R. (2019). A virtual engineering-based approach to verify structural complexity of component-based automation systems in early design phase. *Journal of Manufacturing Systems*, *53*, 18–31. doi:10.1016/j.jmsy.2019.09.001

Chen, L., Xu, G., Zhang, S., Yan, W., & Wu, Q. (2020). Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks. *Journal of Manufacturing Systems*, *54*, 1–11. doi:10.1016/j.jmsy.2019.11.008

G̈unther, J., Pilarski, P. M., Helfrich, G., Shen, H., & Diepold, K. (2016). Intelligent laser welding through representation, prediction, and control learning: An architecture with deep neural networks and reinforcement learning. *Mechatronics, 34*, 1–11.

Glaessgen, E., & Stargel, D. (2012). The digital twin paradigm for future nasa and us air force vehicles. *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, 1818.

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper, 1*(2014), 1–7.

Gupta, S., & Godavarti, R. (2020). Iot data management using cloud computing and big data technologies. *International Journal of Software Innovation*, *8*(4), 50–58. doi:10.4018/IJSI.2020100104

Hu, L., Liu, Z., Hu, W., Wang, Y., Tan, J., & Wu, F. (2020). Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network. *Journal of Manufacturing Systems*, *55*, 1–14. doi:10.1016/j.jmsy.2020.02.004

Kochunas, B., & Huan, X. (2021). Digital twin concepts with uncertainty for nuclear power applications. *Energies*, *14*(14), 4235. doi:10.3390/en14144235

Kucukoglu, I., Atici-Ulusu, H., Gunduz, T., & Tokcalar, O. (2018). Application of the artificial neural network method to detect defective assembling processes by using a wearable technology. *Journal of Manufacturing Systems*, *49*, 163–171. doi:10.1016/j.jmsy.2018.10.001

Lai, Z.-H., Tao, W., Leu, M. C., & Yin, Z. (2020). Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing. *Journal of Manufacturing Systems*, *55*, 69–81. doi:10.1016/j.jmsy.2020.02.010

Liu, C., Vengayil, H., Lu, Y., & Xu, X. (2019). A cyber-physical machine tools platform using opc ua and mtconnect. *Journal of Manufacturing Systems*, *51*, 61–74. doi:10.1016/j.jmsy.2019.04.006

Lu, Y., & Asghar, M. R. (2020). Semantic communications between distributed cyber-physical systems towards collaborative automation for smart manufacturing. *Journal of Manufacturing Systems*, *55*, 348–359. doi:10.1016/j.jmsy.2020.05.001

Mayani, M. G., Svendsen, M., & Oedegaard, S. (2018). Drilling digital twin success stories the last 10 years. In *SPE Norway One Day Seminar*. OnePetro. 10.2118/191336-MS

Moghaddam, M., Cadavid, M. N., Kenley, C. R., & Deshmukh, A. V. (2018). Reference architectures for smart manufacturing: A critical review. *Journal of Manufacturing Systems*, *49*, 215–225. doi:10.1016/j.jmsy.2018.10.006

Mohr, J. (2018). *Digital twins for the oil and gas industry. Hashplay*. Tech. Rep.

Nadhan, D., Mayani, M. G., & Rommetveit, R. (2018). Drilling with digital twins. In *IADC/SPE Asia pacific drilling technology conference and exhibition*. OnePetro.

Napoleone, A., Macchi, M., & Pozzetti, A. (2020). A review on the characteristics of cyber-physical systems for the future smart factories. *Journal of Manufacturing Systems*, *54*, 305–335. doi:10.1016/j.jmsy.2020.01.007

Negri, E., Berardi, S., Fumagalli, L., & Macchi, M. (2020). Mes-integrated digital twin frameworks. *Journal of Manufacturing Systems*, *56*, 58–71. doi:10.1016/j.jmsy.2020.05.007

Padiya, T., Bhise, M., & Rajkotiya, P. (2015). Data management for internet of things. In *2015 IEEE Region 10 Symposium* (pp. 62–65). IEEE. 10.1109/TENSYMP.2015.26

Park, H.-S., & Tran, N.-H. (2012). An autonomous manufacturing system based on swarm of cognitive agents. *Journal of Manufacturing Systems*, *31*(3), 337–348. doi:10.1016/j.jmsy.2012.05.002

Parrott, A., & Warshaw, L. (2017). *Industry 4.0 and the digital twin*. Deloitte Insights.

Poddar, T. (2018). Digital twin bridging intelligence among man, machine and environment. In *Offshore Technology Conference Asia*. OnePetro.

Prices, C. O. (2017). *Year historical chart.* http://www.macrotrends.net/1369/crude-oil-price-history-chart

Ren, L., Sun, Y., Cui, J., & Zhang, L. (2018). Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *Journal of Manufacturing Systems*, *48*, 71–77. doi:10.1016/j.jmsy.2018.04.008

S̈anchez, F., & Hartlieb, P. (2020). Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation. *Mining, Metallurgy & Exploration, 37*(5), 1385–1399.

Saini, G., Ashok, P., van Oort, E., & Isbell, M. R. (2018). Accelerating well construction using a digital twin demonstrated on unconventional well data in north america. In *Unconventional Resources Technology Conference*, *Houston, Texas*, 23-25 July 2018 (pp. 3264–3276). Society of Exploration Geophysicists, American Association of Petroleum. 10.15530/urtec-2018-2902186

Sharma, P., Knezevic, D., Huynh, P., & Malinowski, G. (2018). Rb-fea based digital twin for structural integrity assessment of offshore structures. In *Offshore Technology Conference*. OnePetro. 10.4043/29005-MS

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018a). Digital twin-driven product design, manufacturing and service with big data. *International Journal of Advanced Manufacturing Technology*, *94*(9), 3563–3576. doi:10.100700170-017-0233-1

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. (2018b). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405–2415. doi:10.1109/TII.2018.2873186

Tuptuk, N., & Hailes, S. (2018). Security of smart manufacturing systems. *Journal of Manufacturing Systems*, *47*, 93–106. doi:10.1016/j.jmsy.2018.04.007

Uhlemann, T. H.-J., Lehmann, C., & Steinhilper, R. (2017). The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia CIRP*, *61*, 335–340. doi:10.1016/j.procir.2016.11.152

Vongsingthong, S., & Smanchat, S. (2015). A review of data management in internet of things. *Asia-Pacific Journal of Science and Technology*, *20*(2), 215–240.

Wanasinghe, T. R., Wroblewski, L., Petersen, B. K., Gosine, R. G., James, L. A., De Silva, O., Mann, G. K., & Warrian, P. J. (2020). Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 104175–104197. doi:10.1109/ACCESS.2020.2998723

Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020a). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, *55*, 317–324. doi:10.1016/j.jmsy.2020.03.009

Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020b). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, *55*, 317–324. doi:10.1016/j.jmsy.2020.03.009

Zhidchenko, V., Malysheva, I., Handroos, H., & Kovartsev, A. (2018). Faster than real-time simulation of mobile crane dynamics using digital twin concept. *Journal of Physics: Conference Series*, *1096*, 012071. doi:10.1088/1742-6596/1096/1/012071

Zhou, M., Yan, J., & Feng, D. (2019). Digital twin framework and its application to power grid online analysis. *CSEE Journal of Power and Energy Systems*, *5*(3), 391–398.

Chapter 3
Cybertwin-Driven Resource Provisioning for IoE Applications at 6G-Enabled Edge Networks

Hemapriya K. E.
https://orcid.org/0000-0003-2184-6286
Sri Krishna Arts and Science College, India

Saraswathi S.
Sri Krishna Arts and Science College, India

ABSTRACT

6G is the latest in wireless communications network technologies supportive for cellular data networks. 6G networks use complex frequencies unlike 5G networks and will empower higher data rates to be achieved and for the 6G network to have a superior global volume. Lower latency levels will almost definitely be a requirement. 6G radio networks will deliver the communication and data congregation essential to accrue data. However, a systems method is mandatory for the 6G technology. It will include data analytics, AI, and next-generation computation abilities using HPC and significant computation.

DOI: 10.4018/978-1-6684-5722-1.ch003

1. INTRODUCTION

Recent advances in the Internet of Things (IoT) and sensing devices have accelerated the rise of data-intensive applications across a wide range of application domains, including online gaming, augmented/virtual reality, path navigation, autonomous driving, video streaming, and so on. The pervasiveness of IoT devices connects millions of physical items, allowing for seamless data collection and sharing via improved wireless technologies. The number of IoT devices is growing, and according to a study undertaken by the International Telecommunication Union (ITU), 125 billion IoT devices will be part of our digital eco-system by 2030, creating about 4395 Exabytes of data. 6G technology was introduced in response to increased traffic from various data-intensive applications and for seamless data transfer among faraway computer devices. 6G is a next generation of wireless technology (Adhikari et al 2021).

In recent years, millions of IoT devices and apps have been able to connect to the internet thanks to advancements in network architecture. A lot of data is produced by gadgets like smartphones, RFID tags, smart refrigerators, smartwatches, smart fire alarms, and others (Khan et al 2020). These IoT devices generate enormous volumes of data, necessitating tremendous computational, storage, and processing capabilities. The construction of centralised cloud data centres has made all of these services viable.

Data transport between remote cloud data centres and IoT applications with a variety of service requirements is also made easier with the latest 6G network technology. However, because cloud data centres are located far away from IoT devices, it poses new issues (location unawareness, security, privacy, trustworthiness, and increased network congestion), making it unsuitable for delay-sensitive applications. Cisco invented the phrase fog/edge computing in 2012 to address the limitations of cloud computing. Edge computing reduces the latency of user requests by bringing communication, processing, and caching capabilities to the network edge (Jain et al 2021). Furthermore, the edge-cloud collaborative architecture makes advantage of the features to improve user experience and enhance QoS for delay-sensitive applications.

On the other hand, the limited capabilities of the edge-cloud cooperation framework cannot keep up with the changing resource needs of data-intensive and latency-sensitive IoT applications. Emerging strongly is a new generation of technology called Cybertwin for a 6G-enabled future communication network. Cybertwin is a term used to describe the digital representation of real-time phenomena in virtual space, such as tangible objects or things. It is a notion for next-generation communication systems(Dhiman et al 2022). The underlying network architecture offers scalability, security, accessibility, and resilience when combined with Cybertwin's features.

6G is the sixth generation presently further down progress for wireless communications network technologies supportive for cellular data networks. 6G networks are determination to use complex frequencies than 5G networks and it will empower higher data rates to stand achieved and for the 6G network to have a much superior global volume. A much lower latency levels will almost definitely be a requirement.

How Will 6G Architecture Look Like?

It is no doubt we may witness ultra-high speed and power that would entice customers for 6G services, but what would be the 6G architecture and components that will pave the way on creating robust and resilient 6G network. Connectivity and automation with advanced versions can assist in development of such networks predicted in previous discussions. It is also suggested to overcome shortcomings of current networks, improve network design and optimise network(Qadir et al 2022). 6G's quicker lightning-speed will be used by advanced AI systems and powerful edge computing to cooperate with complicated systems and provide seamless Internet connectivity.

Who Are Working on 6G?

- Many organisations and universities are showing interest and involving themselves in researching the new technology as given below:
- University of Aveiro released a whitepaper on 2019 'Why 6G?' which discusses the driving forces behind the development of new network like 6G, what are the latest features and key technologies that can be expected.
- Samsung is also keen in participating in the research race as it commenced 6G research in 2019 in the month of June.
- SK Telecom, a South Korean telecom organisation has signed agreements with Ericsson, Samsung and Nokia to conduct research and development in 6G mobile network technology jointly.
- TeraView, terahertz test equipment manufacturing organisation was recently supported with £191 million funds from the Sustainable Innovation Fund with Innovate UK, an innovation agency based in United Kingdom. This is considered to be a crucial step forward in making 6G a reality. With a focus on using its expertise and intellectual property, TeraView will help in building the blocks for future 6G network and accelerate its development.
- Google and Apple have expressed their interest in 6G research and joined the Next G Alliance, that was built in October 2020 to create a 6G roadmap

and motivate North American companies to prepare themselves and be at the forefront of adopting 6G across the globe.

- Korean MNC, LG Electronics has stepped up to develop 6G technology with the establishment of research centre. Company's CTO Park II-pyeong mentioned to support R&D for next-generation 6G network and lead the role with global standardisation and creation of new business opportunities.

Differences Between 5G and 6G Network

Table 1.

	5G	6G
Use of different spectrum	Low band frequencies	High band frequencies
Faster than 5G technology	Low performance and speed	Higher performance and expected to be 100 times faster than 5G
Low latency in both G's	5G networks had ten times lower latency than 4G	Lowering latency to five times than that of fifth-generation network making massive data transmissions possible in less than a second.
6G wireless accelerates IoT after 5G	Lakh of connecting devices	Increase in number of connected devices

6G radio networks

6G radio networks will provide the communication and data gathering capabilities required to collect data. For 6G technology, however, a systems method is required. It will contain capabilities such as data analytics, artificial intelligence, and next-generation computation leveraging HPC and substantial computation.

6G networks resolve considerably more information than 5G networks, and computation willpower evolve to contain management among edge and central platforms. In response to those two deviations, data centres will have to evolve. 6G capabilities in detection, imagery and position determination will produce huge amounts of statistics that must be accomplished on behalf of the network owners, service providers and data owners.

Edge Computation

Edge computation refers to the communication of IoT models with cloud models for the execution of analytical data. Nodes that can be accessible by network devices in order for them to interact with each other. Edge devices assemble, categorise, and analyse data before storing it in the appropriate location. The data calculation takes place at the device's edge, or at the point where data is generated. It ostensibly improves data processing performance by transferring data rootless between policies and centralised cloud data centres. For a long time, federal cloud computing was stable, and it has transformed the way people read, write, save, and retrieve data. It has a powerful data centrical architecture with effective and climbable source storing and computational operations. Data processing and storage have become increasingly reliant on the cloud.

Cyber twin atmosphere develops more flexible, given the enormous productive, and efficient resource development. Motive of cyber twin to become flexible production environment as faster. To minimization of response time from source to destination is enriching network performance.

The scaling problem of the edge-cloud framework has a long-term solution in the Cybertwin-based network architecture. However, wireless network resource consumption is uneven because of diverse end devices, a wide range of applications with various resource requirements, and conventional offloading mechanisms. The interaction of heterogeneous terminal devices will also raise new security and privacy issues as we move beyond edge clouds. All of these problems would consequently reduce the overall QoS of the underlying network architecture and must be resolved right away.

Cybertwin directs resource distribution and computational offloading in the edge-cloud context. Cybertwin, an intelligent agent and digital representation of terminal nodes, is hosted in one of the edge clouds, where it manages computing, communication, and caching resources to offer high-quality user services. We use deep reinforcement learning (DRL) with the Cybertwin-based edge environment to address the problem of timely offloading optimization.

Cyber twin architecture introduces first for 6G. It can be followed by cloud centric network and radio access networks.it will maintaining an enormous statistics information. Extensively used for communication concluded the complete networks. Two parallel phase implementations are current system and prevailing system conditional upon the recent behaviour.

In traditional networks, the total system hardware involved has a lifespan of about a generation, after which the hardware is replaced by a new component. The entire communication mechanism is halted during the replacement procedure. However, in the case of a cyber twin system, the generation of any of the network's components may be foreseen, and it can be replaced before the network reaches its end.

To adhere mobile traffic in congestion network path cyber-twin manner characterized through a digital ID to identify the node data transfer by network path.

In background numerous authenticating procedures are provided by the architecture networks. Dissimilar security devices can be accessing the data.

IoT devices generate data and save it to a centralised cloud server for further processing in cloud computing. When new data is generated, this happens every time. The edge interface sits between IoT devices and cloud servers in edge computing. This edge could connect directly to the IoT device or be configured as near to the device as possible(Alkinani et al 2022). The goal of this technology is to move computation and data to the edge network, so that data does not have to be sent back to the central server every time an IoT device performs a function.

Using IOT strategies health-care segments are nowadays mounting exponentially in cyber-twin. Hereby using cyber-twinprobable to enlarging the amount of collected data through high-speed node transmission with accuracy and rapidity. Next strategy is used for tracking product in manufacture industry. Simulation methods are used for testing future products performances and measuring of life time(Munusamy et al 2021). The next step is to employ AI approaches to simulate smart buildings, smart traffic management, smart farming and smart livestock, and other aspects of smart city development.

Challenges by cyber-twin with 6G technologies are collaboration identical AI and IOT effective execution with end-to-end process. Next Encounters to be data transmission with high quality of processing node transfer. To be overcome by Interrupted and low-quality data centrals to the deprivation in performance of cyber-twin. To maintain complex data with highly confidential that leads to data security. Using IOT and AI concepts to maintain authorized access thus will defending the cyber-twin methodology. To develop standard methodology and uniformity by using cyber twin.Next challenge is maintenance of network traffic and congestion. Lack of resource availability at the time of network traffic.

The methodology employed in cyber twin is cloud-based cyber-twin architecture, which is unique. The development of a digital twin has been determined by considering data analytics and virtual authenticity. It is proposed that a five-layered cyber-physical framework be implemented in Industry 4.0 for effective production management. The concepts of cyber-physical doubles and their connections are re-examined. In 6G technology, a four-layer reference model for digital twin was employed. This reference perfect includes techniques for hazard prevention and prediction using a digital twin for improving employee workplace security. The issues of modern logistics have been identified using a combination of detached occurrence simulation and digital twin.

Cyber twin technologies are used to enlarging data transmission with high security and productive of information's with 6G wireless communication methods.Data will be transfer from end-to-end process without lacking of information. Attacks can be prevented by using effective algorithm measuring. IOE methods are securing data sharing with high-speed networks, it maintains network traffic and congestion.

The following are the primary contributions of the proposed framework:

1. **Model:** We developed an end-edge cloud architecture powered by Cybertwin that enables user devices to execute computationally demanding apps that are typically offloaded to the core or edge cloud. In the suggested framework, an offloading technique is developed for mobile users that efficiently maintains compute, communication, and cache resources while minimising end-to-end latency and energy consumption. In order to guarantee that only authorised users' representation is distributed in remote places, we have developed an authentication method for Cybertwin.
2. **Algorithm:** The offloading optimization problem is stated as a markov decision process to reduce latency and energy (MDP). We developed a multi-agent deep reinforcement learning algorithm that makes use of the twin delayed deep deterministic policy gradient (TD3) method to maximise rewards obtained.
3. **Simulation:** Extensive simulations are used to determine how effective the proposed approach is. The benchmark multi-agent deep deterministic policy gradient (MADDPG) technique, which was previously created, is used to compare the simulation results.

Numerous computationally intensive applications have been developed recently as a result of the expansion of smart IoT devices. Due to limited resources, resource allocation is frequently difficult near the network's edge. Resource allocation (caching, computation, and communication) at the network edge is a key area of focus for business and academia in order to enhance consumers' low latency experiences. Resource allocation has been a big issue in the past when finishing a variety of projects. Allocating computing power and bandwidth in the context of mobile edge computing reduces the burden of data-centric Internet of Things must focus on cutting down on how much energy is used and how long it takes to execute tasks.

To increase the frequency of resource allocation and information sharing, take into account applications that are delay-sensitive. Experimental results show that it performs much better than the baseline method in terms of latency and resource utilisation. Lv and Qiao investigated several resource allocation optimization algorithms (LV el at 2020). For resource allocation and computation offloading in an edge computing setting, the researchers proposed a two-stage heuristic optimization algorithm based on a genetic approach. The algorithm led to a decrease in overall energy use as the number of end devices increased. Different fuzzy strategies have been used to solve different optimization problems because of their ability to convey ambiguity.Ant colony optimization, artificial chemical reaction optimization, fuzzy c-means optimization, and other approaches are among them. The application of such fuzzy-based strategies can also aid with resource allocation and offloading issues.

However, all of the above-mentioned research utilises either matching auction theory or a one-shot optimization strategy to resolve the offloading problem. Existing works in both situations ultimately fall short of catching up with the changing computer environment. In the past, nontrivial mathematical equations have not been adequate to solve non-convex and NP-hard optimization problems. In order to change the overall problem to be solved, new assumptions are reviewed, such as a change in the type of the decision variable or a restructuring of the objective function. Recently, machine learning has become more well-liked for leveraging simple network topologies to solve complex optimization problems.

The issue of resource allocation has been studied extensively using supervised or unsupervised learning methodologies. On the other hand, these techniques study the complexity of networks using static datasets that don't interact or change in response to their surroundings. Also employed is reinforcement learning (RL), a machine learning technique whose main goal is to train the agent by continuously observing its surroundings.

Resource allocation is one of the many contexts in which reinforcement learning can be employed. This is a relatively recent application field. The resource provisioning problem has lately been addressed by many academics using reinforcement learning techniques. The provisioning of network edge resources is enhanced via a deep deterministic policy gradient (DDPG) strategy. The suggested system finds an optimal decision policy and outperforms the PG, DPG, and AC approaches. An energy-efficient resource allocation method based on MADDPG policy.

The effectiveness of the suggested technique is assessed while taking into account a variety of factors, including convergence, reward, and energy consumption. We employed federated learning and DRL for resource allocation in ocean federated learning IoT networks. The technique might be applied in federated distributed learning systems. a deep reinforcement learning-based resource allocation method for vehicle-to-vehicle communication networks. Using a deep recurrent Q-network (DRQN) approach, take into account the QoS requirements and resource demand characteristics. The given DRQN approach performs significantly better than the traditional Q-learning approach. Q-table hybrid learning method or Deep Q-network (DQN) (opt-QL). Performance of the suggested algorithm and conventional approaches are contrasted. According to the results, the hybrid learning approach works better than pure DQN.

Fast communication and excellent security are just two of the features that Cybertwin on the Border has been designed to meet. Contrarily, edge computing collects data from the cloud centre using device-based processing. As end-user resources get closer to users, there is less latency between devices and the cloud data centre. As a result, service quality could perhaps be slightly enhanced (QoS). Network speed and capacity will also become a major hurdle for cloud computing as the number of devices that can connect to the internet increases. Similar to this, service allocation is complicated by the complexity of end-user requirements. The most important factor in fulfilling end-user needs is wise resource selection.

When deploying large-scale distributed neural networks in the beginning, the low computing resources at edge devices provide many challenges. Due to a lack of resources, storage is constrained. Two of these limitations include a shortage of energy and structural problems. The edge lacks the capabilities to leverage cloud computing, despite their complementary relationship and the advantages of lower latency. Application performance, job scheduling, and end-user QoS are all significantly impacted by the allocation or forecast of available resources. A resource plan that employs particular factors to estimate the amount of resources consumed will be created after providing an estimate of the resources necessary for each end-user.

Network model

The fundamentals of edge computing, which sits in between the cloud and edge devices, are covered in this section. Storage, compute, and network services can all be aided by edge computing. In order to provide distributed computation with severely constrained memory and processing power in the edge nodes and edge devices, the distributed FNN wireless healthcare model is employed in edge computing. Edge devices' proximity to resources makes real-time communication possible.

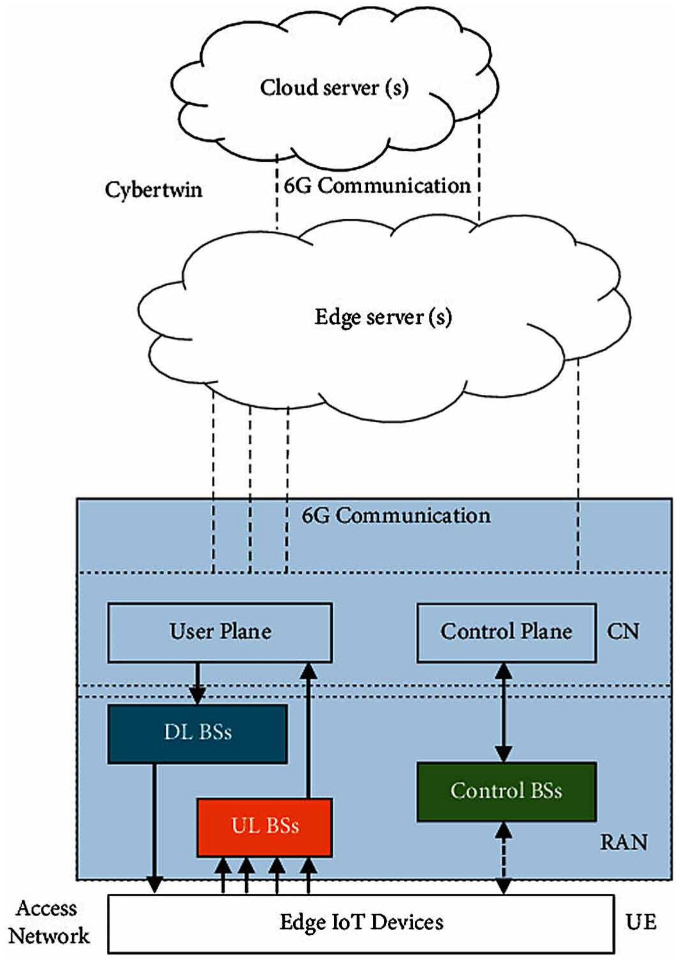
The cyber twin's goal is to handle data while allowing users to freely communicate and execute computations. As illustrated in, the current study employs a three-tier design with an edge computing approach.

6G Edge Computing Framework

The command The BSs is in charge of supplying data to the control plane, which determines the required resources for edge IoT devices. The user plane can distribute resources for data communication from edge devices via edge IoT devices using this resource allocation with uplink and downlink BSs.

This edge device is designed to generate data and encourage clients to consume it. Clients are encouraged to consume more resources from edge nodes rather than the cloud. Smartphones, IoT sensor nodes, intelligent vehicles, and even smart cities are examples of devices (Ramezani et al 2022). The edge devices use a sensor network to collect data and communicate with one another.

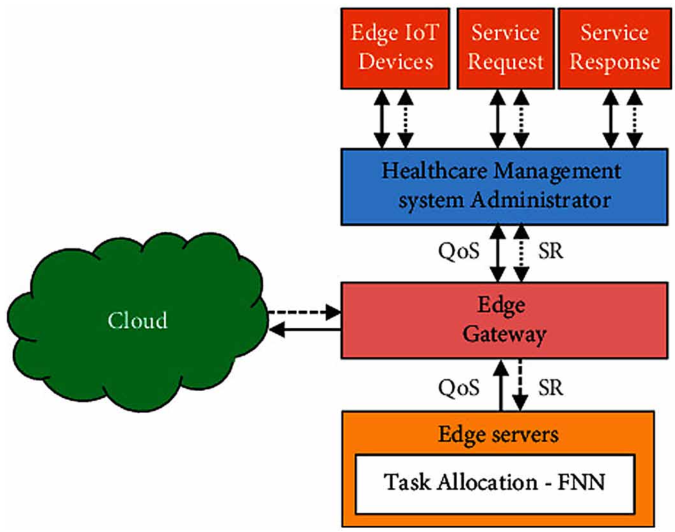
Figure 1.



On the edge nodes, you'll find switches, routers, and local servers, which are often used for specialised services. These nodes have computation, storage, processing, and data transmission capabilities. Edge devices can be connected to edge nodes or edge servers using a single or multihop connection. The microdata centre has computing, networking, storage, and software resources (MDC). At the top of the cloud layer are cluster servers and data centres that operate as storage and processing locations for data received from edge devices.

The process of submitting a service request when using edge computing. Users can send requests to the administrator starting with edge devices. In order to satisfy QoS user needs, the query is saved in the edge nodes and subsequently transmitted to the cloud centres via the edge nodes. Resources, sensor availability, services, and applications all have an impact on the data that the monitoring equipment provides. In order to give each user with the right levels of service, this equipment analyses the data provided to the edge nodes and the QoS service levels for their demands.

Figure 2.



The service request in cyber twin for resource allocation.

Previous research have emphasised the necessity of the cloud, edge nodes, and end-user devices for neural network implementation. Neural networks are less latency-sensitive when more resources are allocated, yet they may be able to transmit the original source in latency-critical applications. Edge customers can access dispersed services from several hidden layers of neural networks. The neural network model makes predictions with the least amount of delay and then allocates energy resources almost instantly. The distributed neural network design can identify when jobs are distributed to various edge partitions and will always select the partition with the fewest resources when distributed edge nodes are utilised.

To enable edge devices to estimate workload intelligently, high-performance distributed neural networks (DNNs) are being developed and deployed on edge networks. A long-standing constraint on computation and resource allocation, symbolised by the trade-off between computational load and resource job, must be upheld while using distributed neural networks.

Performance Evaluation

To evaluate the utility of the proposed model in this section, the complete simulation is done in Matlab. This section focuses on the testing done to see whether the FNN wireless healthcare model may help with judgments regarding resource allocation based on elements like average success rate, task response time, and resource utilisation levels. The Euclidean distance between edge devices is calculated in accordance with the distance model to estimate latency. When assigning resources, reaction time is considered, and an edge device with a slow response time is labelled as a failure-allocated task. 100 servers, each with six cores and five hosts, are spread across three data centres for the research. The location of the data centre is thought to be arbitrary.

The efficacy of the FNN wireless healthcare model is evaluated using three different performance measures. The average response time of each resource provided to the edge device is evaluated for the first time while calculating the effect of cloud servers on the network node. The likelihood of tasks allocated per failure and the average task utilisation at the node where the work is produced are also estimated. Network throughput is computed as a result of analysing each of these three replies in the context of reaction time constraints.

6G Challenges

Instead of being a single technology, 6G is a group of mobile system technologies. A fundamental idea of the 6G future opportunity is the concept of "New Services," from spectrum to infrastructure. The densification of small cell networks is the main focus of 6G wireless evolution. Applications System Innovation 2021, 4, 11 7, of 28 It is extremely difficult to densify networks without causing more inter-cell interference. Another important technological aspect of 6G networks will be the use of high-frequency spectrum, particularly millimetre wave (mmW) spectrum with exceptionally low latency. Examining the discrepancy between data traffic demand and network costs for various network scenarios is the challenge of 6G cost modelling.

6G tracks the tempo of telecommunications technology implantation in space. Studies frequently focus on a particular crucial aspect, like spatial viability modelling. There hasn't been much spatial modelling of potential new technology adoption in mobile systems(Wang et al 2021). Network densification and additional spectrum will be essential for delivering 6G, and this will have a significant impact on the economics of delivery. Only a portion of the 6G deployment's spatial and temporal dynamics have been studied. While recent research on 6G networks sheds light on the hazy technical specifications of rival technologies, the deployment implications for diverse international infrastructure projects have received less attention.

6GIIoE Architectural Framework

According to the literature review, there are no standards, specifications, or reference architectures for 6G-enabled IIoE systems. Scholars are hard at work on 6G specs and standards. We believe that global 6G standards and specifications will be released by the ITU (International Telecommunication Union) before 2030. IIoE devices, sensors, and machines should be able to connect to the Internet as well as communicate with one another autonomously thanks to the 6G communication architecture. Several 6G and IoT/IoE topologies are described in the present literature [43,44]. (p. 134010)

We believe that multiple layers are necessary to capture the complexity of the 6GIIoE system architecture because, unlike the 3 and 4-layer architectures, no global 6G and IIoE standard architecture has yet been proposed(Dash et al 2022). This is because the authors have extensive experience designing IoT/IoE architectures that are 5G enabled, as well as a literature review of related research work.

- The physical components of the perception layer are all sensors, devices, and machines.
- The information communication is handled by the transmission layer.
- Communication protocols, 6G, Wi-Fi, edge (cloud) computing, and data analysis software make up the middleware layer.
- IIoE intelligent applications are handled by the application layer.
- The business layer is responsible for defining the business model.

Wireless Connectivity

A factory typically employs and operates with antiquated industrial systems that are infrequently connected by wide area networks or cutting-edge wireless technologies like 5G or the upcoming 6G. It is also challenging to build a factory-wide monitoring and control network since legacy industrial systems automate using proprietary communication protocols that are not interoperable with one another. On the production floor, there are typically several data silos because data is frequently locked away in separate control loops.

Software-Defined Radio

Both wired (LAN) and wireless (Wi-Fi) networks are becoming common in industrial settings. A traditional IIoT architecture employs a variety of radio protocols and standards. Most businesses are considering implementing new forms of connection, like 5G or the impending 6G, to take advantage of high-value use cases.

In order to support software-defined radio (SDR), a radio communication technology in which the majority of signal processing is performed by software, it is imperative to build a 6GIIoE architecture that is backward compatible and effective enough. SDR or 6G-based IIoE gateways can implement and decode a wide range of protocols, which lowers the infrastructure's cost and complexity(Jain et al 2022). The above design may enable simple software updates with 6G wireless connectivity, allowing businesses to flexibly react to future operational and technical changes.

Priority Areas

According to the conclusions of the extensive literature research, IoT/IoE affects a variety of businesses, with customer experience, asset management, and financial decision making being the top priority areas. The first focus area is to improve the client experience. The customer experience can be improved in the following ways with a 6G equipped IIoE system:

- Constant monitoring of the customer's experience with the company's products and services.
- Making each customer's scenario unique.
- Adapting products and services depending on what we've learned from consumer interactions over time.
- Access to new products and a positive purchase experience are always being improved.

Asset management and financial decision-making, respectively, are the second and third priority areas. By allowing asset tracking and maintaining real-time visibility, the 6G enabled IIoE system can help financial decision makers make better decisions. Depending on the industry, the top priority areas may change. Customers' experiences and financial actions (asset management tracking) help practitioners and academic researchers assess a solid foundation of functional domains.

Furthermore, the Following Specific Priority Areas of 6G and IIoE Needed Attention

- 6G should entrust MTRLLC with mobility management at high frequency mmWave bands and THz beyond.
- The transition to 3CLS (Convergence of Communications, Computing, Control, Localization, and Sensing), powered by artificial intelligence, and end-to-end co-design is necessary for 6G. (AI).

- The 6G vision should be fueled by a diverse portfolio of applications, technology, and techniques.
- The 6G high density tiny cell base station concept might usher in a new era of smart surfaces communicating with human-embedded implants.
- 3D space should be used for the analysis and optimization of 6G performance targets.
- IIoE: customer-centric perspectives—the term "IIoE" refers to how the Internet of Things (IoT) can be employed in industrial applications. The Industrial Internet of Things (IIoE) focuses on enhancing device connection, time/cost reductions, and industry efficiency optimization. With the number of linked devices expected to reach 75 billion by 2025, the industrial internet of things is undergoing a major transition. Previously, IoE interconnection was only available to businesses with the financial resources to construct an IoE system. Companies nowadays want a pre-built and fully managed IoE solution that can provide the appropriate hardware, software, and edge (cloud) infrastructure to handle billions of concurrent device or sensor connections to edge (cloud) network infrastructure like 6G.
- Predictive maintenance and analytics (PMA) in the IIoE — Industries are employing IoE-enabled devices to send maintenance teams notifications when equipment is about to malfunction. This cuts maintenance time in half and gives maintenance staff more time to focus on other responsibilities. IoE-enabled systems collect a lot of information about how frequently these notification failures occur. Industries may utilise 6G to develop maintenance timetables and save money and time by allowing teams to service equipment before it malfunctions. PMA predictive maintenance and analytics is quickly becoming one of the most important applications of industrial analytics.

Theory Development for IS Innovation

The creation of a new theoretical framework is necessary because IS innovation is a complex, multi-dimensional process with dynamic interacting characteristics. Finding a theoretical framework that can navigate the intricate and varied landscape of IS innovation is the aim of our research. The dynamic interaction between the innovator and his or her surroundings is not fully taken into consideration by the current IS theories. Researchers can utilise an ecological systems theory as a valuable framework to study the subject of innovation and how it relates to information systems. The ecological systems theory, which is not studied because it is outside the purview of this study, can be mapped using the principles learned from the literature review.

It's critical to remember that IS research must situate internal organisational processes within a larger philosophical, economical, and ethical framework.

6G Future Benefits

Due to 6G mobile systems, we believe a new set of service classes and IIoE applications will develop. Several cutting-edge technologies will evolve at the same time as 6G deployment, with the potential to play a key role in the global standardisation of 6G. Quantum computing and communications (QCC) is one such renowned technology that can provide highly secure long-distance networking at Terabit per second speeds, vast network capacity, and extremely low latency. Integration of non-RF and RF communications will be one of the other 6G technologies. Industries are motivated to invest in order to realise a significant value proposition from the deployment of 6GIIoE, which will enable new business models(Padhi et al 2021). The value drivers of 6G connection for industries offer the following unique value proposition in addition to the benefits listed above: I increased dependability, (ii) increased bandwidth, and (v) improved operating efficiency The following is the DNA of the 6G value proposition:

- Beyond 6GHz with compact high-density cell stations for data-rich situations ● Human-centric services with better decision-making
- Large intelligent surfaces for communication.
- Automation agility for MTRLLC.
- The effectiveness of intelligence.
- Latency and dependency on TRLLC connectivity

6G Enabled IIoE System Theory Development

Taking into account the priority areas, issues, applications, and major prospects for connected intelligent industrial applications on the horizon, there is a void in the literature that establishes theoretical development on 6G-enabled IoE/IIoE system adoption and execution(Alsenwi et al 2021). We advise researchers to widen and cultivate a full study and theoretical framework components where research holes or gaps exist in the technology fields of 6G, IoE/IIoE, and CIIW.

IS innovations

Information and communication technology (ICT), operational technology (OT), and the fundamental components of the IIoE all intersect with the IS discipline in terms of their technological, business, and social implications. An developing digital breakthrough called 6GIIoE can give clients a more individualised experience. The development of 6GIIoE heralds the beginning of an uncharted chapter in digital history. There are several opportunities for scientific investigation in this environment. Therefore, in order to truly benefit society, major future research is needed to develop theories on 6G and IIoE that support philosophical, economical, and ethical components specifically for linked intelligent industrial applications. We emphasise the following future research agenda as a result:

- A multilevel investigation of heterogeneity and IS.
- Impact on organisations, technology, persons, organisations, and society are the four thematic influence domains.

Smart Service Contracts (SSC)

While 6GIIoE-related applications are being looked into, the potential of smart service contracts (SSC), distributed ledgers, and other blockchain-related technologies is far bigger. The research gaps in this field are as follows: Increased transaction rates per second; architectural support for Internet of Everything (IoE) devices or sensors to use pertinent technologies; secure methods for fixing unintentional SSC errors; and clarification of hidden costs when using blockchains to enable objective comparisons with conventional alternatives.

IIoE/Industrial Control Systems (ICS)

They are no longer a closed network, and they are vulnerable to a variety of threats and/or attacks. IoE is evolving into IIoE with the arrival of the fifth industrial revolution. ICS is responsible for the supervision and management of industrial processes carried out by critical infrastructures. The rise in assaults and/or threats affecting ICS has highlighted the importance of using Intrusion Detection Systems (IDS). Malware used to attack critical infrastructure has advanced to the point where it can imitate normal network traffic. As a result, it is critical to pursue research into how to protect against threats and/or attacks using machine learning and deep learning, as well as novel detection techniques for identifying threats and/or attacks.

Securing the 6GIIoE Infrastructure

The coming together of 6G, IoE, and other upcoming digital technologies will open up a world of new possibilities and ubiquitous experiences. Many fascinating phenomena are shaping the future and ushering in a new era of digital destiny. End-user appetite for network capacity and speed, such as 6G, is a significant difference in the digital journey. The IIoE will be delivered using network technologies in the end. The following are industry-specific suggestions for achieving safe and secure connectivity: I Manage security at all levels of the 6GIIoE system, (ii) Protect the identity of objects and users, (iii) Implement multifactor authentication, and (iv) Protect identities rather than gateways.

Cyber Physical System (CPS)

Most sectors are transitioning to an era of self-contained IIoE, and security must be regarded a critical component. Future threats or attacks must be given top priority in order to preserve the IIoE's increasing growth rate. Interconnectivity and new application scenarios such as Machine to Machine (M2M) must also be taken into account. This propels the IIoE's growth of devices and sensors. Threats to cyber security via 6G wireless communications are on the horizon. As a result, there is an increasing need for novel communications systems for industrial applications. As a result of the convergence of technologies in the industrial sector, Cyber Physical Systems are required. As a result, research into resource-saving, low-cost, and effective alternatives to traditional cryptographic approaches is critical if IIoE systems are to be trusted.

Security of IIoEvia ICMETRIC Technology

IIoE shows great promise for enhancing supply chain efficiencies, quality assurance, communications, and corporate productivity for the majority of sectors. The tremendous potential of IIoE data in industrial applications is being addressed by 6G, a paradigm-shifting network technology. On the other hand, the adoption of IIoE results in more vulnerabilities and dangers. In order to fully adopt and implement the 6GIIoE system, it is essential to enhance and increase the security of 6G communications in IIoE(Niyato et al). Future industrial manufacturing must embrace security to manage risks by integrating it into the IIoE ecosystem at an early stage of the product design life cycle in order to deliver security in a cost-effective manner.

Feature Extraction Strategy

It is difficult to analyse and store the enormous amount of data gathered from IoE devices at the local edge device. The majority of the data provided by these devices is filled with redundant and pointless features, which raises processing costs for edge networks, reduces prediction accuracy, and reduces feature dimensionality. In order to decrease processing overhead, shorten training times, and improve the predictive accuracy of AI models, feature extraction techniques can be applied at the edge networks. Inspiring by this idea, we provide the IGs feature extraction technique to precisely extract data features from IoE applications. Let U and V represent the difference between two feature variables. Following that, IGs selects a distinctive trait, as illustrated below. IG(U|V) = H(U) H(U|V) where U and V are discrete random variables (19).

Average Energy Consumption

The energy consumption of 6G-enabled edge networks is driven by compute (E user k) and communication (E upload k) when processing and transmitting data from IoE applications on remote computing devices. While keeping track of dynamic changes in the network and the characteristics of the input tasks, the proposed DRL-driven resource provisioning technique distributes incoming tasks via the Cybertwin server to nearby edge devices (Jk,Lk). In this study, the baseline methods and the proposed DRL resource provisioning method were compared. The comparison demonstrates that, in terms of processing incoming tasks on 6G-enabled edge networks, the suggested technique performs up to 15% better than the baseline techniques.

CONCLUSION

Moving towards better hosting, development, and management of the various new era services that are backed up by 5G and 6G, the requirement for availability of the various computer resources that MEC delivers. Because MEC has a limited number of resources, the infrastructure for various applications, particularly the cellular network, is considered a difficulty. Keeping an eye on this work, a suggestion about IScaler is made. IScaler is a multi-application that aids in scaling and may handle the various obstacles that come with working in a dynamic environment. It is clear that DRL-based apps that use 5G or 6G are more expensive. Some suggestions for the optimizer, IScaler, and long solution switch are provided here.

However, it is clear that ISP is effective at making decisions and performing intelligence with the aid of multi-application decision making.When using IScaler, the optimizer is employed.Understanding IScaler's unique capabilities.As a result, it can be combined with current methods such as model-based scaling.

REFERENCES

Adhikari, M., Munusamy, A., Kumar, N., & Srirama, S. N. (2021). Cybertwin-driven resource provisioning for IoE applications at 6G-enabled edge networks. *IEEE Transactions on Industrial Informatics*, *18*(7), 4850–4858. doi:10.1109/TII.2021.3096672

Alkinani, M. H., Almazroi, A. A., Adhikari, M., & Menon, V. G. (2022). Artificial Intelligence-Empowered Logistic Traffic Management System Using Empirical Intelligent XGBoost Technique in Vehicular Edge Networks. *IEEE Transactions on Intelligent Transportation Systems*, 1–10. doi:10.1109/TITS.2022.3145403

Dash, S. P., Joshi, S., Satapathy, S. C., Shandilya, S. K., & Panda, G. (2022). A cybertwin-based 6G cooperative IoE communication network: Secrecy outage analysis. *IEEE Transactions on Industrial Informatics*, *18*(7), 4922–4932. doi:10.1109/TII.2021.3140125

Dhiman, G., Nagar, A., Vimal, S., & Rho, S. (2022). Guest Editorial: Cybertwin-Driven 6G for Internet of Everything: Architectures, Challenges, and Industrial Applications. *IEEE Transactions on Industrial Informatics*, *18*(7), 4846–4849. doi:10.1109/TII.2022.3151914

Jain, D. K., Tyagi, S. K. S., Neelakandan, S., Prakash, M., & Natrayan, L. (2021). Metaheuristic optimization-based resource allocation technique for cybertwin-driven 6G on IoE environment. *IEEE Transactions on Industrial Informatics*, *18*(7), 4884–4892. doi:10.1109/TII.2021.3138915

Jain, V., Kumar, B., & Gupta, A. (2022). Cybertwin-driven resource allocation using deep reinforcement learning in 6G-enabled edge environment. *Journal of King Saud University-Computer and Information Sciences*.

Ji, B., Wang, Y., Song, K., Li, C., Wen, H., Menon, V. G., & Mumtaz, S. (2021). A survey of computational intelligence for 6G: Key technologies, applications and trends. *IEEE Transactions on Industrial Informatics*, *17*(10), 7145–7154. doi:10.1109/TII.2021.3052531

Khan, L. U., Han, Z., Niyato, D., & Hong, C. S. (2021). Socially-aware-clustering-enabled federated learning for edge networks. *IEEE eTransactions on Network and Service Management*, *18*(3), 2641–2658. doi:10.1109/TNSM.2021.3090446

Khan, L. U., Tun, Y. K., Alsenwi, M., Imran, M., Han, Z., & Hong, C. S. (2021). A dispersed federated learning framework for 6G-enabled autonomous driving cars. arXiv preprint arXiv:2105.09641.

Khan, L. U., Yaqoob, I., Imran, M., Han, Z., & Hong, C. S. (2020). 6G wireless systems: A vision, architectural elements, and future directions. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 147029–147044. doi:10.1109/ACCESS.2020.3015289

Munusamy, A., Adhikari, M., Balasubramanian, V., Khan, M. A., Menon, V. G., Rawat, D., & Srirama, S. N. (2021). Service deployment strategy for predictive analysis of FinTech IoT applications in edge networks. *IEEE Internet of Things Journal*, 1. doi:10.1109/JIOT.2021.3078148

Padhi, P. K., & Charrua-Santos, F. (2021). 6G enabled industrial internet of everything: Towards a theoretical framework. *Applied System Innovation*, *4*(1), 11. doi:10.3390/asi4010011

Qadir, Z., Le, K. N., Saeed, N., & Munawar, H. S. (2022). *Towards 6G Internet of Things: Recent advances, use cases, and open challenges*. ICT Express.

Ramezani, P., Lyu, B., & Jamalipour, A. (2022). Toward RIS-Enhanced Integrated Terrestrial/Non-Terrestrial Connectivity in 6G-Enabled IoE Era. arXiv preprint arXiv:2203.11312.

Chapter 4
Edge Computing on IoT:
Architectures, Techniques, and Challenges

Mahalakshmi R.
Presidency University, India

Uzra Ismat
Presidency University, India

Praveena K. N.
Presidency University, India

ABSTRACT

The internet of things (IoT) is escalating into diverse aspects of our lives with innovative technologies and solutions. In general, IoT devices are restricted to storage and processing power, which results in the lack of performance, reliability, and privacy of IoT applications. The applications in various sectors like agriculture, healthcare, smart cities, smart homes, and production units are enriched by twining the IoT and cloud computing. Cloud analytics is the process of extracting actionable business insights from the data stored in the cloud. Cloud analytics algorithms are applied to large data collections to identify patterns, predict future outcomes, and produce other useful information to business decision makers. Edge computing has arisen to support this intense increase in resource requirements by leveraging the untouched potential away from the enterprise data cores. Processing power is gained by a collective process between various entities at the network edge including the user devices, mobile-based stations, and gateways and access points.

DOI: 10.4018/978-1-6684-5722-1.ch004

The Internet of Things "IoT" has been a buzzword for recent years, the business organization has adopted the technology in an increased way. Current expansion of IoT demands intelligent models to be deployed at the edge. The Internet of Things (IoT) is escalating into diverse aspects of our lives with innovative technologies and solutions. In general, IoT devices are restricted to storage and processing power, which results in the lack of performance, reliability and privacy of IoT applications. The applications in various sectors like agriculture, health care, smart cities, smart homes, and production units are enriched by twining the IoT and cloud computing. Integrating AI techniques into cloud analytics, is called as cloud intelligence. Edge Computing is a archetype to push cloud services from the network core to the network edges. The goal of Edge Computing is to host computing tasks as close as possible to the data sources and end-users. The edge computing and cloud computing are not mutually exclusive. The edge computing is the extension of the cloud and which complements instead.

The core functions of edge computing are, to perform analytics at the edge, data needs to be viewed as real-time flows. Edge computing continually processes streaming flows of data in motion, whereas big data analytics is focused on large quantities of data at rest.

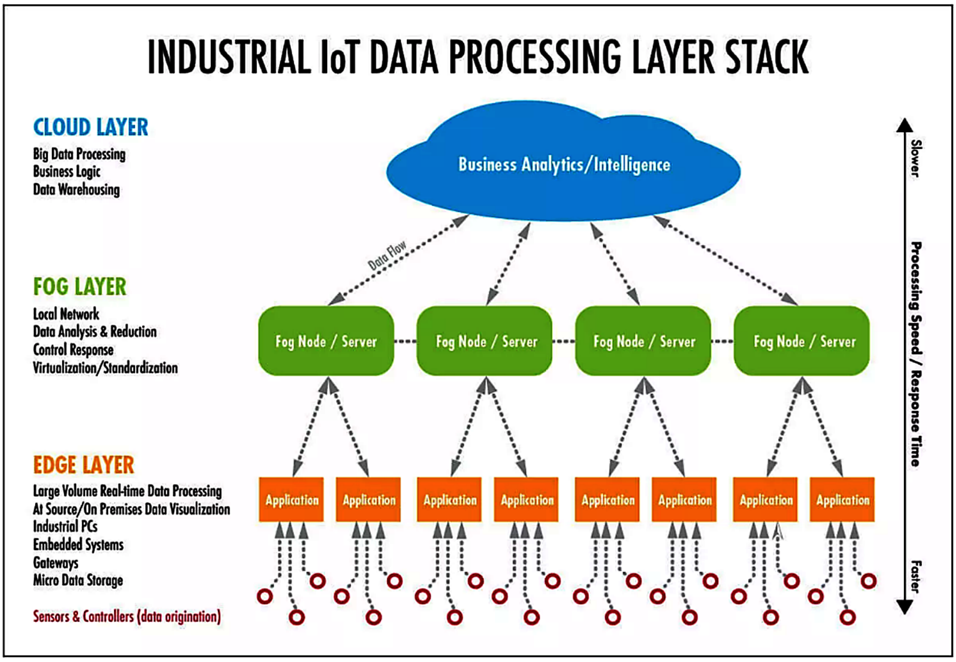
Organization of the work will be as follows, Introduction, overview of edge computing, machine learning and deep learning algorithms, the importance of edge intelligence, enabling techniques and advantages, applications of edge intelligence, finally the conclusion.

INTRODUCTION

Advanced computer technologies such as big data, Artificial Intelligence (AI), cloud computing, digital twins, and edge computing have been applied in various fields as digitalization has progressed.

The most important inspiration for Digital Twins (DTs) comes from the need for feedback between real physical systems and the digital cyberspace model. People try to recreate what occurs in the material world in digital space. Only the whole life tracking using cyclic feedback is the true concept of the whole life cycle. This way, digital consistency with the material world may be truly ensured throughout the life cycle. Various simulations, analysis, data accumulation, mining, and even artificial intelligence applications based on digital models can ensure that it is suitable for real physical systems. An intelligent system's intelligence must first be observed, modeled, evaluated, and reasoned. If there is no accurate modeling description of the actual production system by the digital twins, the intelligent manufacturing system cannot be realized.

Figure 1.



Data acquisition, data modeling, and data application are three principal parts of the advanced twins. Data collection refers to the full utilization of satellite remote detecting, shifted ethereal photogrammetry, lidar estimation, cameras and different innovations to get three-layered information from a total actual space scene19. The capability of the sensor is to acquire various types of genuine information in the genuine world20. The specialized trouble and key of information assortment is the high accuracy and effectiveness of information assortment, which decides the quality, proficiency and cost of information assortment. In the wake of getting a lot of unique actual world information, information demonstrating was done, and programmed displaying devices were utilized for additional handling to produce a three-layered model of the genuine recuperation of the actual world. Notwithstanding high-accuracy virtual reproduction of the climate, advanced twin information is more successful in supporting different working cycles. Information displaying can be separated into two sections: visual 3D modeling and semantic modeling. Visual 3D displaying is a 3D propagation of the actual world.

Rapid technological advancements, along with severe competition and the necessity to survive, are driving an increasing number of businesses to embrace the Internet of Things (IoT), According to a Gartner survey, there will be anywhere from 25 and 50 billion "things" connected to the IoT by 2020. To put things in perspective, that equates to around 7 linked devices every human on the planet. By 2020, the amount of data created by all of these devices is estimated to reach 2.3 zettabytes. Google how many zeros there are following '1' - we can't imagine the amount of data! Let's go over the fundamentals of IoT before we get into the problems of handling this massive volumes of data. IoT stands for "internet of things." The Internet of Things (IoT) is a term that refers to billions of physical objects linked to each other and sharing data throughout the world. Iot is a giant network of different objects ("things") sensor technology, actuators, and software that communicate and exchange communication between devices and networks through the internet, according to a more technical definition. Linking all of these gadgets and embedding sensors in them gives machines that would otherwise be stupid a level of technological intelligence. With the Internet of Things in place, these gadgets become sentient, allowing them to convey real-time data without the need for a person.

Ken Sakamura of the University of Tokyo coined the term "computers everywhere" in 1984 (Zhou, 2013), while Mark Weiser of Xerox PARC coined the term "ubiquitous computing" in 1999 (Weiser, 1999). Kevin Ashton, however, was the first to coin the word "Internet of things" (IoT) in the domain of supply network managing in 2009 (Ashton, 2009). David Brock expanded on the Internet of Things notion in a paper published by the Massachusetts Institute of Technology's Auto-ID Center in 2001 (Uckelmann, 2011). IoT in the real world isn't as simple as snapping on a light with a phone. Plants and machinery are intricately designed entities. Each is frequently reliant on the other for smooth operation. All of the gizmos and gadgets in a conventional plant are equipped with a slew of sensors and actuators that constantly churn out a massive amount of data. Organizations almost always use cloud hosting to handle IoT, especially if their physical plants are spread out throughout the country. Furthermore, the IoT is defined by Gubbi et al. (Gubbi, 2013) as the "connectivity of sensing and actuating devices that allows for the transmission of information across systems using a coherent framework, resulting in an unified operational picture that allows for innovative applications." Cloud computing serves as the unifying framework for seamless ubiquitous sensing, database management, and information representation. Compatibility between devices connected is commonly the responsibility of each individual vendor, who ensures it through the use of proprietary interfaces and protocols, the installation of add-on software customers on gadgets, or the use of a gateway device. (Kotis, 2012). However, devices are projected to develop enough intelligence in the future to interoperate directly without the use of specific gateways (Mazhelis, 2013). The gadget ("thing"), as well as its sensors and actuators, make up a typical IoT Implementation framework. The controller, that is nothing more than a firmware element that communicates electronically or conceptually with the device's sensors and actuators, is in charge of these. An integrated application that operates on or near the Iot system is known as a software agent or gateway. This gateway's job is to inform the controlling software about the status of a module or the condition of the system. Between the controller and the cloud, the agent serves as a link. The agent's job is to figure out which data should be sent to the cloud and which should be discarded. The agent also analyses and adapts to cloud-based instructions and upgrades supplied by the central IoT application, making it a two-way mechanism.

Cloud-Based IoT's Limitations

Cloud computing was the most prevalent foundational innovation for many organisational structures for years, and firms chose to shift all of their data, computation, processing, and other services from "on-premise" infrastructure to the cloud (P. Escamilla-Ambrosio, 2017). From the viewpoint of a company that can arrange for all elements to change automatically, the cloud looks to give infinite storage space and computing scalability without any restrictions. As a result, it's possible that less time and money will be spent managing "on premise" systems (Patierno, 2016). Cloud computing offers on-demand network access to a distributed pool of configurable computing resources, such as networks, hosts, memory, apps, and services, that may be swiftly created and released with little effort or contact from service providers (Cox, 2011). A cloud-only design, on the other hand, is not the ideal solution for many enterprises. Despite the fact that cloud storage costs are dropping, transferring and keeping massive amounts of data for analysis on the cloud becomes prohibitively expensive very quickly (P. Escamilla-Ambrosio, 2017). In most cases, all data is supplied in its present fidelity, with no method of knowing which data has the greatest economic worth and which data may be discarded. Cloud-based IoT systems also require a constant network connection, making them less than ideal for organisations with remote operations who can't afford to retire functioning if their link fails (Watson., 2016).

As we've seen, the Internet of Things (IoT) is made up of networked devices that communicate data. The software that manages the IoT system takes in the information and chooses what to do next. Whilst it is all well and well in principle, there are constraints in actuality.

Bandwidth: IoT systems create massive amounts of data (Big Data), which will strain existing bandwidth. IoT devices, regardless of connection, put a burden on bandwidth. An IIoT-connected chemical reactor, for example, may create hundreds of gigabytes of data each day with all of its sensors and actuators. Using a software platform to properly handle this data is a problem in and of itself.

IoT data must go from one place to another, despite the fact that it is digital. While bandwidth refers to the pace at which data is sent over a set length of time, latency refers to the time it takes for data to travel that distance. The data transfer trip from transmitter to the receiver is anything from simple, thanks to the numerous routers engaged. Latency is the term for this lag.

ISPs are notorious for being capricious when it comes to connectivity. In certain nations, power disruptions are common. If the data stream is disrupted for any circumstances, the IoT system will suffer.

Data security is one of the most significant challenges that IoT systems face. With cyber assaults becoming all too common, it's more important than ever to secure IoT data.

The Necessity of Edge Computing

While IoT benefits organisations, it also has its own set of difficulties, particularly when employed in the cloud, as we've seen. Edge computing is a computer infrastructure that processes data near to the data source at the network's edge. If an edge IoT solution is installed on the sensor rather than sending the data to the cloud, it will quickly make a reasonable decision and take appropriate remedial action. As a gateway to the Internet, the Internet edge infrastructure is vital in facilitating the events and programs that are vital to the advanced enterprise's operation (Internet Edge Solution Overview, 2010). Popular Internet edge components include routers, switches, and firewalls. The edge of the Internet of Things, on the other hand, is made up of system terminal points that communicate with and transmit real-time info generated by smart products and services such sensors, actuation, and smart objects (Biron, 2016). Rather than transmitting all of the data to a cloud in Eu, the sensor activates a mechanism that immediately solves the problem. As a result, edge computing in IoT avoids bandwidth limits, latency, and connection problems. It's a technique of getting computation and storage services closer to an institution's hardware objects, or at the Cloud network's edge. Edge computing is a decentralised computation architecture that allows assets and application services to be spread throughout the data source to cloud communication line. That is, computing requirements can be met "at the edge," where data is gathered or where the consumer performs certain tasks. The strong and seamless interface between IoT and Cloud; between both the physical world and the realm of computation, is a key component of Edge computing. It's critical to remember that the Edge is a conceptual layer rather than a physical split, therefore "where" the edge is is a matter of personal opinion and interpretation.

The following are some of the benefits of edge computing in IoT:

1. Provides protection for equipment and machinery against extremes of heat, cold, pressure, and other factors.
2. Keeps track of the plant's / manufacturing lines' performance.
3. Optimizes a location's or plant's supply chain

4. Prevents equipment failure by arranging preventive maintenance in advance.
5. Assists in mitigating data security risks.
6. Lowers operating expenses

Where Should IoT Data Be Hosted: Cloud or Edge?

While edge computing is considered a cutting-edge solution, it is not always the best option. IoT systems benefit from data scalability provided by cloud hosting. A central repository is highly important for large enterprises in gaining significant information into the performance of several plants / manufacturing units. Industrial IoT platforms that require speed, security, and scalability are best supported by cloud solutions.

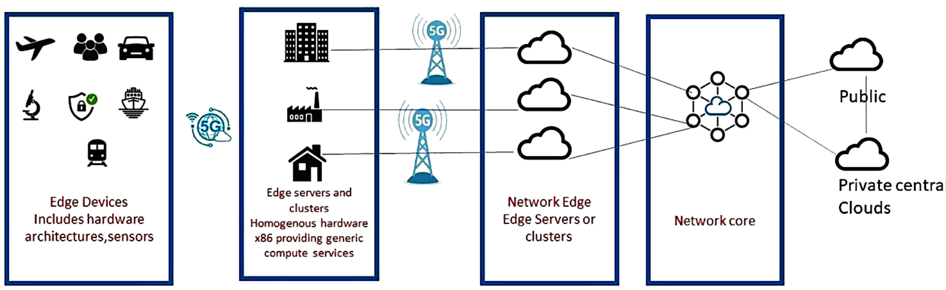
Edge computing is beneficial for applications when data delay is not an option, such as data analysis. As a result, edge computing is the preferred method for operations that require real-time data processing and response without relying on a centralised server farm. Edge computing, on the other hand, is more expensive than cloud-hosted IoT due to the difficulties of maintaining and upgrading distributed hardware.

OVERVIEW OF EDGE COMPUTING

Edge computing is the practise of bringing computer workloads (including hardware and software) as near to the edge as possible—to the point where data is produced and actions are performed. Customers benefit from edge computing environments because they get faster reaction times, more data privacy, and lower data transport costs. The phrase edge computing was initially created about 2002 and was primarily linked with the deployment of applications across Content Delivery Networks, according to (Garcia-Lopez, 2015). The main purpose of this strategy was to achieve massive scalability by using CDN edge servers' proximity and resources. Edge nodes that transport network traffic include routers, mobile base stations, and switches (Varghese, 2017). As previously stated, the term "edge" in this case refers to the Internet's edge. These devices do complex processing to manage the packets that come in across the multiple subnetworks they aggregate. Edge devices include devices at a datacenter's frontend that conduct operations such as XML acceleration, load balancing, or other content processing, as well as devices at an enterprise's entrance point that provide security-related services such as firewalls, intrusion detection, and virus scanning (Reinders, 2007). Rather of being centralised, this strategy distributes applications, data, and processing to the network's logical edges. The amount and distance that data must be transported is minimised when data and data-intensive applications are positioned near the edge (OpenFog reference architecture for fog computing, OpenFog Consortium., 2017). Edge devices, also known as Internet edges in (Varghese B. W., 2016), are devices with a direct Internet connection. A smart phone, for example, serves as a barrier between the body and the cloud, a smart home gateway serves as a barrier between the house and the cloud, and a cloudlet and a small data centre serve as a barrier between a mobile device and the cloud.

A rudimentary edge computing architecture is shown in Figure 2.

Figure 2. Simplified architecture of edge computing.



Edge computing models are in great demand as the Internet of Things (IoT) grows, and they ‘ve emerged as attractive fields of research.

2.1 Edge Computing Concepts

Edge computing is not at all like traditional cloud computing. It’s a new computer paradigm in which processing happens at the site’s edge. The main idea behind it is to move computing nearer to the information provider (Satyanarayanan, 2017). Various scholars describe edge computing in different ways. “Edge computing computing is a new computing mode of network edge execution,” Shi et al. (W. S. Shi, 2019) (W. Shi, 2017)– (W. Shi J. C., 2016) The unconstrained computing and network facilities seen between data source and the cloud computing center’s path are referred to as the edge of edge computing.”The edge of edge computing relates to the random computing and network resources in between data source and the cloud computing center’s route, and the cloud facility is depicted by downlink data of edge computing, the Internet of Everything is depicted by uplink information of edge computing. Edge computing is defined as “a novel computing paradigm that distributes computing and storage resources (such as cloudlets, mini data centres, or fog nodes, etc.) at the network’s edge closer to mobile devices or sensors,” according to Satyanarayanan, a professor at Carnegie Mellon University in the United States. (Satyanarayanan, 2017). Based on the above two definitions, Zha et al. (Z. M. Zha, 2018) proposed that "Edge computing is a revolutionary computing paradigm that integrates resources that are close to the user in geographical distance or network distance to deliver computation, storage, and network for applications service." According to China's edge computing industry alliance (Wang, 2018), edge computing is delineated as an open framework that combines core capabilities such as networking, computing, storage, and applications, and provides edge intelligent services nearby to meet business flexibility essential components in link, actual business, information optimization, app intelligence, security, and privacy. In other words, edge computing is the deployment of services and calculations at the network and information generation's edge. Edge computing is the migration of cloud network, computing, storage, and resource capabilities to the network's edge, as well as the provisioning of intelligent services at the edge, to meet the IT industry's critical needs in nimble connections, real-time business, data optimization, application intelligence, security, and privacy, as well as the network's low latency and high bandwidth requirements. Edge computing has become a research hotspot (The Internet of Things reference model, 2014) (Ansari, 2016) (J. Kang, 2018) (A. Alrawais, 2017)– (C. Mouradian, 2018).

Edge computing (Varghese B. W., 2016) is a concept that aims to move a piece of the network's processing power to the network's edge. The computing capacity in edge nodes like switches, routers, and stations may be put to good use in this fashion. Edge computing is a concept for fine - tuning cloud computing systems or applications by transferring services or data from the system's heart or centre to the Internet's edges or conceptual extremities. The various nodes that reside in between edge device and the cloud can hypothetically enable edge computing. Nodes include switches, gateways, routers, and base stations. Edge nodes can have their resources increased so that they can be utilised for broad sense computing (Varghese B. W., 2016). DSPs can also be substituted with broad sense processors to allow computation, albeit this will require a large investment. Many commercial vendors have lately attempted to create software solutions to implement the Edge Computing concept. Cisco's IOx16 (Meurisch, 2015) programme is an example of this type of initiative. It provides a runtime environment that may be used on its system's integrated service routers. Any software built in this domain is clearly hardware-specific, and therefore cannot be forced to work well in a variety of scenarios.This brings readers to 1 of the key critical points in this discipline. There is a demand for portable software that may be used in a range of situations.

2.2 Edge Computing Research Content

Edge computing has aided the rapid expansion of the Internet of Things and has made a substantial contribution to the realisation of a smart society. As a result, edge computing has become a hot issue amongst researchers both at domestically and overseas.

2.2 1 Essential Technologies

Edge computing's primary technologies include various levels of computation offloading, mobility management, and security.

cache acceleration, traffic offloading technology, and network
Control, for example.

1) Offloading computing

Processing offloading is a scarce resource technology that shifts resource-intensive computation from mobile devices to nearby architecture with more resources to overcome mobile device limitations in resource retention, computing efficiency, and energy consumption (Becvar, 2017). Computing offloading technique reduces transmission delays while also relieving demand on the primary network. Edge computing relies on compute offloading, and mobile edge computing (MEC) may run new, complex apps on user devices (UE).

Many relevant research accomplishments have been made, with the two key challenges being offloading choice and resource allocation. How to offload computing tasks, how much to unload, and what to offload are all considerations when offloading for mobile devices. Allocation of resources is all about figuring out where to deposit resources.

A) Decision on offloading

The offloading choice made by the UE refers to the problem of determining how to offload computing work, how much to offload, and what to offload. A code parser, a system parser, and a decision engine are commonly included in the UE of an offloading system. It follows a three-step approach to carry out the offloading decision: (1) The code parser determines what can be offloaded, with the specific offloading content determined by the application type and code data partition; (2) The system parser monitors various parameters, including available bandwidth, the size of the data to be offloaded, and the energy used to execute a local application; (3) The decision engine then decides whether or not to discharge.

Local execution, total offloading, and partial offloading are the three forms of UE offloading decisions.

The unique choice results are influenced by the UE's energy consumption and time to finish computational processes. Based on the optimization goals of the offloading choice, computing offloading may be divided into three categories: lowering latency as a goal, reducing energy consumption as a goal, and balancing energy consumption and latency as a goal.

B) Allocation of resources

We must evaluate the question of acceptable resource allocation, or where to offload, after we have made our offloading decision. If the UE's computing job is indivisible or may be divided but the divided parts are connected, the offloading task must be performed on the same MEC server; otherwise, computing duties that can be divided but are unrelated to the divided component can be distributed over several MEC servers. At the moment, there are two types of resource allocation nodes: single-node allocation and multi-node allocation.

2) Management of mobility

Edge computing focuses on the geographical spread of resources to support application mobility. Users in the local neighbourhood are served by an edge computer node. Mobile application administration in edge computing is a novel way, since the cloud computing mode supports application mobility by establishing the server's position and delivering data to the server over the network. Resource finding and resource switching are the two fundamental concerns at hand.

Users must rapidly find the resources available around them and select the most appropriate resources during their travel, which is known as resource discovery. Edge computing asset revelation must acclimate to a heterogeneous resource atmosphere while also making sure resource detection pace so that applications can proceed to provide services to users without disruptions; resource switching - mobile applications' computing resources may swap among various devices when users move (C. T. Ding, 2019). To maintain service continuity, the operation site for the service programme will be migrated. One of the most important considerations in MEC is how to ensure that consumers' access to services remains uninterrupted during the transition. Some apps anticipate continuing to serve customers even if their physical location changes. Because of the variety of edge computing assets and network diversity, the migration process necessitates flexible device computing functionalities and variations in network bandwidth. By predicting the user's movement, reference (A. Nadembega, 2016) refined the virtual machine migration technique and proposed the MSMP (mobility-based service migration prediction scheme), which followed an expenditure-quality compromise.

3) Further essential techniques

Edge computing also includes traffic unloading, cache acceleration, and network control, in addition to the two fundamental technologies mentioned above.

A) Technology for Offloading Traffic

Economic applications in cordless networks can benefit from traffic offloading in order to achieve localisation, narrow installation, and minimal-latency, elevated bandwidth transmitting capabilities. Unloading (traffic offloading) of edge network traffic is crucial in mobile edge computing. Traffic offloading is used to unload traffic that meets specific offloading rules to mobile edge networks in order to save backhaul bandwidth, reduce latency, and stimulate the proliferation of more MEC services (that is, a local specialised network, either - intranet or the Internet). Reference (F. Lu, 2020) proposed an energy-efficient traffic offloading system for mobile users in a 2 different heterogeneous wireless network. In normal network conditions, the technique can save up to 34% of energy, according to test results.

B) Acceleration of caching

Mobile edge caching solutions include ground station caching, mobile content transmission networks, and straightforward caching. Caching acceleration technology can boost information delivery efficiency while also improving the user experience. Users can access material locally once it has been cached on the edge of the mobile network, minimising content transmission and relieving burden on the backhaul and core networks.

At the same time, edge caching can minimise the network latency required by the user, resulting in a better network experience for the user. Edge caching also offers the ability to extend the mobile network resource work atmosphere and provide more functions to tenants and users (Wang D. Z., 2017). A cognitive agent (CA) was proposed in reference (R. Wang, 2020) to aid users in caching and completing MEC actions beforehand, as well as synchronizing interaction and cache to reduce MEC strain.

C) Control of the Network

An edge network is a network that isn't officially defined . The edge network, which includes portion or all of the accumulated layer network and the access layer network, is the final component of the network for reaching users. Edge networks are for-profit networks that interconnect existing core networks with high-value customers. In terms of network control, reference (G. S. Aujla, 2018) created a useful workload slicing system that allows users to run data-intensive apps in multi-edge cloud setups using software-defined networks.

2.2.2 Data Secrutiy and Protection of Privacy

One of the most hotly debated topics in research is the security of edge computing. Personal privacy is a concern with network edge data.

Despite the fact that the concept of data processing locally also provides more organised assistance for data security and privacy protection, the distributed architecture of edge computing enhances the dimensionality of threat vectors. As edge computing clients get smarter, they become more vulnerable to malware attacks and security breaches. Data security protection solutions now in use are insufficient for edge computing systems. Furthermore, the extremely dynamic environment at the network's edge makes it more susceptible and difficult to secure. In edge computing, there are four additional issues to data security and privacy protection:

(1) New requirements for lightweight information encryption and quite well information sharing based on multiple authorised parties have emerged in edge computing. Since edge computing is a computer paradigm that combines numerous trust domains with permitted entities as trust centres, conventional data encryption and transmission methods are no longer applicable.As a result, designing a data encryption solution for numerous authorisation centres is critical. At the same time, the algorithm's complexity should be evaluated.

(2) Control and security management of multi-source heterogeneous data transmission in a distributed computing environment. Users or data owners want to be able to use appropriate information dispersion control and access control strategies to enable data distribution, search, access, and control over the breadth of data authorization. Furthermore, because data ownership and management are separated owing to the outsourced nature of data, a good audit verification method can verify the data's integrity.

(3) Security issues for edge computing services with large-scale interconnected services and terminals with limited resources. (3)Because of the multi-source information aggregation attributes of edge computing, the convolution of mobile and Internet networks, and the resource drawbacks of storage, computing, and available power of edge terminals, conventional and more complicated encryption techniques, access control measures, veriðcation protocols, and privacy protection methods cannot be used in edge computing.

(4) New edge computing mode needs for successful privacy protection, as well as diverse capabilities for the Internet of Things. Apart from the need to design effective data, location, and authenticity privacy protection schemes, future research will focus on how to merge conventional privacy security protocols with edge data processing character traits in edge computing situations to enable data confidentiality protection in a variety of service atmosphere.

Edge computing security and privacy protection research is still in its early stages, with few published findings. Porting existing security solutions from other disciplines to the edge computing environment is one of these research ideas. Scholars in the United States and elsewhere have studied mobile cloud computing and its security extensively. Roman et al. (R. Roman, 2018) investigated the security of a range of mobile edge paradigms, constructed a general cooperative security protection system, and made research suggestions. The theoretical underpinnings for edge computing security research are laid forth in these works.

ARTIFICIAL INTELLIGENCE

We are living in an era of artificial intelligence that has never seen such rapid growth (AI). Thanks to current innovations in algorithm, computing power, and big data, deep learning (Y. LeCun, 2015), the most stunning segment of AI, has made considerable advancements in a broad variety of fields, from machine vision, speech synthesis, and nlp to chess playing (e.g., AlphaGo) and robotics (Yu, 2014). As a result of these improvements, a flood of intelligent applications have come to prominence and gained massive popularity, including intelligent personal assistants, personalised shopping suggestions, video surveillance, and smart home appliances. These adaptive applications are widely recognised as significantly improving people's lives, enhancing human productivity, and increasing social efficiency. As a basic engine that advances AI development, big data has recently undergone a radical shift in source of data from megascale cloud datacenters to more ubiquitous end devices, such as smart phones and Internet-of-Things (IoT) devices. Traditionally, big data was generated and housed largely in massive datacenters. Because of the advent of mobile computing and the Internet of Things, this trend is now reversing. Clearly, the edge ecosystem will deliver a host of new AI usage scenarios by bringing vast volumes of data to AI, assuring AI's continuing growth. However, extending the AI threshold to the edge community that exists at the last mile of the Internet is exceedingly difficult because to considerations about pace, cost, and privacy. The general consensus is that IoT data bulks should be transferred to cloud datacenters for analytics (B. Heintz, 2015). However, when sending a big amount of data across a wide area network (WAN), the monetary cost and transmission time may be exorbitantly high, and privacy breaches can be a severe concern (Pu, 2015). On-device analytics, which runs AI programmes on the device to evaluate IoT data natively, is an alternative, but it has the potential to be inefficient in terms of performance and energy consumption. This is due to the fact that many AI applications demand a large amount of processing power, which much exceeds the capabilities of resource- and energy-constrained IoT devices.

Edge computing (W. Shi J. C., Edge computing: Vision and challenges, 2016), which relocates cloud services from the core of the network to network edges, which are proximal to IoT devices and data sources, has recently been proposed as a solution to the aforementioned issues. While edge nodes can be as little as a credit card to as large as a micro-datacenter with numerous server racks, the most significant feature emphasised by edge computing is physical proximity to information-generation sources. Physical proximity between computing and information-generation sources has several advantages over the typical cloud-based computing paradigm (W. Shi H. S., 2017), (Y. Mao, 2017). Low latency, energy efficiency, privacy protection, lower bandwidth use, on-premises, and context-awareness are only a small handful of them.

Artificial Intelligence

While artificial intelligence (AI) has gotten a lot of press recently, it was initially described in 1956. Simply said, AI is a method for creating intelligent computers that can do jobs in the same way people do. This term has a broad meaning that might include anything from Apple Siri to Google AlphaGo, as well as future technologies. Human intelligence-related behaviours that AI systems frequently replicate include planning, learning, reasoning, problem-solving, knowledge representation, vision, motion, and manipulation, as well as, to a lesser extent, social intelligence and creativity. Artificial intelligence has risen, fallen, and risen again during the last 60 years. Deep learning, an approach that has attained human-level accuracy in several fascinating areas, was partly responsible for the recent surge of AI after the 2010s.

Deep Neural Networks and Deep Learning

Machine learning (ML) is a powerful tool for AI to accomplish its goals. Decision trees, K-means clustering, and Bayesian networks are just a few of the machine learning approaches that have been created to teach machines to make classifications and predictions based on real-world data. Deep learning has exceeded all existing machine learning methods in a variety of tasks, including photo classification and face recognition, by using artificial neural networks (ANNs) (D. Svozil, 1997) to learn the deep representation of data.

The model is dubbed a deep neural network because the ANN used by deep learning models often consists of a series of layers (DNN).

A DNN is made up of layers of neurons, each of which may create nonlinear outputs based on input data.

The data is received by the input layer neurons, who then send it to the intermediary layer (also known as the hidden layer). The intermediate layer's neurons create weighted sums of the input data, which are then output using specific activation functions (e.g., tanh), and the outputs are then transported to the output layer.

The final results are displayed in the output layer. DNNs may learn high-level characteristics with more sophisticated and abstract layers than a traditional model, enabling for high-accuracy job inference. Three popular DNN topologies are multilayer perceptrons (MLPs), convolution neural networks (CNNs), and recurrent neural networks (RNNs).

The MLP model, which is made up of a series of totally connected layers (R. Collobert, 2011), is the most basic DNN. Convolution layers in CNN models, unlike fully connected layers in MLPs, employ convolution processes to extract fundamental properties from input. CNN models are the most common for computer vision tasks such as image classification (e.g., AlexNet (A. Krizhevsky, 2012), VGG network (Zisserman, 2014), ResNet (K. He, 2016), and MobileNet (al., 2017)) because they can capture the high-level representation of the input data by applying various convolutional filters (e.g., Fast R-CNN (H. Mao, 2018), YOLO (J. Redmon, 2016), and SSD (W. Liu et al., 2016)). Another type of DNN is the RNN model, which uses sequential data feeding.The cell, which is made up of layers, is the fundamental unit of RNN, and a sequence of cells enables RNN algorithms to be processed continuously. RNN models are used in natural language processing for a variety of applications, including language modelling, machine translation, question answering, and document classification.

Deep learning is a cutting-edge artificial intelligence technology as well as a resource-intensive activity that accommodates itself well to edge computing.

From deep learning to model training and inference, there's a lot to learn.

Every neuron in a DNN layer has a weight vector proportional to the size of the layer's input data.

The weights of a deep learning model must, of course, be fine-tuned through a training approach.

During the training phase of a deep learning model, the weight values are typically chosen at random. The task result is then represented by the last layer's output, and the correctness of the results is assessed using a loss function that calculates the error rate (e.g., root-mean-squared error) between both the results and the true label. The weights of each neuron in the model are updated using an optimization approach such as SGD (L. Bottou, 2010), and the gradients of the loss function is computed. The backpropagation mechanism (D. E. Rumelhart), (Y. Chauvin and D. E. Rumelhart, 2013) sends the error rate back through the whole neural network, and the weights are modified based on the gradient and learning rate. By feeding in a huge amount of training sets and repeating the process until the error rate falls below a certain threshold, a very precise deep learning model may be created. DNN models are deduced after training. For example, the DNN is given a large number of training samples to learn how to recognise a picture, and then inference takes real-world photos as inputs and quickly draws predictions/classifications. The training operation employs both the feed-forward and backpropagation procedures. The feed-forward and backpropagation processes are both used in the training operation. The feedforward method is used for inference, which means that the input from the actual world is processed through the whole neural network and the model produces the prediction.

Models of Deep Learning That Are Popular

1) A Convolutional Neural Network
2) Recurrent Neural Network (RNN)
3) Generative Adversarial Network
4) Deep Reinforcement Learning (DRL)

EDGE INTELLIGENCE

What is Edge Intelligence and Edge AI?

A new study field called "Edge Intelligence" or "Edge AI" has emerged as a result of the merging of Edge Computing and AI. Edge Intelligence uses widely available edge resources to fuel AI applications rather than depending solely on the cloud. While the name Edge AI or Edge Intelligence is relatively new, activities in this area date back to 2009, when Microsoft developed an edge-based prototype to facilitate mobile voice command recognition.

Despite the fact that the field is still in its early stages of development, there is no precise definition for edge intelligence. Edge Intelligence is currently defined as "the paradigm of running AI algorithms locally on an end device, with data (sensor data or signals) produced on the device" by most organizations and publications.

Edge AI and Edge Intelligence are well-known research and commercialization topics. Edge AI has lately gotten a lot of interest due to the superiority and need of executing AI applications on the edge.

The Cloud Isn't Enough to Run Deep Learning Apps

Due to the enormous advantages of deep learning in the disciplines of Computer Vision (CV) and Natural Language Processing, artificial intelligence and deep learning-based intelligent services and apps have revolutionised many parts of people's life (NLP).

The existing cloud computing service architecture, however, is insufficient to give artificial intelligence to every individual and company in any location due to efficiency and latency difficulties.

Due to the following factors, there are only a limited number of intelligent services available for a larger variety of application scenarios, such as smart factories and cities, facial recognition, medical imaging, and so on.

- Cost: Deep learning model training and inference in the cloud necessitates huge data transmission from devices or users. This necessitates a significant quantity of network bandwidth.
- Latency: The time it takes to access cloud services is rarely guaranteed, and it may be insufficient for many time-sensitive applications.
- Reliability: For linking consumers to services, most cloud computing applications rely on wireless communications and backbone networks. Intelligent services must be very dependable in various industrial contexts, even when network connections are broken.
- Privacy: Deep Learning frequently includes a tremendous quantity of sensitive information. Concerns about AI privacy are crucial in sectors like smart homes and smart cities. Even the transfer of sensitive data may be impossible in some instances.

Edge computing is supposed to overcome many of these difficulties because it is closer to consumers than the cloud.

The Benefits of Bringing Deep Learning to the Periphery

Because there is a clear overlap between AI and edge computing, it makes sense to combine the two. To fully realise the value of data created at the network edge, AI is required. Edge computing may also thrive in more complex data and application environments.

Deep learning calculations are intended to be pushed as far as feasible from the cloud to the edge via edge intelligence. This allows for the creation of a variety of distributed, low-latency, dependable, and intelligent services.

The following are some of the benefits of putting deep learning to the edge:

- Deep Learning services are distributed near to the requesting users, resulting in low latency. The latency and cost of transmitting data to the cloud for processing are dramatically reduced as a result.

- The raw data necessary for deep learning services is stored locally on the edge devices or on the user devices themselves, rather than on the cloud, which improves privacy.
- Increased Reliability: Deep learning computation is more dependable with a decentralised and hierarchical computing architecture.
- Edge computing can support the wider use of deep learning across sectors and increase AI adoption by providing richer data and application scenarios.
- Commercialization: Deep learning services that are diverse and useful extend the commercial value of edge computing and speed up its implementation and growth.

Using resources at the network edge, close to data sources, to unleash deep learning services has emerged as a preferred approach. As a result, edge intelligence strives to make the implementation of deep learning services employing edge computing as simple as possible.

The Key Infrastructure for AI Democratization Is Edge Computing

Many digital goods or services in our everyday lives have seen remarkable success thanks to AI technologies (e-commerce, service recommendation, video surveillance, smart home devices, etc.). In addition, AI is a major driving force behind new inventive frontiers including self-driving vehicles, intelligent finance, cancer diagnostics, smart cities, intelligent transportation, and medical research.

Leaders in AI are pushing to allow a wider range of deep learning applications and to push the boundaries of what is achievable based on these examples. As a result, prominent IT corporations have announced AI democratisation or ubiquitous AI as a goal, with the ambition of "creating AI for every individual and every company worldwide."

As a result, AI has to get "closer" to people, data, and end devices. Edge computing is clearly more capable than cloud computing in achieving this goal:

- Edge servers, as opposed to cloud data centres, are closer to people, data sources, and devices.
- Edge computing is more inexpensive and accessible than cloud computing.
- Edge computing, rather than cloud computing, has the ability to enable a wider range of AI application scenarios.
- Edge computing is a natural facilitator for ubiquitous AI because of these benefits.

Machine Learning Algorithms at the Network Edge

Deep learning technology is enabled by the tremendous amount of data and recent developments in artificial intelligence (AI). Edge Intelligence allows machine-learning algorithms to be deployed at the network's edge.

The main motive for bringing learning to the edge is to give quick access to the massive amounts of real-time data created by edge devices for AI model training and inference, giving the devices human-like ability to respond to real-time events.

On-device analytics use AI programmes to process data locally on the device. Because many AI applications demand a large amount of processing power, the capability of resource- and energy-constrained edge devices is severely limited. As a result, Edge AI has a number of issues, including a lack of performance and energy efficiency.

Edge Intelligence at Various Levels

Because of the significant resource consumption of the training phase, most Edge Intelligence ideas focus on the inference phase (running the AI model) and assume that the AI model is trained in cloud data centres.

Edge Intelligence, on the other hand, takes advantage of all available data and resources across the hierarchy of end devices, edge nodes, and cloud data centres to improve the overall performance of training and inferencing a Deep Neural Network model.

As a result, Edge Intelligence does not necessitate fully training or inferring the deep learning model at the edge. As a result, cloud-edge scenarios involving data offloading and co-training exist.

Because the ideal setting of Edge Intelligence is application-dependent and is established by evaluating several variables such as latency, privacy, energy efficiency, resource cost, and bandwidth cost, there is no "best-level" in general.

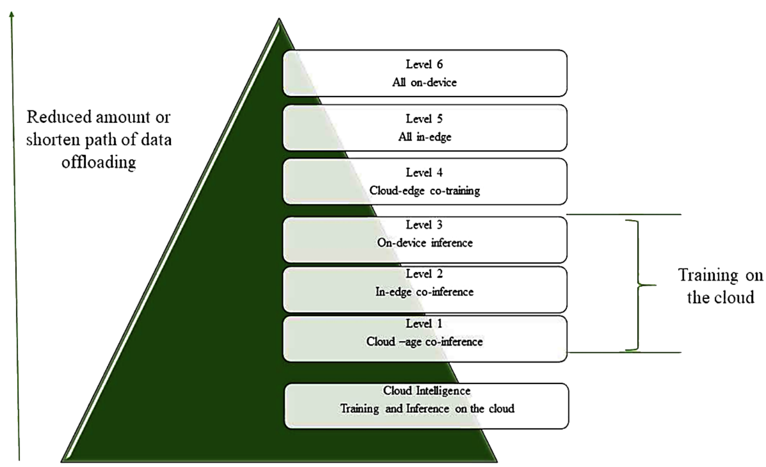
- Cloud Intelligence is the totally cloud-based training and inference of AI models.
- On-device Inference involves cloud-based AI model training as well as entirely local AI inference on the device. No data would be offloaded using on-device inference.
- All On-Device does AI model training and inference entirely on-device.

Data offloading transmission latency is decreased, data privacy is increased, and cloud resource and bandwidth costs are lowered by moving jobs to the edge. However, this comes at the expense of higher energy consumption and edge computational delay.

On-device Inference is a viable technique for a variety of on-device AI applications, and it has been shown to be optimally balanced for a variety of use scenarios. Federated Learning is built on the foundation of on-device model training.

To enhance their accuracy, Deep Learning On-Device Inference at the Edge AI models, especially Deep Neural Networks (DNNs), require larger-scale datasets. This means that computing prices will skyrocket, since Deep Learning models' superior performance necessitates high-end gear. As a result, deploying them to the edge is challenging, owing to resource restrictions.

Figure 3. Levels of Edge Intelligence



As a result, large-scale deep learning models are commonly deployed in the cloud, with end devices simply sending data to the cloud and waiting for deep learning inference results. The cloud-only inference, on the other hand, restricts the widespread usage of deep learning services:

- Latency of Inference It cannot, for example, meet the latency requirements of real-time applications such as real-time detection with severe latency requirements.
- Privacy. Cloud-based inference systems have significant drawbacks in terms of data security and privacy protection.
- Deep learning services frequently use edge computing to overcome these issues. As a result, deep learning models must be tailored to the resource-constrained edge. Deep learning applications, on the other hand, must be carefully designed to strike a balance between inference accuracy and execution latency.

APPLICATIONS OF EDGE INTELLIGENCE

As new developing technologies such as IoT and Big Data mature, more systems begin to integrate sensors for real-time data collection, which must also be analysed in real time or with time criticality in order to identify crises or important occurrences. To enhance the view of the context, the number of sensors for data collection will continue to grow, resulting in a higher volume of data (Hass, 2019). Due to the bottleneck in uploading and downloading data from all of the sensors, standard Cloud Computing will not be able to keep up with the rate at which data is created. This is when Edge Computing delivers a cost-effective solution by processing data at the network's edge (Seraphin B Calo, 2017)- (Hasibur Rahman, 2019).

The application of computer vision Deep Learning is the most cutting-edge invention for image processing (characterization and location), which is a vital task in computer vision. Sensors such as reconnaissance cameras, season of light cameras, and movement identifiers, among others, generate data for these exercises at the organization's edge. In certain circumstances, these sensors now include DL capabilities. When information from these sensors is sent to the cloud to be analysed, the number of edges processed by the (Zhenqiu Huang, 2018) framework per second decreases. Meanwhile, it may cause security concerns, especially if the material contains sensitive data. Another reason to use Edge Computing for this activity is the bottleneck caused by the massive amount of data being transmitted from these sensors. Edge AI has been used in PC vision applications such as "Vigil" (Tan Zhang, 2015). This structure is made up of a network of remote cameras that assess data at the organization's edge. These edge hubs offer information regarding outlines that should be investigated further. Unlike the traditional way of downloading the raw data to be inspected, Vigil reduces data transmission utilisation by preprocessing the images at the edge. Another concept is VideoEdge (Chien-Chun Hung, 2018), which has a multi-tiered architecture to enable Edge and Cloud to collaborate to modify responsibilities while keeping transfer speed requirements low.

Handling of everyday language For Natural Language Processing (NLP) tasks (Tom Young, 2018), such as discourse recognition (Apple, 2017), machine interpretation (Yonghui Wu, 2016) (Guillaume Lample, 2016), and element recognition (Guillaume Lample, 2016), ML, specifically DL, has become quite likely the most unmistakable way. Following the constraints of Edge Computing, ML believes dormancy to be reduced to several milliseconds. Voice colleagues from other companies, such as Amazon's Echo Dot, Google's "Alright Google," and Apple's Siri, are examples of this usage. While some of the data collected by these devices is processed in the cloud, certain critical functions should be done at the organization's edge. This is demonstrated by the watchword finding for framework actuation. The remainder of the message is sent to the cloud to be determined, and the outcome is returned if this watchword is located. In Siri, two phoney brain organisations (ANNs) detect watchwords (Assistant., 2017). The first, a smaller DNN in low-power mode, examines all of the data received by the framework to determine if the second one should be launched. The primary processor concludes that the catch has been detected once the second DNN, which is a more profound DNN than the first, is activated. To adapt to low-asset devices at the organization's edge, catch-and-release computations should be expanded further. Pruning and quantization methods may be included into these DNNs to combine them for edge devices. In light of these and other approaches, Microsoft specialists created a more modest RNN that only requires 1 kB of RAM to detect catches (Aditya Kusupati, 2019). Inactivity is a major concern for other NLP exercises due to the complexity to be reduced.

The Internet of Things (IoT): Information collected by some IoT sensors necessitates an independent cognizance gradually for use cases like as astute lattices, astute metropolitan regions, brilliant autos, or medical care. DL has been shown to provide positive results in a few studies on these topics, such as movement recognizable evidence (Zhenqiu Huang K.-J. L.-L.-S., 2018)- (Y. Zhao A. Zhang S. Yao, 2017) and person on foot recognition (Wang W. O., 2013). It is often necessary to entangle the information obtained from a few sensors, which may need to be preprocessed before being merged, in order to provide accurate results. Various structures, such as DeepSense (Y. Zhao A. Zhang S. Yao, 2017), WuKong edge system (Zhenqiu Huang K.-J. L.-L.-S., 2018), or NeuroPilot (Tung-Chien Chen, 2019), operate with this by exploiting the spatio-fleeting correlation of input from various sensors.Another option that has to be investigated is compressing the DL models to fit into IoT devices with limited computational and battery capabilities.

Another option to consider is fitting the DL models onto IoT devices with limited computational and battery capacity.

Simultaneously, by applying Edge Computing for IoT applications, it is possible to handle information security challenges, as it may contain sensitive data or public information from highways, cities, and other locations where the device proprietor is unable to share the data (Jeans, 2019). As a result, Edge AI may be able to address the issue of security while also reducing transmission speed consumption (which is critical when cameras or muddled sensors are involved). An example of this application is the framework proposed by (Muhammad Sharjeel Zareen, 2019), in which Edge Intelligence was used with respect to brilliant dwellings, with excellent results that offer new lines of investigation concerning its application in various scenarios, for example, clever enterprises or smart urban areas.

CONCLUSION

In light of the evolving requirements of distributed computing environments, several strategies have been developed to facilitate job splitting at various geographic locations. Job separation must be specified explicitly in the majority of circumstances. The key difficulty when releasing computing jobs to edge nodes is allowing the system to split work independently without having to identify the location or capabilities of the edge nodes. Furthermore, the system's intrinsic need for flexibility when building a computational pipeline mandates the creation of planners that allow segmented workloads to be routed to edge nodes. Workflows are commonly utilised in circumstances that need the utilisation of resources from outside the cloud. A workflow is the appropriate solution if the input data has to be gathered from a secure database and processed on a public cloud. Advancement tools and software architectures for data-intensive processes have spurred a lot of scientific study, as previously indicated. The inclusion of edge nodes in the scenario poses a difficulty, but it also provides an opportunity to investigate the creation of toolkits and software frameworks that can effectively implement general-purpose computing edge nodes into distributive system scenarios. It's also worth noting that the majority of scientific techniques are used in astronomy and bioinformatics. On the other side, edge analytics offers its own set of use cases centred on user-driven apps.As a result, traditional ways to characterising edge applications may fall short. The programming paradigm created to harness the capabilities of edge computing must allow data and task-level parallelism, as well as workload execution on several hierarchical hardware levels. In order to achieve such a programming paradigm, the necessary programming language must take into account the workflow's potential and capacity, as well as numerous hardware aspects. The system may encounter situations where the edge nodes are vendor-specific, and any architecture designed to support such a process must account for this. The model's complexity is considerably increased as a result of this situation.

For massive implementation of artificial intelligence on integrated edge devices, edge intelligence is viewed as the way of the future. When contrasted with cloud processing, it has the potential to deliver more security and cheaper costs while maintaining performance, as well as new features for both businesses and consumers. As a consequence of the preceding discussion, it is clear that not only the computer architecture and base processor designs, but also the effectiveness of machine learning algorithms, must be improved.

REFERENCES

Aditya Kusupati, M. S. (2019). *Fastgrnn: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network.* arXiv:1901.02358

Alrawais, A., Alhothaily, A., Hu, C., & Cheng, X. (2017). Fog computing for the Internet of Things: Security and privacy issues. *IEEE Internet Computing*, *21*(2), 34–42. doi:10.1109/MIC.2017.37

Ansari, X. S. (2016). Edge IoT: Mobile edge computing for the Internet of Things. *IEEE Communications Magazine*, *54*(12), 22–29. doi:10.1109/MCOM.2016.1600492CM

Apple. (2017). *Deep Learning for Siri's Voice: On-Device Deep Mixture Density Networks for Hybrid Unit Selection Synthesis*. Retrieved from https://machinelearning.apple.com/2017/08/06/sirivoices.html

Ashton, K. (2009). That "Internet of things" thing. *RFiD J.*, 97–114.

Assistant, A. (2017). *Hey Siri: An On-Device DNN-Powered Voice Trigger for Apple's Personal*. Retrieved from https://machinelearning.apple.com/2017/10/01/heysiri.html

Aujla, N. K. (2018). Optimal decision making for big data processing at edge-cloud environment: An SDN perspective. *IEEE Trans. Ind. Informat., 14*(2), 778–789.

Becvar, P. M. (2017). Mobile edge computing: A survey on architecture and computation offlfloading. *IEEE Commun. Surveys Tuts.*, *19*(3), 1628–1656. doi:10.1109/COMST.2017.2682318

Biron, J. F. (2016). *Foundational Elements of an IoT Solution*. O'Reilly Media Inc.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent, *Proc. COMPSTAT*, 177–186. 10.1007/978-3-7908-2604-3_16

Branch, L. I. (2017). Pilot study on application of MEC local shunting service. *Modern Inf. Technol.*, *1*(3), 65–67.

Chauvin, Y., & Rumelhart, D. E. (2013). *Backpropagation: Theory, Architectures, and Applications*. Psychology Press. doi:10.4324/9780203763247

Chen, . (2019). NeuroPilot: A cross-platform framework for edge-AI. *2019 IEEE International Conference on Artiicial Intelligence Circuits and Systems (AICAS)*, 167-170.

Chen, W. D. (2012). Integration of workflflow partitioning and resource provisioning. *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID 2012)*, 764–768. 10.1109/CCGrid.2012.57

Chien-Chun Hung, G. A. (2018). Videoedge: Processing camera streams using hierarchical clusters. *IEEE/ACM Symposium on Edge Computing (SEC)*, 115-131. 10.1109/SEC.2018.00016

Collobert, J. W. (2011). Natural language processing (almost) from scratch. *Learn. Res.*, *12*, 2493–2537.

Cox, P. (2011). *Mobile cloud computing devices, trends, issues, and the enabling technologies, developerWorks*. IBM.

Ding, C. T., J. N. (2019). Edge computing: Applications, state-of-the-art and challenges. *Zte Technol.*, *25*(3), 2–7.

Escamilla-Ambrosio, A. R.-M.-A.-B.-R. (2017). Distributing Computing in the Internet of Things: Cloud, Fog and Edge Computing Overview. *NEO*, 87-115.

Garcia-Lopez, P. M. (2015). Edge-centric computing: vision and challenges. *ACM SIGCOMM Comput. Commun. Rev.*, 37–42.

Ghafarian, T. J., & Javadi, B. (2015). Cloud-aware data intensive workflflow scheduling on volunteer computing systems. *Future Generation Computer Systems*, *51*, 87–97. doi:10.1016/j.future.2014.11.007

Goodfellow, I. (2014). Generative adversarial nets. *in Proc. Advances in Neural Information Processing Systems*, 2672–2680.

Gubbi, J. B., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660. doi:10.1016/j.future.2013.01.010

Guillaume Lample, M. B. (2016). *Neural architectures for named entity recognition.* arXiv:1603.01360.

Hasibur Rahman, R. R. (2019). The role of mobile edge computing towards assisting IoT with distributed intelligence: a smartliving perspective. Mobile Solutions and Their Usefulness in Everyday Life, 33-45.

Hass, R. (2019). *What's Powering Artiicial Intelligence?* ARM.

He, K. X. Z. (2016). Deep residual learning for image recognition. *Proc. IEEE CVPR*, 770–778.

Heintz, B. A. C. (2015). Optimizing grouped aggregation in geo-distributed streaming analytics. *Proc. ACM HPDC*, 133–144. 10.1145/2749246.2749276

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735 PMID:9377276

Internet Edge Solution Overview. (2010). Academic Press.

Jeans, D. (2019). *Related's Hudson Yards: Smart City or Surveillance City?* Retrieved from https://therealdeal.com/2019/03/15/hudsonyards-smart-city-or-surveillance-city/

Kang, J., Yu, R., Huang, X., & Zhang, Y. (2018). Privacy-preserved pseudonym scheme for fog computing supported Internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *19*(8), 2627–2637. doi:10.1109/TITS.2017.2764095

Kotis, K. K. (2012). Semantic interoperability on the web of things: the semantic smart gateway framework. *Proceedings of the IEEE Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, 630–635. 10.1109/CISIS.2012.200

Krizhevsky, A. I. S. (2012). Imagenet classification with deep convolutional neural networks, *Proc. NIPS*, 1097–1105.

Le, Y., & Cun, Y. B. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539 PMID:26017442

Liu, W. (2016). SSD: Single shot multibox detector. *Proc. Eur. Conf. Comput. Vis.*, 21–37.

Lu, F., Hu, J., Yang, L. T., Tang, Z., Li, P., Shi, Z., & Jin, H. (2020). Energyeffificient traffific offlfloading for mobile users in two-tier heterogeneous wireless networks. *Future Generation Computer Systems*, *105*, 855–863. doi:10.1016/j.future.2017.08.008

Mao, C. Y. (2017). A survey on mobile edge computing: The communication perspective. *IEEE Commun. Surveys Tuts., 19*(4), 2322–2358.

Mao, H., Yao, S., Tang, T., Li, B., Yao, J., & Wang, Y. (2018). Towards real-time object detection on embedded systems. *IEEE Transactions on Emerging Topics in Computing*, *6*(3), 417–431. doi:10.1109/TETC.2016.2593643

Mazhelis, O. W. (2013). *Internet-of-things market, value networks, and business models: State of the art report*. University of Jyvaskyla.

Meurisch, C. S. (2015). Upgrading wireless home routers for enabling large-scale deployment of cloudlets. *International Conference on Mobile Computing, Applications, and Services*, 12–29. 10.1007/978-3-319-29003-4_2

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. doi:10.1038/nature14236 PMID:25719670

Mouradian, D. N. (2018). A comprehensive survey on fog computing: State-of-theArt and research challenges. *IEEE Commun. Surveys Tuts., 20*(1), 416–464.

Muhammad Sharjeel Zareen, S. T. (2019). Artiicial Intelligence/ Machine Learning in IoT for Authentication and Authorization of Edge Devices. *2019 International Conference on Applied and Engineering Mathematics (ICAEM)*, 220-224. 10.1109/ICAEM.2019.8853780

Nadembega, A. A. S. (2016). Mobility prediction modelbased service migration procedure for follow me cloud to support QoS and QoE. *Proc. IEEE Int. Conf. Commun. (ICC)*, 1–6.

OpenFog reference architecture for fog computing. (2017). *OpenFog Consortium.*

Patierno, P. (2016). Hybrid IoT: On fog computing, gateways, and protocol translation. *DZone/IoT Zone.*

Pu, Q. (2015). Low latency geo-distributed data analytics. *Proc. ACM SIGCOMM*, 421–434.

Redmon, J. S. D. (2016). You only look once: Unified, real-time object detection. *in Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, 779–788.

Reinders, J. (2007). *Intel Threading Building Blocks: Outfifitting C++ for Multi-core Processor Parallelism*. O'Reilly.

Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, *78*, 680–698. doi:10.1016/j.future.2016.11.009

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). G. E. (n.d.). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. doi:10.1038/323533a0

Satyanarayanan, M. (2017). The emergence of edge computing. *Computer*, *50*(1), 30–39. doi:10.1109/MC.2017.9

Seraphin, B., & Calo, M. T. (2017). Edge computing architecture for applying AI to IoT. *2017 IEEE International Conference on Big Data (Big Data)*, 3012-3016.

Shi, H. S. (2017). Edge computing-an emerging computing model for the Internet of everything era. *J. Comput. Res. Develop.*, *54*(5), 907–924.

Shi, X. Z. (2019). Edge computing: State-of-the-art and future directions. *J. Comput. Res. Develop.*, *56*(1), 1–21. doi:10.1360/crad20070101

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637–646. doi:10.1109/JIOT.2016.2579198

Svozil, D., Kvasnicka, V., & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, *39*(1), 43–62. doi:10.1016/S0169-7439(97)00061-0

Szegedy, C. (2015). Going deeper with convolutions. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, 1–9.

Tang, W. J. (2014). Data-aware resource scheduling for multicloud workflflows: a fifine-grained simulation approach. *2014 IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom)*, 887–892.

The Internet of Things reference model. (2014). *Internet of Things World Forum,* 1–12.

Tom Young, D. H. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine, 13*, 55-75.

Uckelmann, D. H. (2011). An architectural approach towards the future Internet of things. Architecting the Internet of Things, 1–24.

Varghese, B. W. (2016). Challenges and opportunities in edge computing. *IEEE International Conference on Smart Cloud (SmartCloud)*, 20–26. 10.1109/SmartCloud.2016.18

Varghese, B. W. (2017). Feasibility of fog computing. *arXiv*.

Wang, D. Z. (2017). Mobile network edge computing and caching technology. *Railway Comput. Appl.*, *26*(8), 51–54.

Wang, R., Li, M., Peng, L., Hu, Y., Hassan, M. M., & Alelaiwi, A. (2020). Cognitive multi-agent empowering mobile edge computing for resource caching and collaboration. *Future Generation Computer Systems*, *102*, 66–74. doi:10.1016/j.future.2019.08.001

Wang, W. O. (2013). Joint deep learning for pedestrian detection. *Proceedings - IEEE International Conference on Computer Vision*, 2056–2063.

Wang, X. H. (2018). Edge computing technology: Development and countermeasures. *Chin. J. Eng. Sci.*, *20*(2), 20. doi:10.1016/j.jes.2017.11.002

Watson, I. (2016). *The power of analytics at the edge*. IBM.

Weiser, M. G., Gold, R., & Brown, J. S. (1999). The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal*, *38*(4), 693–696. doi:10.1147j.384.0693

Werbos, P. J. (1990). Backpropagation through time: What it does and how to do it. *Proceedings of the IEEE*, *78*(10), 1550–1560. doi:10.1109/5.58337

Yonghui Wu, M. S. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv*.

Yin, D. K. (2011). A data-aware workflflow scheduling algorithm for heterogeneous distributed systems. *International Conference on High Performance Computing and Simulation (HPCS) IEEE*, 114–120.

Yu, L. D. (2014). Deep learning: Methods and applications. *Found. Trends Signal Process.*, *7*(3–4), 197–387.

Zha, Z. M., F. L. (2018). Edge computing: Platforms; Applications and challenges. *J. Comput. Res. Develop.*, *55*(2), 327–337.

Zhang, T. A. C. (2015). The design and implementation of a wireless video surveillance system. *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 426-438. 10.1145/2789168.2790123

Zhao, Zhang, & Yao. (2017). DeepSense: A uniied deep learning framework for time-series mobile sensing data processing. *Proc. 26th Int. Conf. World Wide Web*, 351-360.

Zhenqiu Huang, K.-J. L.-L.-S. (2018). Building edge intelligence for online activity recognition in service-oriented IoT systems. *Future Generation Computer Systems*, *87*, 557–567. doi:10.1016/j.future.2018.03.003

Zhou, H. (2013). *The Internet of Things in the Cloud: A Middleware Perspective*. CRC Press.

Zisserman, K. S. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv*.

Chapter 5
Disease Analysis and Prediction Using Digital Twins and Big Data Analytics

Rajagopal R.
Narsimha Reddy Engineering College, India

Karthikeyan P.
https://orcid.org/0000-0001-8977-5520
National Chung Cheng University, Taiwan

Menaka E.
Vivekanandha College of Engineering and Technology for Women, India

Karunakaran V.
Jain University (Deemed), India

Harshavaradhanan Pon
Vellore Institute of Technology, Bhopal, India

ABSTRACT

The data generated by the big data-based clinical need analysis plays a key role in improving the consideration feature, decreasing waste and blunder, and reducing treatment expenses. The use of big data analytics (BDA) techniques for analyzing disease and predictions is discussed in this investigation. This precise survey of writing means to decide the extent of BDA in disease analysis and difficulties in treatment in the medical filed. Further, this study has discussed the comparative analysis of heart diseases, predictions using BDA techniques, predicting of breast cancer, lung cancer, and brain diseases. Digital twins will be key to delivering highly personalized treatments and interventions. Intelligent digital twins, combining data, knowledge, and algorithms (AI), are set to revolutionise medicine and public health.

DOI: 10.4018/978-1-6684-5722-1.ch005

INTRODUCTION

BDA plays an essential role in healthcare to improve healthcare service for humans. Big data analytics is used to analyse semi-structured and unstructured information to investigate helpful information. Presently multi day's numerous interpersonal interactions, clients share their health-associated medical information connected data on the web. Such health-related data can use the forecast sicknesses. Ailments like asthma, high/low blood weight, and diabetes are the most popular and expensive ceaseless conditions in the world, which cannot be relieved. Anyway, precise and timely observation information can control illnesses.

Digital twins will be an essential part of the process, even more so when you factor in the actionable (and proactive) insights that can be gained from their integration. Moreover, while the reaction curve of a particular medication could find commonality among patients, simply being able to learn from previous experiences will be invaluable. Analytics is the procedure of investigation to foresee disguised examples and relationships among information. Enormous information investigation has been connected to procedure consideration conveyance and ailment investigation. Nonetheless, the investigation's degree of data acceptance and improvement is disturbed by some necessary natural problems among the enormous data universe.

Late research enables the use of vast quantities of clinical data when analyzing multimodal link information from entirely different sources. Potential analytical regions within this area could provide a critical outcome on similarly examined medicinal services conveyance square measure. In light of this, we need to determine a strategy that could give agreeable outcomes in anticipating malady patterns. Investigation strategies concentrated on different viewpoints dependent on applications and information assortment. A portion of the application includes lodging administration, advanced education, human services, information e-administration, and customer directions.

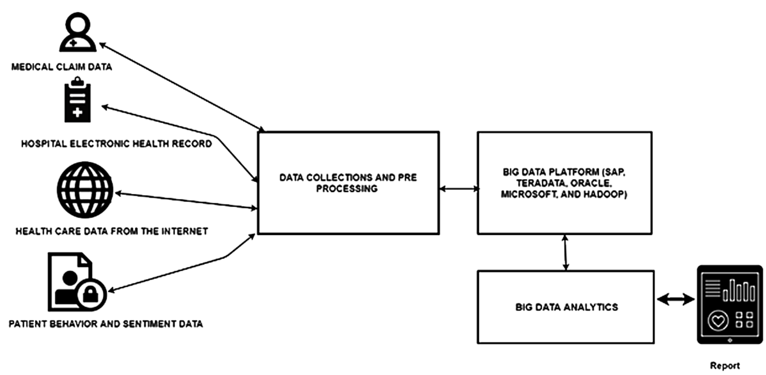
Diseases like asthma, high/low circulatory strain, and diabetes are the most pervasive and expensive endless conditions on the planet, which cannot be restored. Anyway, precise and convenient observation information can control infections. World human services problems, for example, integrative / omics data for better comprehension of harm instruments and reconciliation of genomic learning in the EHR framework for upgraded quiet end and treatment was attempted to investigate the utility of enormous biomedical information since we needed to discover the hole of where and how we can structure a calculation that will investigate and foresee the informational collections, in the various stages.

Choice tree calculations, Support vector machines (SVM), K-Nearest neighbours, K-implies, Artificial Neural networks, DBSCAN, Bayesian and so forth are used to analyze the medical data. A few systems, likewise Map Reduce systems (Spark and

Apache Hadoop), were referenced. Using the recently reviewed approach is generally made available through various instruments, libraries, and stages, such as Elastic Search, Weka, R-cran, Kibana, MOA, Python Sci-Kit, and so on. It is possible to use the Hadoop Distributed File System (HDFS) to store information. For Head Component Analysis and tensor-based methodologies, Singular Value Decomposition and highlighting extraction are helpful. Channel and wrapper-based approaches are useful. All these are systems designed to increase dimensionality. Important stages of massive data processing are IBM Cloud, Apache Hadoop, Tableau, Apache Spark Streaming and other graphical investigative methods. So there are numerous methods to deal with and break down this enormous measure of information. The test we face while putting away this tremendous measure of information is an examination, sharing, stockpiling, and so on. Fig 1 shows the architecture of disease analysis and prediction using BDA. The first step in disease analysis is to collect the data from a different source and remove the redundant data. Prepared data is passed to a big data platform which uses big data analytic techniques to generate the data.

The digital twin is a virtual or living representation of a physical object and one that draws heavily on data. That makes it a good fit for the healthcare sector, even more so when considering how much information doctors or clinicians need to do their job effectively. Integrating a digital twin into treating a disease such as Chronic obstructive pulmonary disease (COPD) allows one to visualize both simple and complex medical responses to a patient condition without administering medication to the physical aspect.

Figure 1. The architecture of disease analysis and prediction using BDA



1. LITERATURE REVIEW

Swain et al. (2016) discussed the utilization of data mining in the study of enormous medicinal services information. Considering the considerable effect of weight on the ever-growing social protection costs in the U.S.A, this examination distinguishes key statistics and way of life attributes related to grown-up stoutness. The example for this investigation was taken from the Centers for sickness management and Prevention's activity Risk issue police work Framework database. Two discerning models were worked on victimization SAS Enterprise manual labourer to form a profile of a grown-up bunch of individuals WHO area unit at risk of being fat.

Current work focuses on the two types of information in enormous medicinal knowledge science. Chen et al. (2017) proposed a new multimodal disease risk expectation-based convolutional neural network (CNN-MDRP) approach using organized and unstructured emergency clinic data. Appearing differently about two or three run-of-the-mill desire techniques, the conjecture exactitude of arranged count accomplishes 94.8% with a joined speed quicker than that of the CNN-based unimodal disease probability desire.

Asri et al. (2015) summarized AI strategies and factual investigation, and reality mining would now be able to give a public image of our individual and aggregate lives. Reality Mining is tied in with utilizing vast information to examine our conduct through cell phone sensors. New Mobile telephones are furnished with various types of sensors (for example, movement, area, and haptic). Each time an individual uses his/her cell phone, some data is gathered on how reality mining can improve the medicinal services framework. Here are a few models: Some determination data is eliminated by outstanding victimising sensors in mobile devices, like accelerometers or mouthpieces. Characteristic variations during a state of mind are conceivable from how a person talks in conversations; therefore, the finisher understands depression.

Ram et al. (2015) analyzed the plausibility of utilizing different information hotspots for anticipating the quantity of asthma-related ED visits. A starter expectation model was worked for this reason. Semi-directed order models were connected to information streams from Twitter to recognize tweets showing asthma tribulation from tweets only, including watchwords related to asthma. We additionally handled air quality information acquired from sensors, authentic electronic wellbeing records showing asthma-related visits to a crisis room, and Google search inclines, all from a similar explicit geographic zone, during a similar timeframe.

Sahoo et al. (2016) developed probabilistic information securing technique intended for the cloud-based human services framework. Also, IeCE and IaCE calculations are intended for the bury bunch connection investigation of the social insurance enormous information. An FHCP calculation is expected to envision the future prosperity condition of the patients reliant on their present prosperity status with a precision of 98%. Likewise, the cloud-based MapReduce model is used as the handling system for our extensive information investigation. It has been seen that the projected convention is often used for varied applications related to medicative services and patient assessment, like predicting coronary unwellness or ordering malignant growth gravity.

Viceconti et al. (2015) suggested a technique to predict the risk of an osteoporosis-affected woman's bone fracture, an over-the-top loss of her mineralized bone mass. The goal is to create markers that decide whether the patient will likely break over a given period (usually in the course of ten years). This is the certifiable value used to determine if a split occurs in that time.

Table 1. BigData Analytical Techniques in Disease Analysis.

BigData Analytical Technique	Disease Analysis Application
Decisions tree	Bio-signal observation for obesity
New convolution neural network	Multimodal disease risk prediction
Electronic Healthcare Predictive Analytics (e-HPA)	Recognize depression
Semi-supervised classification models	Predicting Asthma-Related Emergency
Probabilistic data acquisition method	Predict patients ' future health status
Cluster Analysis	Predicting the obesity groups for classifying high-risk clusters
Mechanistic models	Predicting the danger of bone fracture in an exceeding girl plagued by pathology

In medicines, machine learning algorithms are primarily used to evaluate and forecast diseases in the early stages. There are several life-threatening diseases in the country. Heart disease is one the hazardous infections. One of the challenging assignments in the therapeutic field is the determination of coronary illness. The conclusion of coronary illness depends fundamentally on the patients' signs, side effects and physical assessment. There are several reasons for increasing the risk of heart attack, such as lack of physical exercise, smoking, high blood pressure, alcohol, and so on. This work has discussed various classification algorithms used for analyzing and predicting heart disease.

Vivekanandan et al. (2017) used a hybrid differential evolution algorithm to predict heart disease. The experiment was conducted with UCI database reference data set. Utilizing the feed-forward neural system, the execution of the proposed framework is assessed. The framework proposed beats as far as precision, execution time, and a few ages expected to pick includes and limits essential traits.

Nahar et al. (2013) used the rule mining association to identify factors that contribute more to male and female heart disease. Three of the rule generation algorithms are used to classify these variables. One is Apriori, the other is Apriori predictive, and the fourth is Tertius. The experiment is conducted with the Cleveland dataset obtained from the UCI repository. Finally, the authors concluded that women are less risky than males for coronary heart disease.

Uyar et al. (2017) suggested training recurrent fuzzy neural networks based on the Genetic Algorithm to diagnose heart disease. The research is performed with the heart disease dataset of Irvine Cleveland University of California. The dataset consists of 297 instances; 252 instances for learning were taken for the study, and 45 instances for analysis were taken. The proposed system provides a classification accuracy of 97.78 per cent. A genetic algorithm-based recurrent fuzzy neural network trained compared to ANN-Fuzzy AHP, the proposed system is better in all aspects, such as tolerance, specificity, reliability, F-score, and precision.

Anooj (2012) has proposed weighted fuzzy rules to predict heart disease risk levels. Tautomated system to produce a weighted fuzzy policy was developed in the first phase. A fuzzy rule-based decision support model was developed in the second phase. Weighted fuzzy rules were obtained from data mining, feature collection, and the weighted function method in the first step. The proposed system has been compared with a Neural Network-based system. The proposed system is better in all aspects, like sensitivity, specificity, and accuracy. The studies were performed with the following datasets collected from the UCI database, such as the Cleveland, Hungarian, and Switzerland datasets.

El-Bialy et al. (2015) performed an observational analysis to examine the characteristics of data sets for coronary artery heart disease. In this paper, the work is divided into two phases. In the first phase, the experiment with four heart disease datasets identifies what features contribute more information during classification. In the second phase, those features are integrated, and a new derived dataset is obtained. Then classifier is trained with the derived dataset. The classifier is tested with any dataset. The classifier provides better classification accuracies.

Ilayaraja and Meyyappan (2015) used frequent item-setting methods to predict heart disease risk levels. Usually, frequent itemsets have used for market basket analysis. Authors tried frequent item sets for predicting heart disease. In this work, the frequent item sets are generated through signs and symptoms. The proposed framework is tried utilizing different databases of coronary illness. The exploratory outcome shows that the proposed framework distinguishes patients' hazard levels.

Chadha and Mayank (2016) conducted a study using data mining techniques to forecast heart disease. The experiment is conducted using three different classifiers: Naive Bayes, Artificial Neural Network and Decision Tree. The experiment's outcome suggests that Artificial Neural Networks offer better classification accuracies with less computational time than Decision Tree and Naive Bayes.

Singh et al. (2018) used data mining techniques to build an effective heart disease prediction method. The heart disease database is taken from the UCI archive in this article. The dataset consists of 303 instances, 40 per cent of the data is trained on the system, and 60 percent of the data is checked on the system. The training and testing phases are conducted through the multilayer perception neural network with backpropagation. The authors concluded from the experimental result that the proposed system provides better classification accuracies and prediction rates than existing systems. Anbarasi et al. (2010) established an improved predictor of heart disease. In this paper, the selection of the subset of features is made by a genetic algorithm. The test is performed with two conditions, one before the sub-set selection of features, and another after the sub-set selection of features. The performance is evaluated by three classifiers: Naive Bayes, Classification by clustering and Decision Tree. The experiments are carried out with a reduced dataset. The result shows that the Decision tree classifier outperformed in terms of sensitivity, specificity, precision, and accuracies than Naïve Bayes and Classification by clustering classifiers.

Amin et al. (2013) proposed a genetic neural network for predicting heart disease using risk factors. In this paper, the authors used two well-known data mining tools, a neural network and a genetic algorithm. This hybrid system provides global optimization through a genetic algorithm by assigning weights to neural networks. The hybrid system is more accurate and stable when compared with the backpropagation method, and it provides heart disease prediction accuracies of 89 per cent.

Haq et al. (2018) suggested an intelligent hybrid system to use machine learning methods to predict heart disease. This paper's study is performed with the Cleveland heart disease dataset. The experiment is carried out with two criteria one is before feature selection, and another one is after feature selection. The experiment is conducted with seven well-known classifiers such as Support Vector Machine, logistic regression, Naive Bayes, K-Nearest Neighbor, Random Forest, Artificial Neural Network, and Decision Tree with three well-known feature selection methods: LASSO, and Relief. The experimental shows logistic regression method with Relief feature selection provides better classification accuracy. The proposed method is validated with K-fold validation.

Bhatla and Jyoti (2012) proposed a novel approach for predicting heart disease using a Decision tree and Naïve Bayes classifier with fuzzy logic. The proposed system provides better classification accuracies with a reduced number of features. This paper uses fuzzy logic to find the best subset of highlights. The best subset of features was trained and tested with Naïve Bayes and Decision Tree classifiers. The experiment result clearly shows that mean absolute error was minimized in both the classifiers. Table 2, 3 and 4 lists some of the algorithms and toolkits utilized in different research papers.

This report presents the results of a qualitative study into the socio-ethical benefits and risks of using digital twins in healthcare. A digital twin combines various emerging technologies such as AI, the Internet of Things, big data and robotics, each component bringing socio-ethical issues to the resulting artefacts. Thus, the question of which of these socio-ethical themes surface in the process and how stakeholders perceive them in the field.

Table 2. List of the algorithms used in various research papers.

Algorithms	Used
Random Walks Distributed Hash Tables, Bulk Synchronous Parallel (BSP)	The random walk was built into mobile and sensor networks to tackle various issues. The hash table is used for that device to know which data resides. A collection of processors connected by a communication network compiles the BSP program.
CART, Recursive Partition Trees	Algorithms for decision-making tree.
Bayesian, K- Nearest neighbour, K-means and SVM	A survey was conducted to classify, estimate, and simulate the various computer algorithms.
MapReduce, Linear regression	The main goal was to improvise precipitation forecasting accuracy.

Table 3. Description of standard application methods and toolkits.

Category	Toolkit name	Application
Pathway analysis	Onto-Express	Breast cancer
Pathway analysis	GoMiner	Pancreatic Cancer
Pathway analysis	Cluedo	Colorectal tumours
Pathway analysis	GSEA	Diabetes
Pathway analysis	Pathway-Express	Leukemia
Reconstruction of Metabolic Networks	Recon	Drug target prediction studies
Reconstruction of Gene Regulatory Networks	Boolean methods	Cardiac differentiation
Reconstruction of Gene Regulatory Networks	ODE models	Cardiac development

Table 4. Bigdata analytics systems and tools in healthcare

Tool/ Platform	Description
The Hadoop Distributed File System(HDFS)	HDFS enables the major storage for the Hadoop pack. It segments the data into tinier parts and scatters it over the various servers or centers.
Zookeeper	Zookeeper empowers a unified base with various organizations to synchronize over an accumulation of organizations. Large applications for data analysis use these administrations to coordinate parallel plans cross-sectionally over large bunches.
PIG and PIG Latin	The language of pig writing computer programs is intended to adapt a wide scope of data. This contains two fundamental modules: the language alluded to as pig Latin, and the runtime method of execution of the pig Latin code.
MapReduce	MapReduce provides the system for sub-assignment appropriation and yields social incentives. MapReduce monitors the handling of each database or hub when undertakings are executed.
Jaql	Jaql is a pragmatic, definitive inquiry language proposed to continue huge educational lists. To empower parallel getting ready, Jaql changes over “irregular state” investigation into “low-level” questions, including MapReduce assignments.
Hive	Hive is a runtime Hadoop support building that utilises structure question language (SQL) with the Hadoop arrange.
Cassandra	Cassandra is also a scattered database structure. It is doled out as a top-level undertaking shown to manage massive data passed on across various utility servers.
Mahout	Mahout is one more apache adventure whose goal is to make accessible employments of appropriated and versatile AI counts that help tremendous data assessment on the Hadoop stage.
HBase	HBase is a section that arranges databases of the official structure that sit over HDFS. It uses a non-SQL approach.
Lucene	The Lucene adventure is used for the most part for content assessment or look and has been joined into a couple of open-source adventures. Its expansion fuses natural substance requesting and library search for use inside a java application.
Oozie	Oozie, an open-source venture, streamlines the work process and coordination among the errands.

Health care services might become more proactive and individualized thanks to digital twins. Digital twins have the ability to identify irregularities and evaluate health risks before a disease manifests or becomes symptomatic because to prediction algorithms and real-time data. It is difficult to use the current bioethical framework to capture the unique characteristics and associated ethical risks of digital twins for personalized health care services because they represent a convergence of health care, artificial intelligence, information and communication technologies (ICTs), and personalized health care.

2. ANALYZING AND PREDICTING CANCER

Despite decades of work and significant investments, the reasons for malignancy are still inadequately comprehended. Treatment arranging regularly continues through experimentation. Omic information is utilized to make an SVM order model. When another patient's information is obtained, the SVC (Support Vector Classifier) utilizes the arrangement model and patient information to group whether the patient has malignant growth. If an individual is discovered contaminated, customized medication will be proposed by the derivation motor by getting to the learning base and clinical data. SVMs have better execution in the actual application, for example, content order, picture grouping, transcribed character acknowledgement and the bio-arrangement examination. It is currently well known in information mining. SVMs are utilized to determine how to break down the information, perceive examples, arrangement and relapse investigation. When another informational collection comes, SVM maps information to that space and then sees that point on the high side of the hole. SVMs can perform straight just as nonlinear characterization.

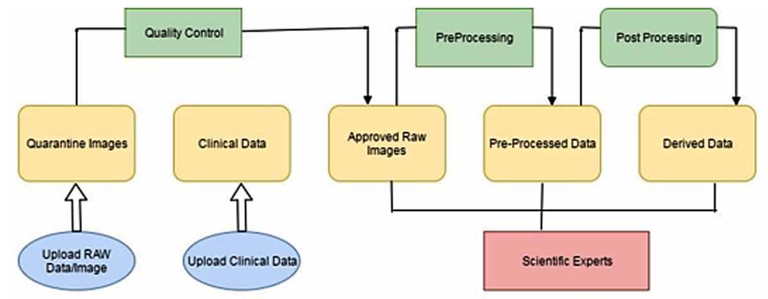
2.1 Breast Cancer

A filter that can improve the virtual screening process in the medication disclosure field. The bosom malignant growth is a perilous illness that takes an ever-increasing number of lives. To this end, we have understood huge information handling systems, such as MapReduce and Mahout. For the most part, five AI calculations intended for enormous information examination to arrange ligands into dockable and non-dockable ones. Three calculations are the best in exactness, as indicated by our involvement with the bosom malignant growth receptor. We have made an outfit of classifiers utilizing these calculations. The acquired outcomes have demonstrated that range of about 80% exactness. As we accept that more information will improve the model structure, we will move it to a more incredible group of machines that will almost certainly perform it on a higher number of ligands in a generally better execution time. Fig 2 shows the architecture of the Process of Data flow.

2.2 Lung Cancer

Lung cancer growth is the first typical disease among men from the Indian subcontinent. It is the leading killer of men failing horrendously because of any malignant growth-related torment. It is, in every case, critical to check whether your information coordinates your business destinations. On the off chance that it doesn't, many inquiries are tended to in learning disclosure from enormous information.

Figure 2. Process of Data flow



Diellza et al (2018) in their work, have made models for anticipating the survivability of separated cases utilizing the SEER chest illness dataset. Two computations, counterfeit neural framework (ANN) and C5.0 decision tree, were utilized to make desired models. C5.0 gave a precision of 93.6%, while ANN gave an exactness of 91.2%. They investigated three data mining methodologies: the Naïve Bayes, the back spread neural framework and the C4.5 decision tree. They have uncovered that the C4.5 estimation gave the best execution of 86.7% precision. Multilayer Perceptron (MLP) neural framework, Probabilistic Neural Network (PNN) and Generalized Regression Neural Network (GRNN) were related to request chest illness into two social occasions; benevolent and undermining tumour. A precision of 100% was worked on utilizing GRNN and PNN. Then again, unequivocally gained in this assessment cannot be named accurate considering that the measure of careful cases in the database was not nicely high.

2.3 The Prescription to Beat Cancer

To those occupied with Big Science®Big Data®Big Collaboration exercises, Big Technology will keep on energizing bioinformatics that will prompt the remedy to beat Cancer. Enormous Technology will empower the virtual development of the researcher to the information instead of the information to the researcher. In that capacity, Big Technology gives active vitality to the R&D malignancy network as life science and therapeutic research combine.

An efficient mapping study planned to break down and combine the utilization of Data mining systems in bosom malignant growth. Four hundred three articles distributed in the range of 2000 and 2016 were chosen and examined by year and wellspring of production, research, therapeutic assignment, exact sort and DM systems utilized. The outcomes demonstrated that countless specialists were keen on applying DM methods. (1) This sickness is the primary source of female demise worldwide and remains a significant general wellbeing challenge, and (2) DM procedures have demonstrated their adequacy and proficiency in a few fields, for example, business, instruction and examples acknowledgement. The highest number of articles on the utilization of DM in bosom malignancy has been distributed in the most recent decade, particularly since 2006. Also, the chosen articles were distributed in various wellsprings of productions representing considerable authority in the accompanying fields: software engineering, the therapeutic field and software engineering connected to prescription. Then again, most papers were named answer propositions or assessment papers. Most of those papers have a place with authentic based assessment and contextual analyses. Wisconsin database and the Mammographic Image Analysis Society (MIAS) database are the two most much of the time utilized databases in the chosen papers. For the most part, the specialists dedicated their exertion to determination, forecast and screening. At the same time, a low number of studies handled the treatment task. The executives and checking assignments were not treated in the chosen papers. NN, SVM and DT were the most now and again DM procedures utilized to improve DM models for bosom disease. Characterization and forecast were the most examined destinations of DM in bosom malignancy.

3. BRAIN DISEASES ANALYSIS

Brain diseases come in various structures. Contaminations, injury, stroke, seizures, and tumours are a portion of the actual classes of Brain diseases. The essential one is tumours, in which the dangerous development starts in the cerebrum, representing 50 percent of mind tumours. Helper or metastatic personality harmful development, in which hazardous cells have spread to the cerebrum from a tumour discovered elsewhere in the body. Twenty to thirty percent of patients with metastatic illness with harmful development that spreads from its origin make discretionary cerebrum dangerous growth. Revealing over-the-top character changes related to Alzheimer's disease (AD) is an inconvenient undertaking, mainly since individuals do not display indications of dementia until it is late. Over the previous years, neuroimaging strategies arranged for PC-based end and portrayal to empower the mechanization of therapeutic choice assistance and help clinicians perceive rationally perfect subjects in danger of growing AD. As a dynamic neurodegenerative issue, examiners investigated how AD impacts the cerebrum utilizing various frameworks: 1) picture-based strategies where, on a fundamental level, neuroimaging modalities are utilized to give early AD biomarkers, and 2) organize set up together philosophies which center concerning practical and essential character frameworks to give bits of data into how AD changes mind wiring. In this assessment, we watched out for neuro imaging-based specific strategies made for AD and fragile mental obstacle (MCI) depiction and want attempts, picked by screening all MICCAI procedures that appropriated some spot in 2010 and 2016.

The down-to-earth Magnetic Resonance Imaging (fMRI) methodology allows us to assess valuable joining by methods for evaluating the association between's intrinsic Blood-Oxygen-Level-Dependent (BOLD) signal changes of scattered cerebrum areas still. The BOLD sign is fragile to unconstrained neural development inside mind territories. Thus it will, in general, be used as a viable and non-invasive course for analysing neurological issues at the whole cerebrum level. The utilitarian system (FC), described as the transient relationship of the BOLD banner in different personality areas, can show how fundamentally confined and specific cerebrum locales collaborate. This way, the cerebrums orchestrating examination using fMRI data will give particular central focus to a robotized finding of neural or mental ailments.

Using the diagram theoretic technique, two or three analysts model the FC data as a particular system. Contrasts among standard and aggravated FC systems accomplished by despondent person assaults give essential biomarkers to understand over the top underpinnings, to the degree the topological structure and association quality. The system assessment has become a beneficial mechanical gathering for understanding the cerebral working portion and tunnelling delicate biomarkers for neural or mental insecurities. Sub-nuclear imaging helps coordinate the organization of mind tumours. Specialists use PET examinations to portray the degree of danger, choose the disease's level and perceive harmful development rehashes (Velliangiri et al 2019).

The despondent person system related to neurological issues constantly makes explicitly spatial-regular models. These models can be imaged with positron discharge tomography (PET) or single-photon spread figured tomography (SPECT) utilizing genuine radiotracers. In any case, standard quantitative PET and SPECT picture assessment estimations, for example, the institutionalized take-up worth (SUV) and non-displaceable keeping potential (BPND), are a significant part of the time assessed as midpoints over a particular area of premium (ROI). This way of thinking rejects the spatial dissipating of radiotracer complete, which might be influenced by illness inside the ROIs. There is a making assertion that better frameworks for spatial picture evaluation are required to accomplish a persistently complete sickness portrayal and improve figure and following of malady advancement. So far, the essential strategy to oversee spatial assessment regarding these destinations has depended upon the utilization of surface and shape-based picture highlights, inclining vivaciously on the making field of radiomics. Following a general radionics approach, an epic number (some spots in the extent of tens and hundreds) of various highlights are enrolled from the photographs inside pre-depicted ROIs. The highlights are then utilized as responsibilities to one of the set-ups AI models (neural nets, choice timberlands, etc.) to anticipate clinical degrees of intrigue.

Regardless, evaluations that utilise radiomic highlights are obliged in two. In the first place, PET and SPECT cerebrum imaging thinks about from time to time, have a generally minimal model size (10 to 50 subjects) showed up distinctively in connection to the measure of endeavoured highlights. The estimations of highlights may likewise rely on inside factors required to depict the highlights, picture re-trying check and ROI definition criteria. This point may accomplish an astoundingly huge model-parameter search space and incrementally the danger of model overfitting, especially if the utilized precise testing method is not eager or genuine. Second, a decently high multifaceted nature and space of the radiomic highlights upsets depiction and characteristic translation of the tainting noteworthy spatial models(Velliangiri et al 2021).

4. DIGITAL TWINS IN PERSONALISED MEDICINE

IoT has significantly aided the healthcare industry by gathering real-time data streams from linked clinical, health, and other (such as environmental) sensors and devices to facilitate communications between equipment, machines, and humans (Tao et al. 2019). It has helped make critical data available through electronic medical records, diagnostic processes, remote monitoring, and patient-generated reports. Other technologies that are blooming in parallel, led by artificial intelligence (including machine learning), which offers advanced data analytics, and cloud computing, which provides powerful, on-demand networked computational resources, which have also become crucial tools for not only processing large quantities of (IoT) data and new knowledge discovery but also for doing so in real-time. Combining all these technologies has a tremendous synergistic effect. They can produce timely and valuable insights for medical professionals and individual patients, helping them make more informed and proactive decisions. These technologies can undoubtedly serve as the backbone for shifting healthcare to a direction that focuses more on precision and preventive care. However, by adding digital twins to the combination, the vision of achieving high-quality care will be even more complete.

5. DIGITAL TWINS IN PRECISION PUBLIC HEALTH

In the previous section, we discussed the opportunities of creating digital twins for personalized medicine at length. Human digital twins in personalized medicine are constructed to represent organs or micro-structures within the body. They can be expanded to incorporate external factors, including the environment surrounding the individual and social interactions during some observed period. This would be sufficient to analyse and produce personalized care recommendations for the individual. Public health, on the other hand, tends to be more concerned with person-to-person interactions, ongoing medical conditions of patients within a community, and factors affecting health at the population level. This is because public health mainly works to promote population wellness and track, control, and prevent disease outbreaks and epidemics. However, this Technology plays an equally crucial role. It presents essential research opportunities for precision public health, especially after the lessons learned from the COVID-19 pandemic.

6. DIGITAL TWINS FOR SMART HEALTHY CITIES

Besides those above COVID-19 city-wide application examples, digital twin technology coupled with geographic information systems (GIS) can deliver much desired intelligent decision support functionality for urban and public health planners in a wide range of applications, from road traffic management (e.g., to reduce road traffic congestion, air and noise pollution, and injuries/accidents) to flood monitoring and flood situation services, in the context of intelligent, healthy cities. These city twins go beyond traditional 3D models of cities, allowing smart cities to dynamically integrate critical factors, such as time (temporal dimension, especially using accurate- and near-real-time data) and human behaviour, to better monitor conditions of interest, test various intervention scenarios *in silico*, and predict how a city system will react to changes and modifications and how its population will be impacted. Urban and public health planners can then make informed decisions to take appropriate courses of action and revisit policies to achieve desired public health and well-being outcomes (Deren et al., 2021).

7. CONCLUSION

This research discusses how BDA and digital twin applied in analyzing the disease. The use of Big Data Analytics (BDA) techniques for analyzing disease and predictions is discussed in this investigation. Medical research has been explored based on studying and predicting breast cancer, lung cancer, and brain disease. Intelligent digital twins, combining data, knowledge, and algorithms (AI), are set to revolutionize medicine and public health. Applied to medicine and public health, they can drive a much-needed radical transformation of traditional electronic health/medical records (focusing on individuals) and their aggregates (covering populations) to make them ready for a new era of precision (and accuracy) medicine and public health.

REFERENCES

Amin, S. U., Agarwal, K., & Beg, R. (2013, April). Genetic neural network based data mining in prediction of heart disease using risk factors. *IEEE Conference on Information & Communication Technologies*, 1227-1231.

Anbarasi, M., Anupriya, E., & Iyengar, N. C. S. N. (2010). Enhanced prediction of heart disease with feature subset selection using genetic algorithm. *International Journal of Engineering Science and Technology*, *2*(10), 5370–5376.

Anooj, P. K. (2012). Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University-Computer and Information Sciences.*, *24*(1), 27–40.

Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2015, June). Big data in healthcare: challenges and opportunities. In *2015 International Conference on Cloud Technologies and Applications (CloudTech)* (pp. 1-7). IEEE.

Bhatla, N., & Jyoti, K. (2012). A Novel Approach for heart disease diagnosis using Data Mining and Fuzzy logic. *International Journal of Computers and Applications*, *54*(17).

Chadha, R., &Mayank, S. (2016). Prediction of heart disease using data mining techniques. *CSI Transactions on ICT, 4*(2-4), 193-198.

Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. (2017). Disease prediction by machine learning over big data from healthcare communities. *IEEE Access: Practical Innovations, Open Solutions*, *5*, 8869–8879.

Deren, L., Wenbo, Y., & Zhenfeng, S. (2021). Smart city based on digital twins. *Computers & Urban Society*, *1*, 4.

El-Bialy, R., Salamay, M. A., Karam, O. H., & Khalifa, M. E. (2015). Feature analysis of coronary artery heart disease data sets. *Procedia Computer Science*, *65*, 459–468.

Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). *A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms*. Mobile Information Systems.

Ilayaraja, M., & Meyyappan, T. (2015). Efficient data mining method to predict the risk of heart diseases through frequent itemsets. *Procedia Computer Science*, *70*, 586–592.

Nagavci, D., Hamiti, M., & Selimi, B. (2018). Review of Prediction of Disease Trends using Big Data Analytics. *International Journal of Advanced Computer Science and Applications.*, *9*, 8.

Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, *40*(4), 1086–1093.

Ram, S., Zhang, W., Williams, M., & Pengetnze, Y. (2015). Predicting asthma-related emergency department visits using big data. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1216–1223.

Sahoo, P. K., Mohapatra, S. K., & Wu, S. L. (2016). Analyzing healthcare big data with prediction for future health condition. *IEEE Access: Practical Innovations, Open Solutions*, *4*, 9786–9799.

Singh, P., Singh, S., & Pandi-Jain, G. S. (2018). Effective heart disease prediction system using data mining techniques. *International Journal of Nanomedicine*, *13*, 121.

Swain, A. K. (2016). Mining big data to support decision making in healthcare. *Journal of Information Technology Case and Application Research*, *18*(3), 141–154.

Tao, F., Liu, W., Zhang, M., & Hu, T. (2019). Five-dimension digital twin model and its ten applications. *Computer Integrated Manufacturing Systems*, *25*, 1–18.

Uyar, K., & İlhan, A. (2017). Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks. *Procedia Computer Science*, *120*, 588–593.

Velliangiri, S., Anbarasu, V., Karthikeyan, P., & Anandaraj, S. P. (2021). Intelligent Personal Health Monitoring and Guidance Using Long Short-Term Memory. *Journal of Mobile Multimedia,* 349-372.

Velliangiri, S., Karthikeyan, P., Joseph, I. T., & Kumar, S. A. (2019, December). Investigation of deep learning schemes in medical application. In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 87-92). IEEE.

Viceconti, M., Hunter, P., & Hose, R. (2015). Big data, big knowledge: Big data for personalized healthcare. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1209–1215.

Vivekanandan, T., & Iyengar, N. C. S. N. (2017). Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Computers in Biology and Medicine*, *90*, 125–136.

Chapter 6
Knowledge Discovery Through Intelligent Data Analytics in Healthcare

Kowsalya S.
Sri Krishna Arts and Science College, India

Saraswathi S.
Sri Krishna Arts and Science College, India

ABSTRACT

The chapter aims to embed the demanding computing concepts to attain intelligent data analytics in the domain of healthcare. The targeted outcome provides the pathway to design the brainy decision support system needed to have efficient prediction with trained input patterns. The usage of IoT devices is increasing tremendously to overcome the challenges existing in handling the data related to human-relevant happenings. The volume, velocity, and variety of data are emerging newly and dominating the decision support characteristics. This scenario happens almost in all the computing fields, but more attention is expected to implement in the healthcare sector due to the existence of sensitive data. The traditional data analytics methods are deviating in the performance due to the unpredictable dynamic challenges emerging in the day-to-day operation. The efficient features of demanding computing strategies are motivated to embed together to discover crucial knowledge through intelligent data analytics.

DOI: 10.4018/978-1-6684-5722-1.ch006

INTRODUCTION

The Internet of things (IoT) is comprised of physical objects that are integrated with sensors, processing capacity, hardware, software, and other communicating components. These comprised units are connected to exchange the domain-specific generated data with other devices over the Internet networks and initiate the data analytics process.(Zhang, 2004, p.93)

The following processes are planned to merge with strategies and computing techniques to ensure efficient data handling and analytics, (A.Adhikari & J.Adhikari, 2017, pp.189-192)

- Domain-specific Data Generation
- Data Storage and Security
- Data Preprocessing and Handling
- Data Communication over the Internet
- Data Accessibility and Maintainability
- Data Analytics
- Data Extraction

The outcome will be the knowledge extracted through intelligent data analytic strategies from the hidden information of the transactional data. The objective is attained through the IoT-based architecture that consists of web-based smart devices to execute the sub-tasks like collecting the data from a different source, send to the targeted destination, and initiate the data analysis process which creates an impact in dominating the timely decision over the system process. (Maimon & Rokach, 2014, p.172). The data analysis can be carried out either using cloud storage or local storage depending upon the level of security needed in accessing the stored data.

Advantages of IoT Based Data Analytics

The current emerging computing concept internet of things offers unmeasurable services to the environment where it is implemented and used. The services are categorized as industry-specific called unique services and common services (Zobaa & Bihl, 2018, pp.203-211). The common services of IoT enabled processes are listed below as:

- Centralized Monitoring Services
- Maintenance Service based on Customer Experience
- Handling Time and Cost-Effectively.
- Yield High Productivity support Services.

- Business Model Integration Services.
- Business Decisions support Services.
- Secured Data Management Services.

Day to day organizations are rethinking adopting the methodologies that are essential to attain higher productivity with minimal risk. The business strategies need to redefine constantly to ensure the dynamic challenges are handled measurably. IoT is most implemented in the production sector, transportation of products, and utility service-based organizations such as health care, education sector, agriculture, infrastructure, and home automation industries due to the need for high security while communicating the data and motivates the organizations to think of digital transformation.

The IoT-enabled data handling architecture is greeted in the healthcare sector due to the collection of data from different sources, combined and analyzed to take timely decisions. The health care professionals need to be the accessible location of medical instruments like wheelchairs, oxygen pumps, defibrillators more proactive in handling the patient data. Some sources of data are tracking the nebulizers and other monitoring instruments.

IoT in Healthcare Sector

The secured intelligent data analysis can be done over the data collected and communicated through IoT devices by utilizing the new opportunities for healthcare professionals to monitor and handle patients or the patients to handle themselves. The knowledge discovery process can become more efficient by using wearable IoT devices holds an array of services and opportunities for healthcare professionals and their patients alike. Patient monitoring in Remote: Devices support to collect the crucial health metrics like temperature, pressure, heartbeat rate automatically and communicated for analysis by the professionals or software applications. Here, the incorporation of efficient algorithms must be useful to process the data to discover the existing fact and recommend the treatments or produce alerts. The alerts can be defined in all the monitoring services like Glucose monitoring, Heart-rate monitoring, Hand hygiene monitoring, Depression monitoring, Connected inhalers, and Ingestible sensors. (Ghosh et al., 2016, p.167)

Integration of Data Analytics in Healthcare

Nowadays, due to the emergence of digital technologies, a massive amount of healthcare data is created, collected, and stored more effectively than traditional methods. The role of Big data in healthcare is demanded to use of prescriptive, descriptive, and prescriptive analysis to acquire depth insights from captured data. The captured data can be analyzed in three aspects. They are

- Patient data
- Operational data
- Financial data

Identification of crucial ways to incorporate big data in the healthcare sector:

- Sufficient Training - Data-driven mindset
- Effective Collection and Secure Storage Mechanism
- Intelligent algorithms – Data Analytics

Big Data Excellence in Healthcare Sector

- Minimizing the Cost
- Advanced Services
- Health Tracking
- High-risk patients Guidance
- Avoiding Human Errors
- Electronic Patient Health Records
- Reduce unauthorized access and enhance security
- Real-Time Alerting or Warning
- Impact of health data for planning strategic decisions
- Efficient Prediction

Big Data Challenges in Healthcare

The research has to extend the utilization of computing performance with the support of big data evolution features like unpredictable challenges, conditions, and issues due to the massive growth of healthcare data. (Andreev et al., 2022, pp.89-95). Data analytics needs the big data computing concepts to update the changes of data presents a lot of efficient challenges in storing, analyzing, accessing, and extracting the knowledge from the massive amount of data.

The key opportunities are encountered by healthcare are as follows,

- Quality and Quantity of data
- Higher-end Data analysis

- Excellence in data analytics
- Data security and confidentiality
- Multiple sources of data

The Big data features are the main problems that are needed to be addressed immediately to ensure the performance and non-conformity of the deviation. This is a vital role to move towards the big data techniques to provide efficient medical facilities.

The beneficiaries from the use of big data in healthcare

- Providers (Clinics, Hospitals)
- Payers (Insurance)
- Patients
- Device Manufacturers
- Research and Development Cell

COMPONENTS OF IoT

There are four main components used in IoT

- **Low-power embedded systems:** Less battery consumption, high performance are the inverse factors that play a significant role during the design of electronic systems.
- **Cloud computing:** Data collected through IoT devices is massive and this data has to be stored on a reliable storage server. This is where cloud computing comes into play. The data is processed and learned, giving more room for us to discover where things like electrical faults/errors are within the system.
- **Availability of big data:** We know that IoT relies heavily on sensors, especially in real-time. As these electronic devices spread throughout every field, their usage is going to trigger a massive flux of big data.
- **Networking connection:** To communicate, internet connectivity is a must where each physical object is represented by an IP address. However, there are only a limited number of addresses available according to the IP naming. Due to the growing number of devices, this naming system will not be feasible anymore. Therefore, researchers are looking for another alternative naming system to represent each physical object.

Two Ways of Building IoT

- Form a separate internetwork including only physical objects.
- Make the Internet ever more expansive, but this requires hard-core technologies such as rigorous cloud computing and rapid big data storage (expensive).

IoT Enablers

- RFIDs: uses radio waves to electronically track the tags attached to each physical object.
- Sensors: devices that can detect changes in an environment (ex: motion detectors).
- Nanotechnology: as the name suggests, these are extremely small devices with dimensions usually less than a hundred nanometers.
- Smart networks: (ex: mesh topology).

Characteristics of IoT

- Massively scalable and efficient.
- IP-based addressing will no longer be suitable in the upcoming future.
- An abundance of physical objects is present that does not use IP, so IoT is made possible.
- Devices typically consume less power. When not in use, they should be automatically programmed to sleep.
- A device that is connected to another device right now may not be connected in another instant of time.
- Intermittent connectivity – IoT devices aren't always connected. To save bandwidth and battery consumption, devices will be powered off periodically when not in use. Otherwise, connections might turn unreliable and thus prove to be inefficient. (Pankajavalli & Karthick, 2019, p.97)

Modern Applications

- Smart Grids and energy saving
- Smart cities
- Smart homes
- Healthcare
- Earthquake detection
- Radiation detection/hazardous gas detection

- Smartphone detection
- Water flow monitoring
- Traffic monitoring
- Wearables

Real-Time Example

Suppose the electric cables used on metro train bridges are made up of copper from inside and that copper is replaced by smart copper. Smart copper is nothing but copper having sensors that are wirelessly connected to the cloud which is later connected to machines wirelessly. So in the future, if there will be any breakage or any other kind of physical or functional damage to the wire, the sensors will record and transfer relevant reports to the respective authority wirelessly via the cloud. This whole process will help the system in three ways:

- Physical labor and time can be saved in finding the problem in the wire
- Instead of changing the whole wire, only parts affected can be monitored and changed
- Daily report of wires which may include its temperature, tension, an even amount of current transferred can be received.

Even in the future if the wire is covered with snow, this can also be monitored and dotted down wirelessly.

IMPORTANCE OF IOT IN THE HEALTH SECTOR

Simultaneous Reporting and Monitoring

Remote health monitoring via connected devices can save lives in event of a medical emergency like heart failure, diabetes, asthma attacks, etc.

With real-time monitoring of the health condition in place using a smart medical device connected to a smartphone app, connected medical devices can collect medical and other required health data and use the data connection of the smartphone to transfer collected information to a physician or a cloud platform.

Center of Connected Health Policy conducted a study that indicates that there was a 50% reduction in the 30-day readmission rate because of remote patient monitoring on heart failure patients.

The IoT device collects and transfers health data: blood pressure, oxygen, blood sugar levels, weight, and ECGs.

These data are stored in the cloud and can be shared with an authorized person, who could be a physician, your insurance company, a participating health firm, or an external consultant, to allow them to look at the collected data regardless of their place, time, or device. (Kumar et al., 2019, p. 174)

End-to-End Connectivity and Affordability

IoT can automate patient care workflow with the help of healthcare mobility solutions and other new IoT technologies, and next-gen healthcare facilities.

IoT in healthcare enables interoperability, artificial intelligence machine-to-machine communication, information exchange, and data movement that makes healthcare service delivery effective.

Connectivity protocols: Bluetooth LE, Wi-Fi, Z-wave, ZigBee, and other modern protocols, healthcare personnel can change the way they spot illness and ailments in patients and can also innovate revolutionary ways of treating across different healthcare fields.

Consequently, technology-driven setup brings down the healthcare cost, by cutting down unnecessary visits, utilizing better quality resources, and improving the allocation and planning.

Data Assortment and Analysis

The vast amount of data that a healthcare device sends in a very short time owing to their real-time application is hard to store and manage if the cloud access is unavailable.

Even for healthcare professionals to acquire data originating from multiple devices and sources and analyze it manually is a tough bet.

IoT devices can collect, report, and analyses real-time information and cut the need to store the raw data. (Kumar et al., 2019, pp. 234-245)

This all can happen to overcloud with the providers only getting access to final reports with graphs.

Moreover, healthcare operations allow organizations to get vital healthcare analytics and data-driven insights which speed up decision-making and are less prone to errors.

Tracking and Alerts

The on-time alert is critical in chronic conditions. Medical IoT devices gather vital signs of any disease and transfer that data to doctors for real-time tracking while dropping notifications to people about critical parts via mobile apps and smart sensors.

Reports and alerts give a firm opinion about a patient's condition, irrespective of place and time.

It also helps healthcare providers to make well-versed decisions and provide on-time treatment.

Thus, IoT enables real-time alerting, tracking, and monitoring, which permits hands-on treatments, better accuracy, apt intervention by doctors, and improves complete patient care delivery results. (Kumar et al., 2019, p. 269)

Remote Medical Assistance

In event of an emergency, patients can contact a doctor who is many kilometers away with smart mobile apps.

With mobility solutions in healthcare, the medics can instantly check the patients and identify the ailments on the go.

Also, numerous IoT-based healthcare delivery chains are forecasting to build machines that can distribute drugs based on patients' prescriptions and ailment-related data available via linked devices. (Pankajavalli & Karthick, 2019, p.202)

IoT will improve the patient's care in the hospital. This, in turn, will cut people's expenses on healthcare.

Research

IoT healthcare applications can also be used for research purposes. It's because IoT enables us to collect a huge amount of data about the patient's illness which would have taken many years if we collected it manually.

The data thus collected can be used for the statistical study that would support the medical research.

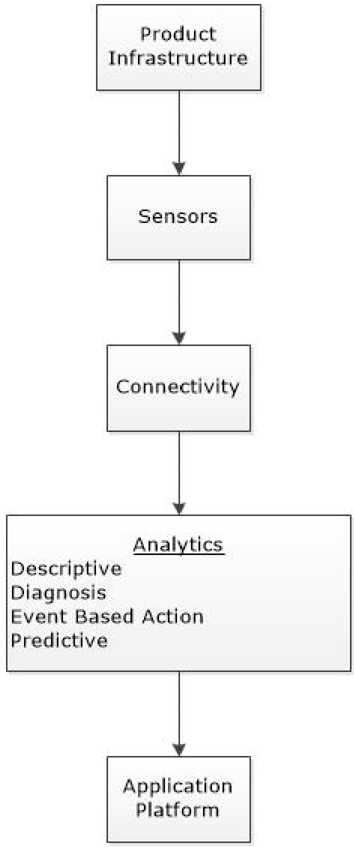
Thus, IoT doesn't only save time but also our money which would go into the research.

Thus, IoT has a great impact in the field of medical research. It enables the introduction of bigger and better medical treatments.

IoT is used in a variety of devices that enhance the quality of the healthcare services received by the patients.

Even the existing devices are now being updated by IoT by simply using embedding chips of smart hospital devices. This chip enhances the assistance and cares that a patient requires. (Andreev et al., 2022, pp.113-116)

Figure 1.



IoT Architecture

Product Infrastructure: IoT product infrastructure such as hardware/software components read the sensor's signals and display them to a dedicated device.

Sensors: IoT in healthcare has different sensors devices such as pulse-oximeter, electrocardiogram, thermometer, fluid level sensor, sphygmomanometer (blood pressure) that read the current patient situation (data).

Connectivity: IoT system provides better connectivity (using Bluetooth, WiFi, etc.) of devices or sensors from microcontroller to server and vice-versa to read data.

Analytics: Healthcare system analyzes the data from sensors and correlates to get healthy parameters of the patient and based on their analyze data they can upgrade the patient health.

Application Platform: IoT system access information to healthcare professionals on their monitor device for all patients with all details. (Maroto, 2019, pp.31-35)

Challenges in IoT

Data Security & Privacy

One of the most significant challenges in healthcare that IoT poses is data security & privacy. IoT security devices capture and transmit data in real-time.

However, most IoT devices lack data protocols and security requirements.

In addition to that, there is significant ambiguity regarding data ownership regulation with electronic devices.

All these factors make the data highly susceptible to cybercriminals who can hack into the system and compromise the Personal Health Information (PHI) of both patients as well as doctors.

Cybercriminals can misuse patients' electronic health records to create fake IDs to buy drugs and medical equipment which they can sell later.

Hackers can also file a fraudulent Insurance claim in a patient's name.

Integration: Multiple Devices & Protocols

Integration of multiple devices also causes hindrance in the implementation of IoT in the healthcare sector. The reason for this hindrance is that device manufacturers haven't reached a consensus regarding communication protocols and standards.

So, even if a variety of devices are connected; the difference in their communication protocol complicates and hinders the process of data aggregation. This non-uniformity of the connected device's protocols slows down the whole process and reduces the scope of scalability of IoT in healthcare.

Data Overload & Accuracy

As discussed earlier, data aggregation is difficult due to the use of different communication protocols & standards. However, IoT devices still record a ton of data. The data collected by IoT devices are utilized to gain vital insights.

However, the amount of data is so tremendous that deriving insights from it are becoming extremely difficult for doctors which, ultimately affects the quality of decision-making. This will eventually lead to patient safety issues. (Andreev et al., 2022, pp.156-161)

Moreover, this concern is rising as more devices are connected which record more and more data.

Cost

Surprised to see cost considerations in the challenge sections? I know most of you would be; but the bottom line is: IoT has not made healthcare facilities affordable to the common man yet.

The boom in Healthcare costs is a worrying sign for everybody, especially the developed countries.

The situation is such that it gave rise to "Medical Tourism" in which patients with critical conditions access healthcare facilities of developing nations which costs them as less as one-tenth. IoT in healthcare as a concept is a fascinating and promising idea.

However, it hasn't solved the cost considerations as of now. To successfully implement IoT app development and to gain its total optimization the stakeholders must make it cost-effective otherwise it will always remain out of everyone's reach except the people from the high class.

Scalability

Billions of internet-enabled devices get connected in a huge network, large volumes of data are needed to be processed. The system that stores, analyses the data from these IoT devices need to be scalable. At present, in the era of IoT evolution, everyday objects are connected via the Internet. The raw data obtained from these devices need big data analytics and cloud storage for the interpretation of useful data. (Maroto, 2019, p.43)

Interoperability

Technological standards in most areas are still fragmented. These technologies need to be converged. Which would help us in establishing a common framework and standard for the IoT devices. As the standardization process is still lacking, interoperability of IoT with legacy devices should be considered critical. This lack of interoperability is preventing us to move towards the vision of truly connected everyday interoperable smart objects.

Lack of Government Support

Government and Regulatory bodies like FDA should come up and bring up regulations by setting up a standing committee for the safety and security of devices and people. (Andreev et al., 2022, p.191)

Safety of Patients

Most Of IoT devices are left unattended, as they are connected to real-world objects. If used on patients as wearable devices, any technical error in security can be life-threatening for the patient.

Design-Based Challenge

With the development in technology design challenges are increasing at a faster rate. There have been issues regarding design like limited computation power, limited energy, and limited memory which need to be sorted out.

IoT DEVICES

Internet of Things Devices is non-standard devices that connect wirelessly to a network with each other and can transfer data. IoT devices are enlarging the internet connectivity beyond standard devices such as smartphones, laptops, tablets, and desktops. Embedding these devices with technology enables us to communicate and interact over the networks and they can be remotely monitored and controlled.

There are large varieties of IoT devices available based on the IEEE 802.15.4 standard. These devices range from wireless motes, attachable sensor-boards to interface-board which are useful for researchers and developers.

IoT devices include computer devices, software, wireless sensors, and actuators. These IoT devices are connected over the internet and enable the data transfer among objects or people automatically without human intervention. (Andreev et al., 2022, pp.208-212)

Properties of IoT Devices

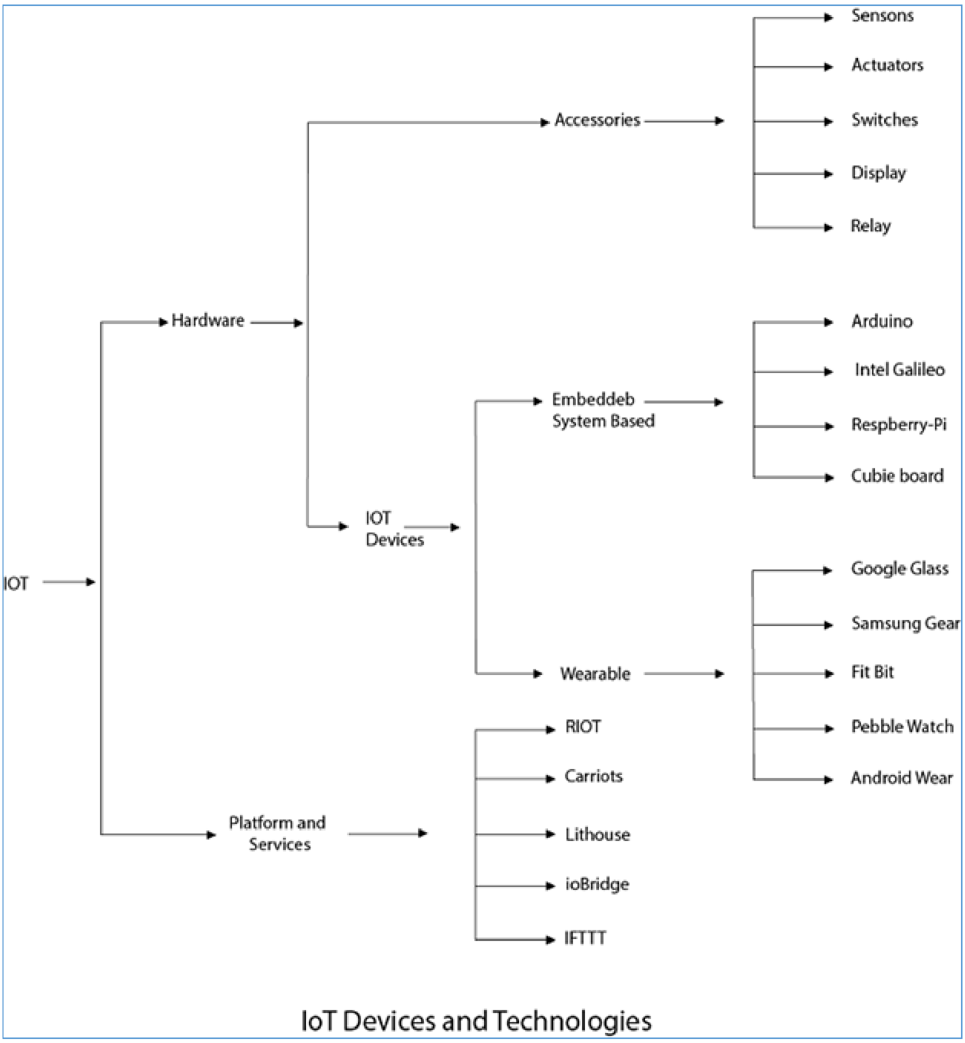
Some of the essential properties of IoT devices are mentioned below:

- Sense: The devices that sense their surrounding environment in the form of temperature, movement, the appearance of things, etc.
- Send and receive data: IoT devices can send and receive data over the network connection.
- Analyze: The devices can able to analyze the data received from the other device over the internet networks.

- Controlled: IoT devices may control from some endpoint also. Otherwise, the IoT devices are themselves communicate with each other endlessly leads to system failure.

Some of the common and popular IoT devices are given in Figure 2.

Figure 2.



IoT PLATFORM

As in IoT, all the IoT devices are connected to other IoT devices and applications to transmit and receive information using protocols. There is a gap between the IoT device and IoT application. An IoT Platform fills the gap between the devices (sensors) and applications (network). Thus we can say that an IoT platform is an integrated service that fulfills the gap between the IoT device and application and offers you to bring physical objects online. (Ghosh et al., 2016, p. 172)

Figure 3.

There are several IoT Platforms available that provides the facility to deploy IoT application actively. Some of them are listed below:

Amazon Web Services (AWS) IoT platform: Amazon Web Service IoT platform offers a set of services that connect to several devices and maintain security as well. This platform collects data from connected devices and performs real-time actions.

Microsoft Azure IoT platform: Microsoft Azure IoT platform offers a strong security mechanism, scalability, and easy integration with systems. It uses standard protocols that support bi-directional communication between connected devices and platforms. Azure IoT platform has Azure Stream Analytics that processes a large amount of information in real-time generated by sensors. Some common features provided by this platform are

- Information monitoring
- A rules engine
- Device shadowing
- Identity registry

Google Cloud Platform IoT: Google Cloud Platform is a global cloud platform that provides a solution for IoT devices and applications. It handles a large amount of data using Cloud IoT Core by connecting various devices. It allows to apply BigQuery analysis or to apply Machine learning to this data. Some of the features provided by the Google Cloud IoT Platform are:

- Cloud IoT Core
- Speed up IoT devices
- Cloud publisher-subscriber
- Cloud Machine Learning Engine

IBM Watson IoT platform: The IBM Watson IoT platform enables the developer to deploy the application and build IoT solutions quickly. This platform provides the following services:

- Real-time data exchange
- Device management
- Secure Communication
- Data sensor and weather data services

Artik Cloud IoT platform: Arthik cloud IoT platform is developed by Samsung to enable devices to connect to cloud services. It has a set of services that continuously connect devices to the cloud and start gathering data. It stores the incoming data from connected devices and combines this information. This platform contains a set of connectors that connect to third-party services

How IoT Platforms Help

- IoT Platform connects sensors and devices.
- IoT platform handles different software communication protocol and hardware.
- IoT platform provides security and authentication for sensors and users.
- It collects, visualizes, and analyzes the data gathered by the sensor and device.

Why It Is Important In Healthcare?

We can consider an IoT healthcare facility as a collection of ubiquitous computing that mainly deals with external activities. In healthcare, IoT-based healthcare systems collect a variety of patient data and get inputs from doctors medical professionals. Continuous glucose monitoring for insulin pens is the best example of this.

All these devices can communicate with each other and take important actions that would provide timely help to save someone's life. After collecting the data, the IoT healthcare devices would send this critical information to the cloud so that doctors can act upon it.

From this, we can say that the potential application of IoT in healthcare can improve not only a patient's health but also the healthcare employee productivity and hospital workflows. (Ghosh et al., 2016, pp.204-206)

Applications of IoT in Healthcare

The rise of IoT is exciting for everybody due to its different scope of use in various sectors. In Healthcare it has several applications. Here are some remarkable IoT applications in healthcare:

- Reducing emergency room wait time
- Tracking patients, hospital staff, and inventory
- Enhancing drug management
- Ensuring availability of critical hardware

IoT has also introduced several wearables & devices which have made the lives of patients comfortable. These devices are as follows.

Hearables

Hearables are new-age hearing aids that have completely transformed the way people who suffered hearing loss interact with the world.

Nowadays, hearables are compatible with Bluetooth which syncs your smartphone with it.

It allows you to filter, equalize and add layered features to real-world sounds. Doppler Labs is the most suitable example of it.

Ingestible Sensors

Ingestible sensors are genuinely a modern-science marvel. These are pill-sized sensors that monitor the medication in our body and warn us if it detects any irregularities in our bodies.

These sensors can be a boon for a diabetic patient as would help in curbing symptoms and provide an early warning for critical health issues. Proteus Digital Health is one such example. (Zobaa & Bihl, 2018, p.203)

Computer Vision Technology

Computer vision technology along with AI has given rise to drone technology which aims to mimic visual perception and hence decision making based on it.

Drones like Skydio use computer vision technology to detect obstacles and navigate around them.

This technology can also be used for visually impaired people to navigate efficiently.

Healthcare Charting

IoT devices such as Audemix reduce much manual work which a doctor has to do during patient charting. It is powered by voice commands and captures the patient's data.

It makes the patient's data readily accessible for review. It saves doctors' work by 15 hours per week.

Insulin Pens and Smart CGM

These devices are used for the real-time monitoring of blood glucose levels and data sharing over a dedicated mobile app. Patients with diabetes can use these devices to track their glucose levels and even send this data to their doctor and the relevant medical staff.

Ingestible Sensors

Medical sensors help the patients to swallow the prescribed medication via a tiny digestible medical sensor that sends a small signal to a wearable receiver on the patient, which sends data to a dedicated smartphone app.

Smart Video Pills

A smart pill travels through a patient's intestinal tract to take its clear-cut picture. It can then send that pictures to wearable devices that is connected with dedicated medical applications. Smart pills are also helpful to visualize the gastrointestinal tract and colon remotely.

CONCLUSION

This chapter proposed to include:

- the working strategies of IoT devices in health care by including the module integration.
- The Intelligent data analytic techniques and strategies using the characteristics of Big data computing in the health care sector.
- The Knowledge extraction process which integrated with the process of IoT devices and Big data analytics to support timely decisions.
- The basic methods, strategies, techniques, and module integrating architectures are represented. Hope this effort will consume the IoT devices with energy harvesting approaches that lead to efficient data collection and maintenance processes.

IoT changes the way facilities are delivered to the healthcare industry. These technologies improve the product, causing a larger effect by bringing together minor changes.

REFERENCES

Adhikari, A., & Adhikari, J. (2015). Advances in Knowledge Discovery in Databases. Springer Publications.

Andreev, S., Balandin, S., & Koucheryavy, Y. (2022). Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer International Publishing.

Ghosh, Shaw, Islam, & Piuri. (2022). *AI and IoT for Smart City Applications.* Springer.

Kumar, R., Srivastava, R., & Balas, V. E. (2019). Recent Trends and Advances in Artificial Intelligence and Internet of Things. Springer International Publishing.

Maimon, O., & Rokach, L. (2005). Data Mining and Knowledge Discovery Handbook. *Springer*.

Maroto. (2019). *Industrial IoT - Edge Computing Vendors Overview.* Independently Published.

Pankajavalli, P. B., & Karthick, G. S. (2019). Incorporating the Internet of Things in Healthcare Applications and Wearable Devices. IGI Global.

Piateski, G., & Frawley, W. (1991). Knowledge Discovery in Databases. AAAI Press.

Zhang. (2004). *Knowledge Discovery In Multiple Databases.* Springer.

Zobaa, A. F., & Bihl, T. J. (2018). Big Data Analytics in Future Power Systems. CRC Press.

Chapter 7
A Novel Dual Image-Based Reversible Hiding Technique Using LSB Matching-Digital World

Kalyanapu Srinivas
Kakatiya Institute of Technology and Science, India

Harshini B.
Kakatiya Institute of Technology and Science, India

K. Mounika
Kakatiya Institute of Technology and Science, India

B. Sai Sreeja
Kakatiya Institute of Technology and Science, India

Vyshnavi Kandukuri
Kakatiya Institute of Technology and Science, India

Abhinay K.
Kakatiya Institute of Technology and Science, India

ABSTRACT

The term internet has become more popular these days, and the whole world is connected virtually. Most people communicate through the internet. Communicating information from one to another in network without disclosing to third party is a typical task, but tracking of information is performed when certain security measures are not taken during transmission between the real-time environment and virtual environment. Creation of virtual representation for real-time process is known for the digital world. Secret data communication is performed under the digital world concept where there is a need of security. The chapter introduces a mechanism providing security using the ISDHR technique with least significant matching and dual images for secret message communication. The results show the proposed method is high enough in providing security to the secret data under transmission.

DOI: 10.4018/978-1-6684-5722-1.ch007

I. INTRODUCTION

As everyone nowadays utilises the internet, protecting data has become critical. Securing the information means protecting information from unauthorized access or disruption. Staying free of threat or feeling secure are both definitions of security. Business or certain organisations take precautions to detect and prevent illegal access from outside adversaries. These days the necessity of security over network for communication has expanded as well as the value of communication security. Since a huge number of images are circulated over the internet, if any data or information is stored in it, then there is a chance of threat to this data. The image that has information embedded in it can be safe when sent from sender to receiver if it is securely encrypted. There are several methods to hide information or privately transmitting information. The technique of hiding single message inside either another message or physical item. Any computer file, text, picture, or video is hidden inside any other file, text, picture, or video in computing/electronic settings is called as Steganography. Because our images are shared and accessed via public networks, using cryptography or steganography to hide information in images and perform data security is very important. Steganography is the greatest aspect of transferring information this regard. Everything that is required for steganography would be a cover text, which is the means in which information would be hidden, a message composed of information, an algorithm that determines how to hide and retrieve Information, and, optionally, a key that will be used to randomise the set - up of both the information and, possibly, encrypt it. Because of wide number of reasons, data security is vital for both public and private sector companies. For example, businesses have a lawful duty to secure its client and customer corruption of data. Data security functions are necessary to avoid information breaches and decrease the fear of data exposure. There are a variety of approaches for hiding Information within ordinary files.

Embedded pictures are the most widely mentioned kind of steganography. It's also the type which has received the most attention and there are various other types of algorithms like LSB, DCT etc., Software can encrypt an image in the same manner that it can encrypt words. Software used for encryption modifies the values of the numbers in an uniform manner by conducting a series of mathematical operations, known as an algorithm, on the binary data that makes up an image. The internet is frequently utilised and traded for images or visual elements like video and audio files. These types of mediums are ideal for concealing messages. Secret message bits are put in a section of the cover file or carrier which won't be recognised by other party. Data hiding is concerned with hiding the presence of a secret communication, whereas cryptography is concerned with safeguarding the substance of secret data. This is classified like a data concealing and reversible data hiding method since the actual product, together with the hidden data, may be recovered again from stego object. Employee records, client information, and transactions are all examples of sensitive data that must be safeguarded. This is to prevent other parties from misusing the information for fraudulent purposes, such as phishing schemes and identity theft.

The smallest (right-most) bit in a binary code is known as the Least Significant Bit (LSB). Modifying the LSB is one of the simplest ways to encode a secret message in picture pixels but also is referred to LSB-based steganography. Every LSB steganography technique will either modify the pixel value or keep it the same. This LSB method works best for media files, in which even a single bit out of place entirely changes the character, but not good for ASCII text, where even a single bit are beyond position totally changes the character.

II. LITERATURE REVIEW

In 2017, Lu et proposed a Dual-image data hiding technique using folding process. The approach is improved in two steps i.e., by repeatedly adding the folding process and secondly identifying the second fold operation using an indicator. The problem of overflow and underflow are efficiently handled by this solution. Testing is implemented to check performance of the proposal using eight methods with various values for B and K values. The hidden messages can only be decoded by someone who has both stego-images.

In 2009, a data hiding technique with Edge detection using fuzzy and canny was proposed by Wen-Jan Chen. Steganography is the process of hiding data. The most common technique of steganography is hiding data in images. The data-hidden images are called stego images, these stego images will produce distortion. To improve stego images in this method we are using hybrid edge detectors. The secret data is hidden using classic LSB steganography then by using fuzzy and canny edge detector l1, l12 edge images are obtained. Hybrid edge image l is obtained by performing OR operation for l1 and l2. By using different algorithms secret data is embedded and extracted from Hybrid edge images.

Shui Chih-Wei proposed a hiding technique using pixel value ordering method for embedding the information, in 2019. The method that followed in this paper is to increase security and provide strong privacy. Just like most of the other methods even in this method data is hidden in images. The original image is divided into blocks of 2 bits, embedding is done using pixel value ordering and prediction error expansion technique. This phase is called image partition. The next phase is image encryption, after encryption the next step is embedding a message. Embedding of secret messages is performed using histogram fin shifting. The final step is image decryption, message extraction and image recovery. To increase content-owner privacy in this method different types of keys are used separately for image decryption and message extraction.

A proposal by Frng jiang is a technique of data hiding using public key concept. This paper supports public key encryption schemes using selective multimedia encryption methodologies. Generally, most of the applications of selected encryption area have used symmetric crypto or a hybrid technique. To test security during encryption, multiple methods like quality test in images, correlative test and security analysis are used. A new model was proposed i.e public-key selective multimedia encryption model based on elliptic curve asymmetry encryption and scalable compression technique. This is successfully implemented by a secured public-key selective multimedia encryption.

A new method was proposed by Shobana et al, 2016, based on image color for hiding messages. Initially, the cover image was split into different colours other than red, green, blue colors (RGB) for security purposes. All colour images are based on red, green and blue layers but this method uses CMYK - stands for cyan, magenta, yellow, black(key) colours. In carrier image, each pixel value is applied with modulus of 4. Now, each pixel of the message image is covered by cover image pixel. Then embedding process is performed where all the pixels in the carrier image are involved. Finally, grayscale image is a secret message that is equal to the size of a cover image. The stego image is used to extract the message image by splitting it into cyan, magenta, yellow, and black layers (C, M, Y, K respectively). Later, the array is split into sub array of size 4 digits. Now, all 4 digits values are converted into its equivalent decimal value. Finally, to get the secret image, all these decimals values are arranged in sequential manner.

A double image-based hiding and reversing technique was proposed by Lu etal. in 2015. A pair of pixels are selected and embedded in the selected images which are later used for extraction process. An improved scheme was presented in 2017 by Wang et al. to increase the capacity.

The ownership of an image can be proven by hiding the secret data in an image using copyright of an image which has become very important (Tian, 2003). In few cases, novel techniques are needed like reversible data hiding mechanisms - binary images (Wu & Hwang, 2017). The pixel value differencing method (Lu et al., 2006; Manoharan, 2016) and SMVQ (Chiou et al., 2011; Chiou et al., 2013) methods are many other data hiding schemes.

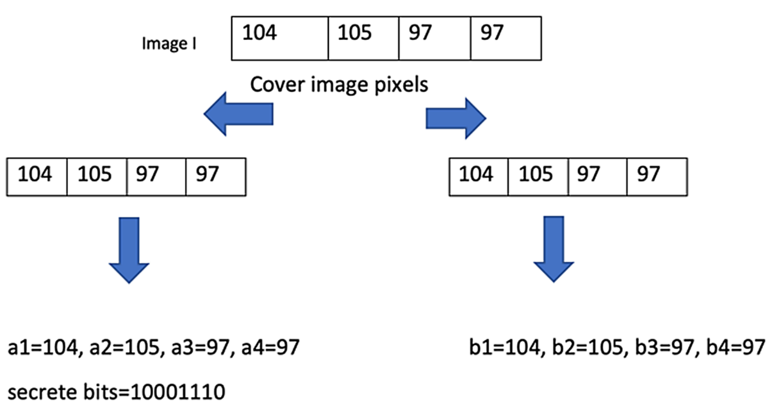
III. PROPOSED WORK AND METHODOLOGY

Implementation

To provide more security to the message under communication from physical to virtual environment, a model is proposed named Image-based Secret data Hiding – Reversing Technique (ISDHR). Using dual images for this proposal adds on more security. ISDHR is a proposed technique used for data hiding and reversing. By using ISDHR both secret information and the original image can be retrieved. The technique uses matching of lease significant bits of secret data in order to embed in the image. By dividing the original image into two copies and substituting secret information in the LSB positions of each pixel value of the images.

The technique invokes a cover image of size m x n, which is further considered as two copies. The pixel values of each image are converted into equivalent binary form. Secret data in binary form which is to be embedded in the images are mapped with the least significant bits of each image. In image copy1, the least significant bit positions 1 and 2 of image are mapped and substituted with secret data bits. In image copy2, the least significant bit position 3 of image is mapped and substituted with secret data bit. The mapping procedure followed shows the strongness of the hiding process. Then to retrieve the secret information take LSB3 in copy image1 and LSB1 AND LSB2 from copy image2. To retrieve original image, take LSB1 and LSB2 from copy image2 and LSB3 from copy image1.

Figure 1. Cover image to Image copy1 and Image copy2



Embedding, extraction, and pixel restoration algorithm

Step 1: Let us consider an image I (I1, I2, I3, I4) and this cover image is again divided into two copies C1(a1, a2, a3, a4) and C2(b1, b2, b3, b4).

I-Cover image

I (I1, I2, I3, I4): First, second, third and fourth pixels of cover image respectively. Similarly,

C1(a1, a2, a3, a4): First, second, third and fourth pixels of copy1 image respectively.

C2(b1, b2, b3, b4): First, second, third and fourth pixels of copy2 image respectively.

Step 2: Using Cover image first pixel value I1 and secrete bits m1 and m2 modify (a1, b1) using equation (1).

$$(a1, b1) = \begin{cases} (I1, I1), \ if\left(LSB(I1) = m1\right) \ and \left(LSB\left(\frac{I1}{2} + I1\right) = m2\right) \\ (I1, I1+1), \ if\left(LSB(I1) = m1\right) \ and \left(LSB\left(\frac{I1}{2} + I1\right) \neq m2\right) = m2) \\ (I1-1, I1), \ if\left(LSB(I1) \neq m1\right) \ and \left(LSB\left(\frac{I1-1}{2} + I1\right) = m2\right) \\ (I1+!, I1), \ if\left(LSB(I1) \neq m1\right) \ and \left(LSB\left(\frac{(I1-1)}{2} + I1\right) \neq m2\right) \end{cases} \quad (1)$$

Step 3: Embedding process: After modifying a1 and b1 pixels of image copy1, new pixel values are obtained. These values are new pixel values of stego image (a1*, b1*). Similarly with image copy2.

$$(a1^*, b1^*) = \begin{cases} (a1+2, b1-1), \ if \ a1 < I1 \ and \ b1 = I1 \\ (a1, b1), \qquad otherwise \end{cases} \quad (2)$$

Step 4: Extraction process: a secret bit m1 can be obtained by LSB(a1) and m2 bit can be retrieved using below equation.

m2=LSB((a1*/2) +b1*) (3)

Step 5: Apply the below equation to obtain the cover image pixel I1*

I1*=(a1*+b1*)/2 (4)

Working Example

An image of size 20 x20 is considered and divided into two copies i.e Image copy1 and Image copy2. Let a1,a2,a3,a4 be the pixels of image copy1 and b1,b2,b3,b4 be the pixels of image copy2 as shown.

a1=104, a2=105, a3=97, a4=97
b1=104, b2=105, b3=97, b4=97

Secret data (m1, m2, m3, m4, m5, m6, m7, m8) = 10001110

Embedding Process

(a1, a2) = (b1, b2) = (104, 105)

Secret data (m1, m2, m3, m4) = 1000.

For Image Copy1

LSB(a1) ^{1}m1 i.e., LSB (104) ^{1}1

So, a1=(a1-1) i.e., a1= (104-1)

Hence, a1=103

and LSB((a1-1)/2+a2) =m2, i.e., LSB (103/2+105) =0

Hence, a2=105(no change).

Similarly,

For Image Copy2

LSB(b1) =m3 i.e., LSB (104) =0

So b1=104(no change)

and LSB(b1/2+b2) ^{1}m4, i.e., LSB (104/2+105) ^{1}0

Hence, b2=b2+1=106

a1=103, a2=105, b1=104, b2=106

Now, consider the other pixels

(a3, a4) = (b3, b4) = (97,97)

Secret data (m5, m6, m7, m8) = 1110.

For Image copy1:

LSB(a3) =m5 i.e., LSB (97) =1

Hence, a3=97

and LSB(a3)/2+a4) =m6, i.e., LSB (97/2+97) =1

Hence, a4=97

Similarly,

For Image copy2:

LSB(b3) = m7 i.e., LSB (97) =1

So, b3=97 (no change).

and LSB(b3/2+b4) ^{1}m8, i.e., LSB (97/2+97) ^{1}0

Hence, b4=(b4+1) =98

a3=97, a4=97, b3=97, b4=98

secrete bits extraction:

m1=LSB(a1) = 1

m2=LSB(a1/2) +105 = 0

m3=LSB (106) = 0

m4=LSB (106/2) +105 = 0

m5=LSB (97) = 1

m6=LSB (97/2) +97 = 1

m7=LSB (97) = 1

m8=LSB (97/2) +98 = 0

Secret bits = 10001110

Pixel restoration of cover image:

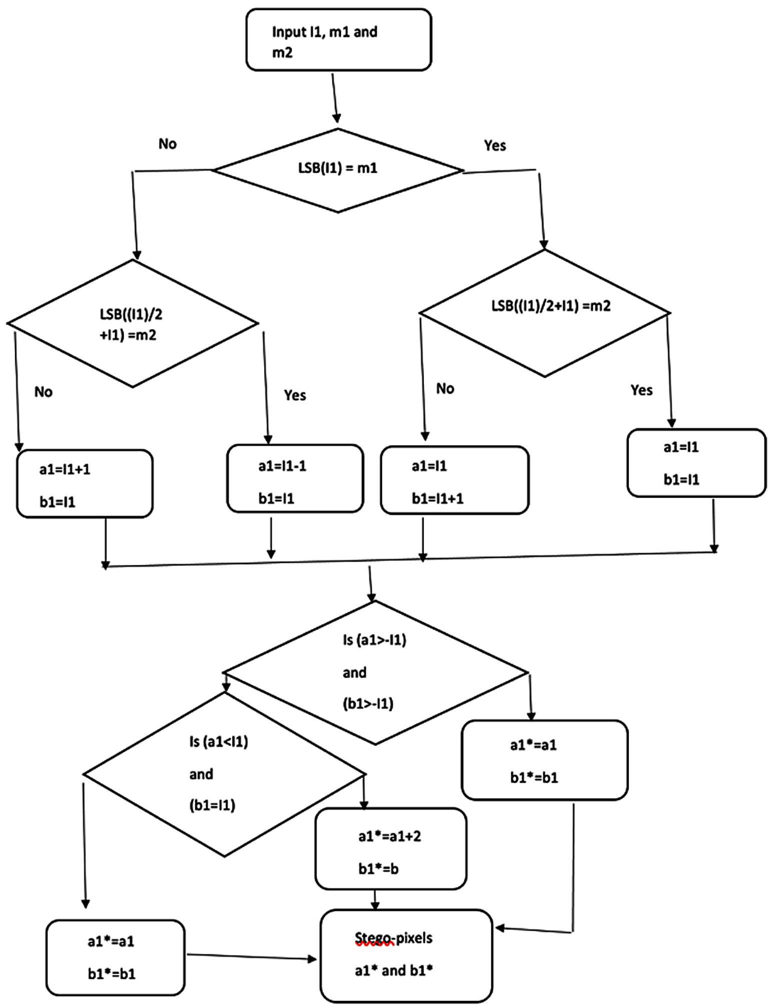
I1= (103+106)/2 =104

I2= (105+105)/2 =105

I3= (97+97)/2 =97

I4= (97+97)/2 =97

Figure 2. Flow diagram of the embedding procedure for the proposed method:



Extraction and Pixel Restoration Process

I1 =104 (Cover image pixel)

Assume, m1m2=11, LSB(I1) ^{1} m1 and

LSB((I1-1)/2) +I1 = m2. Hence, a1=I1-1 and b1= I1=104

As a1 < I1 and b1 = I1

a1* = a1+2 = 105 and b1* =b1-1 =103

I1* = (105 + 103)/2 =104 (pixel restoration)

m1 = LSB(a1*) = LSB (105) =1 and

m2=LSB (a1*/2 +b1*) =LSB ((105/2) +103) =1

(Secrete bit extraction).

IV. CONCLUSION

Image-based Secret data Hiding – Reversing Technique (ISDHR) provide security to the message under communication from physical to virtual environment. Usage of double images for this proposal adds on for performance improvement. Using ISDHR both secret information and the original image can be retrieved. Nevertheless, the experiment also suggests that our approach have an appropriate consequences and better ability. The ability of this method may be strong.

REFERENCES

Chen, W.-J., Chang, C.-C., & Le, T. H. N. (2010). High payload steganography mechanism using hybrid edge detector. *Expert Systems with Applications*, *37*(4), 3292–3301. doi:10.1016/j.eswa.2009.09.050

Chiou, S. F., Liao, I.-E., & Hwang, M.-S. (2011, February). A capacity-enhanced reversible data hiding scheme based on SMVQ. *Imaging Science Journal*, *59*(1), 17–24. doi:10.1179/136821910X12750339175943

Chiou, S. F., Lu, Y. C., Liao, I.-E., & Hwang, M.-S. (2013). An efficient reversible data hiding scheme based on SMVQ. *Imaging Science Journal*, *61*(6), 467–474. doi:10.1179/1743131X12Y.0000000035

Jiang, Salama, & King. (2017). A Public-Key Approach of Selective Encryption for Images. International Journal of Network Security, 19(1), 118-126.

Lu & Leng. (2017). Reversible Dual-Image-Based Hiding Scheme Using Block Folding Technique. *Information Technology And Its Applications*.

Lu, H. C., Chu, Y. P., & Hwang, M. S. (2006). A new steganographic method of the pixel-value differencing. *The Journal of Imaging Science and Technology*, *50*(5), 424–426. doi:10.2352/J.ImagingSci.Technol.(2006)50:5(424)

Lu, T. C., Tseng, C. Y., & Wu, J. H. (2015). Dual imaging-based reversible hiding technique using LSB matching. *Signal Processing*, *108*, 77–89. doi:10.1016/j.sigpro.2014.08.022

Manoharan & RajKumar. (2016). Pixel Value Differencing Method Based on CMYK Colour Model. *Int. J. of Electronics and Information Engineering, 5*(1), 37-46.

Manoharan, S. (2016). Pixel value differencing method based on CMYK colour model. *International Journal of Electronics and Information Engineering*, *5*(1), 37–46.

Shiu, Chen, & Hong. (2019). *Reversible Data Hiding in Permutation-based Encrypted Images with Strong Privacy.* Academic Press.

Tian, J. (2003). Reversible data embedding using a difference expansion. *IEEE Transactions on Circuits System Video Technology*, *13*(8), 890–896. doi:10.1109/TCSVT.2003.815962

Wang, Y. L., Shen, J. J., & Hwang, M. S. (2017). An improved dual image-based reversible hiding technique using LSB matching. *International Journal of Network Security*, *19*(5), 858–862.

Wu, N. I., & Hwang, M. S. (2017). Development of a data hiding scheme based on combination theory for lowering the visual noise in binary images. *Displays*, *49*, 116–123. doi:10.1016/j.displa.2017.07.009

Chapter 8
Using Deep Learning and Big Data Analytics for Managing Cyber-Attacks

Sarabjeet Kaur Kochhar
https://orcid.org/0000-0001-9406-7414
Indraprastha College for Women, University of Delhi, Delhi, India

Anishka Bhatia
Indraprastha College for Women, University of Delhi, Delhi, India

Nandini Tomer
Indraprastha College for Women, University of Delhi, Delhi, India

ABSTRACT

This chapter acquaints the reader to the terms and terminologies of cyber-attacks, cybersecurity, big data, data analytics, and related new age technologies, including deep learning. The types of cyber-attacks, how they become special and different within the big data analytic frameworks, a multi-layer framework for their detection, and the challenges therein are detailed next. Thereafter, an extensive review of some research works has been undertaken to provide an in-depth insight to the various cyber security detection systems using the new age technologies such as naive Bayesian networks in intrusion detection systems, deep learning in Android malware detection, and intelligent malware detection, etc. Conclusions have been drawn from these studies to establish that the emerging technologies, like artificial intelligence, machine learning, deep learning, and internet of Things, are the need of the hour to assist organizations in navigating the increasingly aggressive cyber threat landscape.

DOI: 10.4018/978-1-6684-5722-1.ch008

INTRODUCTION

Cyber-attacks are malicious efforts to steal, breach, alter, disable, or destroy web-based systems, through unauthorized access. According to an article published in Security Magazine a study was conducted by Michel Cukier, Clark Professor of Mechanical Engineering which stated that more than 2,200 cyber-attacks happen per day which equates to about one cyber-attack every 39 seconds (Clark Study School, n.d.). Cyber-attacks are rising with each passing day and therefore the severity of vandalism made by the cyber-attackers is increasing multi-fold. According to Trustwave's 2015 Global Security Report, approximately 98% of tested web applications were found vulnerable to cyber-attack (Trustwave, n.d.). Militia, Science and Research, top government agencies, businesses, healthcare, and even political groups are only some of the top targets for ransom or hacking secured information. Based on the Department of Business, Innovation and Skills' 2015 security survey 90% of the sizable organisations and 74% of the small organisations are affected by security breaches (PWC, n.d.).

Cybersecurity is at a tipping point, with the vast number of cyber-attacks, breaches, and threats increasing the need to respond quickly and precisely, before it's too late. The threat landscape is always evolving; for example, the rapid expansion of malware, ransomware (Richardson & North, 2017), DDoS (Garber, 2000), and social engineering (*Ns*, n.d.) assaults has already posed numerous issues to businesses. As an instance, a standard defence was good enough to protect any organisation from intrusions just a few years ago. Typical malware was easy to identify and targeted thousands of victims. Security solutions focused on blacklisting known malware signatures and were able to guard against the majority of attacks. However, the cybersecurity landscape — as well as modern attackers — have substantially evolved. They're clever and well-organized (many cybercrime operations are conducted like businesses), and they target specific individuals and businesses in search of lucrative targets. These hackers are quiet and sneaky, yet the damage they cause can be quite costly. Therefore, modern solutions and intelligent systems are required to deal with these cybercrimes.

In this challenging scenario, it has become need of the hour to apply Big Data Analytics and technologies like Artificial Intelligence, Machine Learning, Deep Learning and Internet of Things etc., to assist the organizations in navigating the increasingly aggressive cyber threat landscape. Big Data Analytics typically uses the Big Data to inspect, observe, and spot irregularities in the networks by analysing large amounts of data. The security-related information retrieved from Big Data has been successfully employed to cut down on the time it takes to identify and resolve a security problem. According to MeriTalk's recent U.S. government poll, 81 percent of Feds said their agency uses Big Data Analytics for cybersecurity in some form – 53 percent as part of their overall cybersecurity strategy, and 28 percent in a restricted capacity (MeriTalk, n.d.). Artificial Intelligence (AI), Machine Learning and Deep Learning have now become the primary technologies employed worldwide for cybersecurity solutions. With a significant number of data breaches, ransomware assaults, and sophisticated state-sponsored cyber-attacks, 2020 was a year marked by unparalleled cyber-crime. A tremendous increase was seen in the number of attacks as well as the monetary damage caused due to the attacks. Throughout the first half of the year, hackers used staff login credentials at one of Marriott International's franchise locations to get access to 5.2 million records containing personal information about visitors (Gupta, 2020). Following that, notable public figures' Twitter accounts were hacked, including Jeff Bezos, Bill Gates, Elon Musk, Barack Obama, and Joe Biden (Twitter Investigation Report, n.d.). Further, attacks were witnessed by a majority of corporations like the New Zealand Stock Exchange (BBC News, 2020), EasyJet (EasyJet Data Breach, 2021), SolarWinds' Orion software (Jibilian, 2021), etc. According to the Cisco Cybersecurity Reports, 50 percent of major firms with more than 10,000 employees spend nearly $1 million on security each year. According to the report, 43 percent of people spend around $250,000 to $999,999, while 7 percent spend under $250,000. While businesses embark on a digital transformation, it became increasingly important for them to upskill their personnel in order to prepare them for cybersecurity challenges. These technologies have been applied successfully to assist businesses in developing real-time, dynamic, and secured frameworks.

Figure 1. Cybersecurity and Emerging Technologies

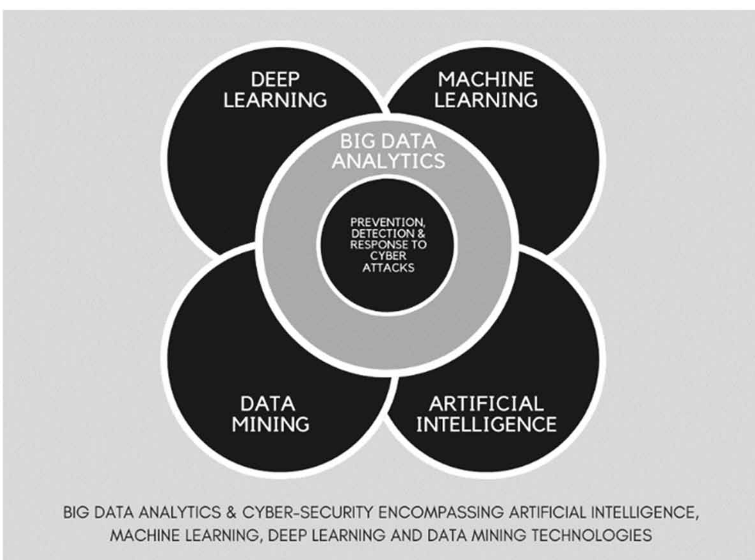
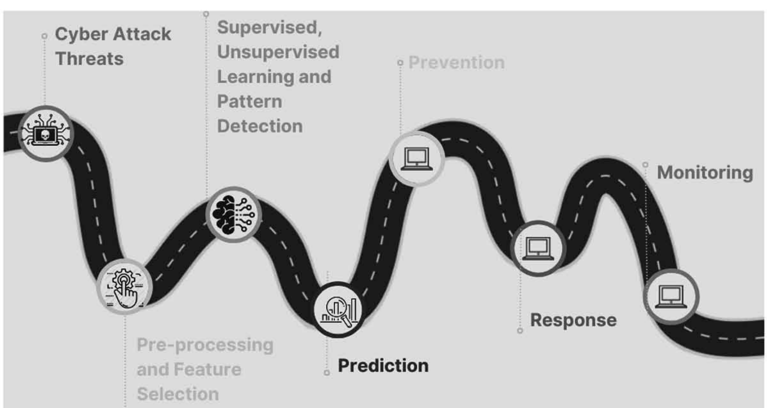


Figure 1 demonstrates how Deep Learning, Machine Learning, Data Mining and Artificial Intelligence technologies are interdependent and encompassed under the umbrella of Big Data Analytics and are being employed to prevent, detect, and respond to the cyber-attacks. By analysing news, articles, and research on cyber dangers, the data can be collected and analysed with very less human intervention. Systems have been designed using these technologies that modify access restrictions based on location or network and employ multi-factor authentication to thwart cyber-attacks. Cyber analysts also use Real-time Analytics and Predictive Analytics to predict and avoid network disruptions, as well as to assess network vulnerabilities and dangers. As a result, real-time and predictive analytics are fantastic ideas for avoiding cybercrimes (Mahmood & Afzal, 2013).

The technologies such as ML, Data Mining, Artificial Intelligence and Cybersecurity are intertwined with one another, as shown in figure 1, and are also been applied to the Big Data generated from IoT products and services, henceforth powering Big Data Analytics & Cybersecurity approaches. The Internet of Things in the layman's language means connecting things i.e., devices to internet services or to other devices via the Internet and this in a sense brings physical comfort in our lives by carrying out automated tasks remotely. Unfortunately, since the IoT-enabled devices primarily support direct object to object communications, they become more prone to cyber-attacks themselves (Shah & Sengupta, 2020). IoT devices can face threats such as DoS, injections, and IP misconfiguration at various layers viz Network Attack, Web-Interface attack, and Application Service layer (Kaur, 2021).

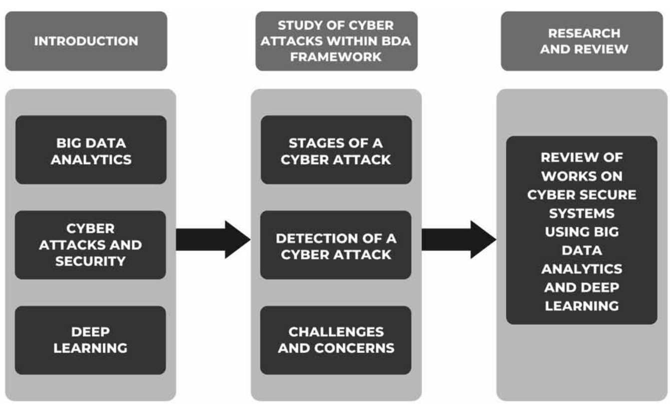
Figure 2. Dealing with Cyber Attacks using Big Data Analytics



As can be seen in Figure 2, analysing and monitoring various cyber-attack threats is a non-trivial process, involving various steps. The attacks are first identified individually, followed by gathering data about them. This is then followed by data pre-processing and feature selection, which reduces dimensionality by selecting only the features which are likely to provide significant insights. Machine Learning, Data Mining and Deep Learning algorithms are then applied for patterns and deviation detection etc., that assist in identification of attacks, in turn leading to brainstorming for a prevention method. Generally, various models are tested, compared on training time, testing time, accuracy, etc., finally choosing the best fit for our solution. The model is then monitored for some time to ensure that there is no breach of security.

The process flow of this chapter is depicted in Figure 3. We begin the chapter by introducing the reader to the basic terms and terminologies of the world of Big Data Analytics, Deep Learning, Cyber-attacks and Cybersecurity. Next, we explore cyber-attacks within the Big Data framework, how traditional methods have now been rendered powerless, given the increasing complexity of cyber-attacks. We discuss in detail, the probabilities, types and detection of cyber-attacks within the Big Data Analytic frameworks, highlight what is special and different therein, and what are the challenges of dealing with Big Data cyber-attacks. We also extensively review some research works, lying at the confluence of cybersecurity, Big Data Analytics, Machine Learning and Deep Learning before drawing conclusions and pointing out the technologies that promise to be part of the future of the cyber security landscape.

Figure 3. Process Flow Diagram



The World of Big Data Analytics, Deep Learning and Cyber Security

Data is everywhere; enormously and exponentially increasing with each passing second. According to research conducted by North-eastern University, by the year 2020 each person produces around 1.7 megabytes of data each second (Northeastern University, 2016).

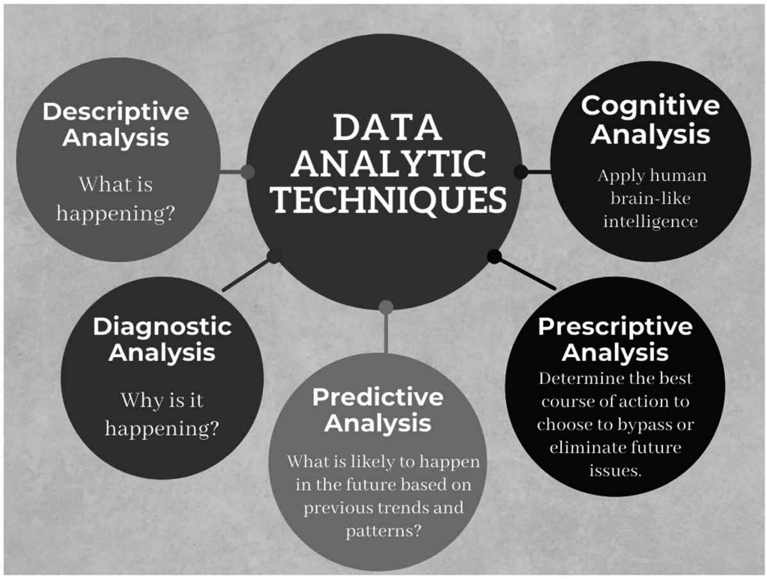
The term Big Data is often used to refer to the gigantic mountains of data, that are being continually fed by endless data sources such as the point of sales transactions around the world, meteorological data sensors, website click streams, biogenetic sequences, satellites scanning the skies and experiments being conducted by the scientists worldwide.

Laney proposed 3Vs to define the Big Data: volume, variety, and the velocity (Kitchin & McArdle, 2016). The first V denotes the sheer high volume of the data. Big Data can be voluminous owing to a very large number of dimensions or simply the high number of tuples. The variety of sources points towards the heterogeneity of data i.e., the sources from where the data is collected. The data may come from geographically different locations and/or in a variety of formats, such as completely structured data (the data stored in company databases or ERP packages), semi-structured data (XML, emails, or web pages) or unstructured data (image, audio, video, or document files etc). The term velocity refers to the speed at which the data is collected. It may be streamed in continuously or collected batchwise.

Big Data Analytics (BDA)

Big Data Analytics (BDA) is the computational branch of study that aims to find patterns, trends, and associations in Big Data. It primarily involves the application of the operators such as Gathering, Selection, Pre-processing, Transformation, Mining, Evaluation, and Interpretation for exploring raw data. Key technologies such as KDD (Data Mining), AI and Machine Learning have been applied to analyse the raw data for discovering the hidden gems of novel, interesting, actionable knowledge. Descriptive, Diagnostic, Predictive, Prescriptive, and Cognitive Analysis, depicted in Figure 4, are some of the prominent data analytic techniques that aid in this endeavour. Healthcare (Raghupathi & Raghupathi, 2014), sales, marketing, planning, and forecasting (Choi et al., 2018), fraud detection, natural language processing (Agerri et al., 2015), social media and market sentiment analysis are only some of the applications which have successfully deployed and benefitted from BDA.

Figure 4. Data Analytic Techniques



A large volume of data is processed at different phases of business analytics, requiring different types of analysis, as seen in Figure 4. Descriptive Analysis analyses real-time data using appropriate visualisation tools such as dashboards, allowing one to look at past behaviours and predict how they will affect future outcomes. Diagnostic Analysis enables analysts or data scientists to go deeper into the data in order to identify the causes or factors affecting the outcome. Predictive Analysis provides advice and answers to queries about what might happen in the future. It is probabilistic and it just provides an estimate of a possible future outcome. Prescriptive Analysis is used to come up with strategies that maximise outcome. Cognitive Analysis is influenced by the way the human brain interprets data, forms conclusions, and converts intuition and experience into knowledge.

A few of the popular software used in the field of Data Analytics are R programming, Tableau, Python programming, SAS, Apache Spark, and Excel etc. In addition, various statistical, Machine Learning and Artificial Intelligence algorithms have been used for analytics purposes for a long time. But with the emergence of age of Big Data, the traditional Data Analytics may not be able to handle such enormous quantities of data because of the following reasons:

- Data Analytics is compatible with structured data that meets a certain criterion but in reality, most data (i.e., Big Data) coming in constantly is unstructured and includes text, audio, video, and web content etc.

- While traditional Data Analytics is predicated on a centralised database architecture, Big Data uses a distributed architecture making it scalable with efficient performance and more economical.
- In Data Analytics, the users are typically expected to know their questions at the start in order to find the answers to the questions asked and then visualisation reports would be generated however Big Data enables a more iterative and exploratory approach.
- With the surge in data, Big Data analysts are encountering a multi-fold surge in data management platforms to support these Big Data initiatives. shown that they are employing a variety of platforms reinforcing Big Data. Hadoop and NoSQL platforms including Cassandra and MongoDB are the most prominent environments in the world of Big Data Analytics.

Figure 5. Traditional Cybersecurity vs Cybersecurity with Big Data Analysis

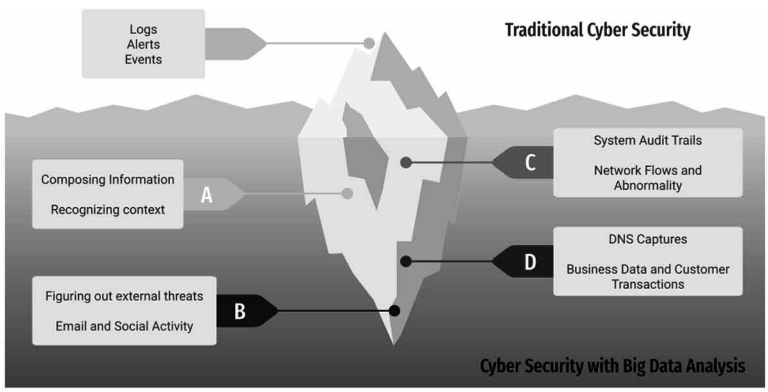


Figure 5 depicts that with traditional cybersecurity methods, one could only monitor a few things such as logs, alerts, and events. But with modern Big Data Analytic techniques there is a lot more that can be covered and monitored. This proves how with traditional methods, efficiency could be achieved only at the surface level but with more modern analytics, one can actually dive deep and see how much more can be analysed.

It can therefore be concluded that Big Data goes hand-in-hand with acclaimed technologies like Machine Learning, Data Science, AI, Deep Learning, etc. and will prove to be remarkable in improvising the furtherance in research as well.

Cyber-Attacks and Cyber Security

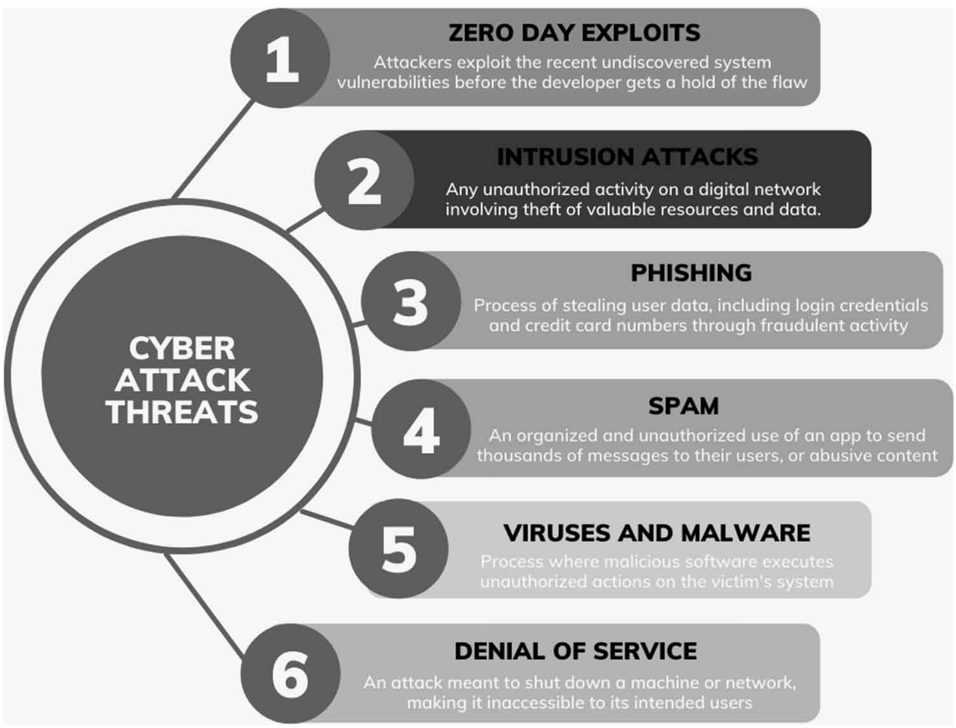
A cyber-attack is a malicious effort made by an individual, such as a cybercriminal or a hacker, or an organisation to steal, breach, alter, disable, or destroy information through unauthorized access to the information system of another individual or organisation. The surge in usage of the Internet has resulted in a paradigm shift in the backbone infrastructure of all the industries to the Cybernet, making these organizations increasingly vulnerable to cyber-attacks.

Cyber-attacks occur in many forms. They could be launched as Malwares, Denial-of-Service (DOS) Attacks, Phishing attacks, Zero-day exploits, IoT attacks, password attacks or Man-in-the-Middle (MitM) Attacks. These forms of attacks are depicted in Figure 6, and detailed as follows:

- **Malware**: stands for a malicious software and acts as an umbrella term encompassing numerous types of attacks including worms, *trojans, spyware and viruses*. **Viruses** are self-replicating applications with a malicious intent whereas **Worms** are like viruses but they self-replicate themselves across computer networks such as the internet. **Trojans** appear to be innocent malicious programs which have embedded malicious content inside a functional program and are difficult to be recognized whereas the spywares gather information for purposes of advertisement and stick to the infected machine for longer durations.
- **Phishing**: The practice of sending fraudulent communications, generally through an email, to steal sensitive data like credit card and login information or to install malware on target machines is called phishing.
- **Denial-of-Service (DOS):** DoS is a cyber-attack where the attacked system or the network becomes inaccessible to the users. The attackers usually aim for high-profile organisations such as banking, commerce, and government (Vemuri, 2005). The two most prevalent Denial-of-Service attacks are:
    - **Smurf attacks**: The Network- layer DDoS (Distributed Denial-of-Service) attack where an attacker sends broadcast packets to a number of hosts with a spoofed source Internet Protocol (IP) address.
    - **SYN flood**: The DoS/ DDoS attack where an attacker sends a connection request to target server but does not complete the connection through a three-way (TCP)/IP network handshake and the attacker keeps on sending requests, overwhelming the system so that valid users are unable to connect and leave the connection port in an occupied state leading to a SYN flood.

- **Zero-day exploits**: The cyber-attacks where the attackers exploit the recent undiscovered system vulnerabilities before the developer gets a hold of the flaw. Due to the nature of these vulnerabilities, these exploits are generally identified after the exploit is identified and patches are released which users must implement.
- **Intrusion Attacks**: Any unauthorized activity on a digital network involving theft of valuable resources and data. Intrusion is the malicious conduct of one or more external or internal nodes in a network, with the goal of causing harm to other nodes in the network (Vemuri, 2005).
- **Spam**: Spam attacks are described as the use of an app to deliver thousands of messages to its users in a coordinated and unauthorised manner. Bogus or hacked profiles send these messages, which frequently include fake adverts and links those actual users are prompted to click on.

Figure 6. Various kinds of cyber-attack threats

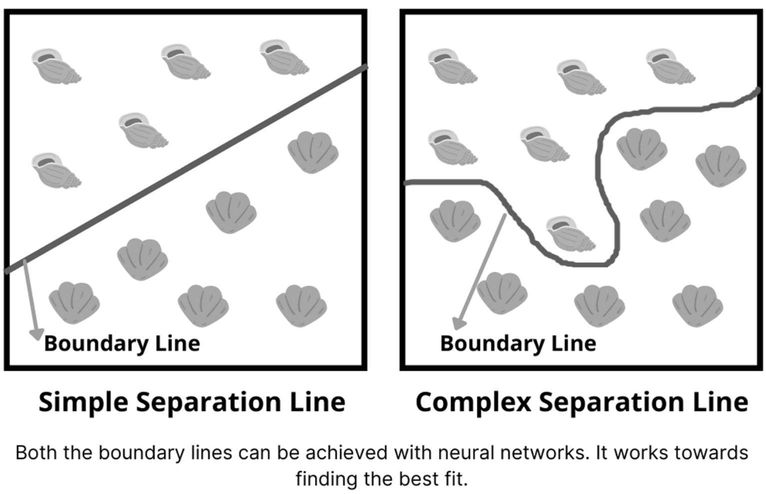


Deep Learning

A fan out of the technologies like Machine Learning and Artificial Intelligence, Deep Learning is a relatively new technology that extensively uses neural networks. A neural network abstractly mimics the operation of the human brain with the neurons that fire bits of information. It can be compared to a robot with an artificial human brain, involved with multiple layers of processing, to extract progressively higher-level features from data.

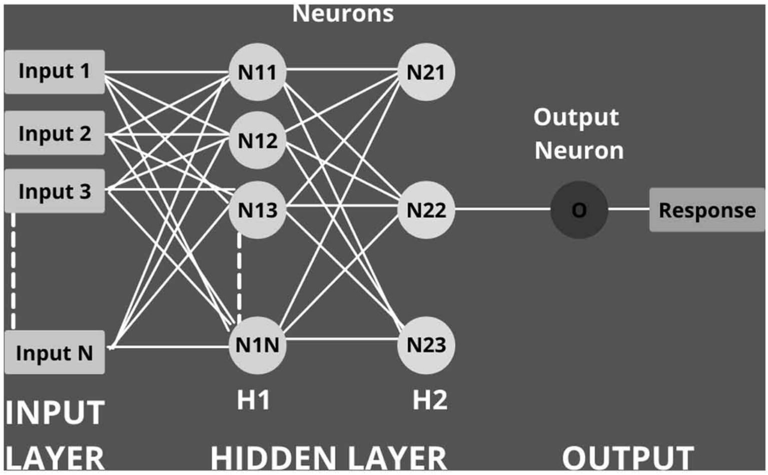
Let's take an example to understand how a neural network works. A child is playing with the red and blue shells, trying to separate the red shells from the blue shells by drawing a boundary line in the sand, as shown in Figure 7. Given some data, say, in the form of differently coloured points, a neural network will try to automatically search for the best line that separates them. For a more complicated data set, a more complicated algorithm, may be required to find a probably more complex boundary that separates the data. The number of layers generally goes up with an increase in the complexity of the problem at hand.

Figure 7. Neural Network example



A neural network comprises of nodes, edges, and layers. The layers in a neural network are defined to implement a certain function. The layers contain nodes, which are responsible for carrying information, and the edges represent the linkages and transfer of information between the different nodes. A representation of a neural network can be seen in Figure 8.

Figure 8. A representation of a neural network.



There are 3 kinds of layers in a simple neural network: input layer, hidden layer and the output layers. The first layer, or the input layer, as the name suggests, is usually responsible for accepting data to be passed on to the model. This input data information passes through the edges, to the hidden layers, implementing an underlying algorithm. The input layer contains n number of nodes, which individually contain n input features. The edges may be weighted, given by a vector W. Followed by the input layer are the hidden layers. The network depicted in Figure 8 contains two hidden layers, namely H1 and H2, which each contain their own respective neurons. The nodes in the hidden layer process the information with some coefficients and may assign new weights to the edges. The products of the input and weight vectors are then added. This process goes on until we reach the last algorithm, also known as the activation function, which determines the final output, also called prediction, at the output or target layer. The output layer may contain a single neuron or more neurons. Figure 8 represents a single neuron output layer, with the output neuron named O, which then gives out the response. The internal layers processing the information are generally represented by mathematical equations of the form:

$$Y = w1x1 + w2x2 + b$$

where, W = (w1, w2) represents the weights vector and X = (x1, x2) represents the input vector. b refers to the bias. So, we simply take a product of the 2 vectors, sum them up and try to predict the output label Y.

The benefits of Deep Learning for cybersecurity include real-time detection and prevention, real-time classification, prediction of unknown (future) threats, operation on any device or operating systems, and connectionless edge deployment. Some applications of employing a Deep Learning empowered cybersecurity systems are as follows:

1. Intrusion Detection and Prevention Systems (IDS/IPS)

These systems detect malicious network activities and prevent invaders from accessing the systems by recognizing known signatures and generic attack forms and then notifying the user. These are useful against threats like data breaches (Vemuri, 2005). Deep Learning neural networks create smarter ID/IP systems by analysing the traffic with better accuracy, reducing the number of false alerts, and helping security teams differentiate bad and good network activities (Handa et al., 2019). One of the shortcomings of the IDS based on Machine Learning is that they generate a lot of false positives, in turn causing problems for the security teams. Deep Learning enabled Next-Generation Firewall (NGFW), Web Application Firewall (WAF), and User Entity and Behaviour Analytics (UEBA) are some of the proposed solutions to this problem.

2. Dealing with Malware

Deep Learning algorithms detect advanced threats and don't rely on remembering known signatures or common attack patterns unlike the firewalls that use a signature-based detection system. Instead, they learn the system and recognize suspicious activities that might indicate the presence of bad actors or malware (Handa et al., 2019).

3. Network Traffic Analysis

Deep Learning artificial neural networks are used to analyse HTTPS network traffic and look for malicious activities, to safely deal with cyber threats such as SQL injections and DOS attacks (*A Deep Learning Based Method for Handling Imbalanced Problem in Network Traffic Classification | Proceedings of the Eighth International Symposium on Information and Communication Technology*, n.d.).

4. User Behaviour Analytics

User and Entity Behaviour Analytics (UEBA) is a great tool against attacks that happen due user activities in an organization. After a learning period, it picks up employee behavioural patterns and becomes more perceptive towards the suspicious activities, such as accessing the system during unusual hours, that could possibly indicate an insider attack, which is very important to track since it bypasses security measures ad raise alerts (Khaliq et al., 2020; R & Babu, 2021).

5. Fog-to-Things Computing

The new Internet of Things applications require new cybersecurity protocols, controls, and models to keep system development flaws, increased attacks, and hacking skills in check. The above-mentioned attacks are becoming inevitable (Atlam et al., 2018). Established machine-learning-based attack detection models have low accuracy and less scalability. Deep Learning models, using fog ecosystem, are better than shallow models at detecting false alarms, providing better accuracy and scalability, in turn increasing the level of sophistication in cybercrime detection (Abeshu & Chilamkurti, 2018).

Cyber-Attacks within the Big Data Analytic Framework

It has become challenging to prevent, detect and respond to the attacks using the traditional techniques, with the multi fold increase in the number of targeted cyber-attacks and advanced persistent threats (APTs). The APTs, are complex attacks, that steal highly sensitive data, occupy the host system for longer periods of time and fall into one of the following categories: Infiltration, Expansion or Extraction. Manual analysis used in the traditional techniques are inherently slow and warrant a lot of human effort. Moreover, as discussed in the previous section, the data is too large and heterogeneous (derived from multiple sources and in multiple formats), the traditional detection methods such as antivirus software and intrusion detection systems are rendered powerless. The manual analysis in traditional security techniques makes the detection process slower, increasing the chances of attackers causing harm to the target systems multi-fold.

Stages of Cyber-Attacks in Big Data Analytic Frameworks

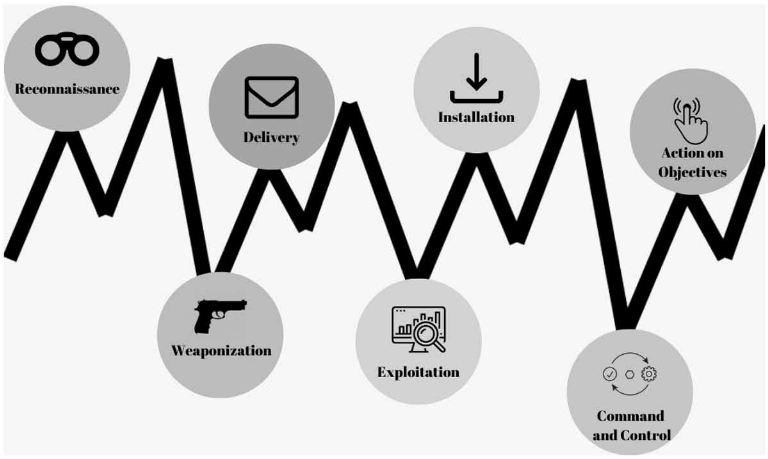
As new threats emerge and existing one's change, the never-ending battle between cyber security professionals and hackers continues. Critical infrastructure cyber-attacks are growing increasingly widespread, complicated, and creative. This is a constant issue for cyber security teams, who must identify where their operations are vulnerable to dangers before hackers do. Rather than attempting to steal data, attackers in recent high-profile cyber-attacks have sought to disrupt services. Thus, understanding the progression of a cyber-attack provides a solid foundation for detecting threats both before and after they occur.

Hackers and their approaches have some similarities. In essence, there are around seven phases to a successful cyber assault, often known as the "kill chain," which refers to the structure of a successful cyber-attack. The hackers may approach these several stages in a variety of ways, but they usually go through all the steps.

Understanding how hackers get access to and exploit critical infrastructure requires knowledge of the seven steps of a cyber-attack. Figure 9 depicts the 7 stages of a cyber-attack. Here is a summarization of these steps, and methods to combat them.

- **Stage 1: Reconnaissance:** Hackers start by conducting a web search of the company, collecting names, jobs, and contact information of employees, choosing one person to attack and then devise a strategy for doing it.
- **Stage 2: Weaponization:** Hackers have access to libraries of code that they use and alter for their assaults. They research about the protocols, Operating Systems, and software that the target organisation might use. The hackers can tailor their programmes to work in those settings by identifying these elements through study.
- **Stage 3: Delivery:** The assault begins during the delivery phase. Phishing e-mails are sent, and 'watering hole' web pages are put on the internet, while the attacker waits for all the information they require.
- **Stage 4: Exploitation:** The hacker begins to reap the benefits of their efforts in executing the assault, tests the details against e-mail systems or VPN connections to the network, got remote access If malware-infected attachments are shared, investigates the targeted network, obtains a better understanding of the traffic flow and connections, and how they'll be abused.
- **Stage 5: Installation:** By establishing a persistent backdoor, creating administrator accounts on the network, and disabling firewall restrictions, the attacker secures continuing access to the network. The hacker's goal at this phase is to ensure that they can stay in the system.
- **Stage 6: Command and Control:** Hackers have unrestricted access to the whole network, admin accounts, etc. for the command-and-control phase. The hacker may then operate it, listen to packets over the network, or crawl across the network. All of this is dependent on the software used and the hacker's goals for the system.
- **Stage 7: Action on Objectives:** The hacker can extract any data they're looking for. Of course, there are several ways to make use of this type of data. This might include stealing information on individuals, clients, product ideas, and other sensitive data. Alternatively, an attacker might begin disrupting the operations of the target firm.

Figure 9. The seven stages of a cyber-attack



Detection of Cyber-Attacks in Big Data Analytic Frameworks

The challenging scenarios, such as the ones documented above, require that the techniques used to detect cyber-attacks should be accurate, efficient, intelligent, that can correlate security data from different sources and prevent attacks as soon as detected. Big Data Analytics comes armed with the prospects of complex analysis of heterogeneous multi-sourced data and discovering trends and patterns in data with the help of Machine Learning (Fraley & Cannady, 2017), Artificial Intelligence, Data Mining and more recently Deep Learning techniques. So, detection and prevention of cyber-attacks using Big Data Analytics, has been much acclaimed for diagnosing/thwarting cyber-attacks. Deep Learning also ushers in unique benefits for dealing with cybersecurity. This includes real-time detection and prevention, real-time classification, prediction of unknown (future) threats, operation on any device or operating systems, and connectionless edge deployment. However, it is noteworthy that though Deep Learning and Big Data technologies have proven to be extremely advantageous for significant progress in a variety of fields, they themselves pose risk. Unlawful individuals can use these very technologies to launch cyber-attacks and pose a threat to enterprises. For instance, these techniques can be used to determine high-value targets.

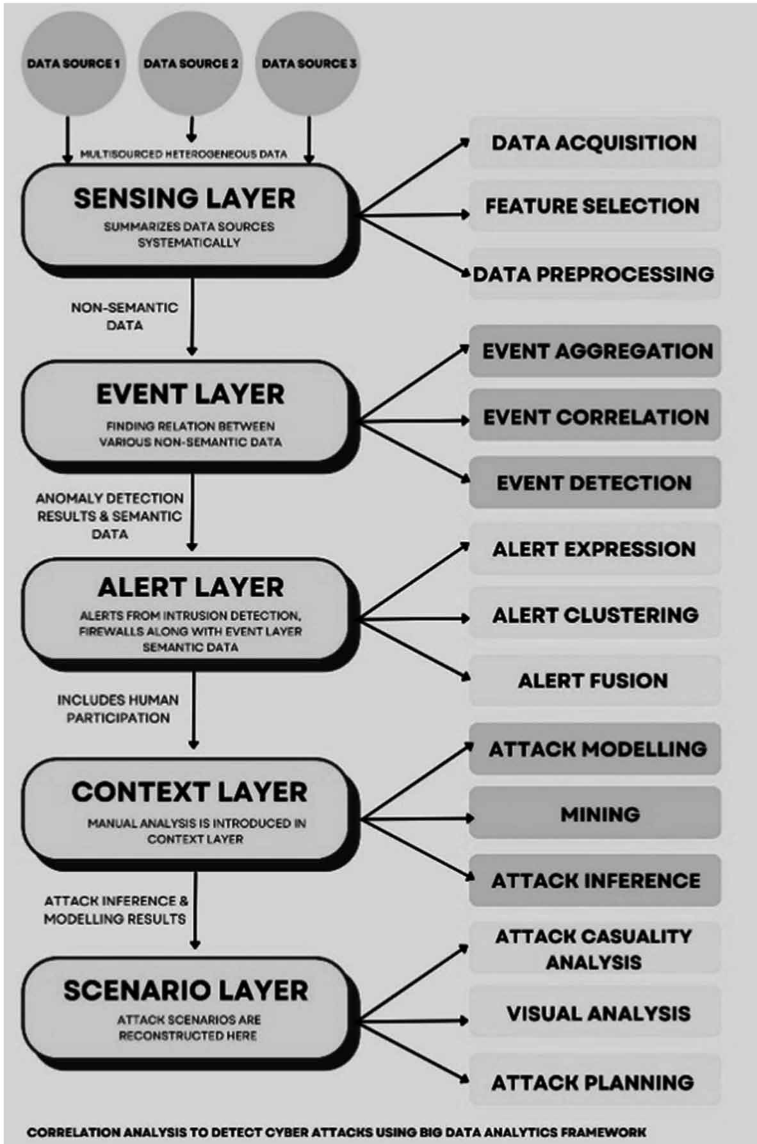
Detection of targeted cyber-attacks from heterogeneous multisource bid data needs to be considered as a systematic analysis process. *Correlation Analysis* is one of the techniques employed by the Big Data Analytics framework for cyber-attack detection, that takes up the attack detection at five different levels: *sensing layer, event layer, alert layer, context layer,* and *scenario layer* respectively, shown in

Figure 10. The lower-level layers serve as input sources whereas the upper layers are used analysis at higher level of granularities. Correlation analysis attempts to enable understanding of the underlying relations between the data from multiple sources by employing processing and visualization techniques.

- **Sensing Layer**: Data sensing is an important step that includes *data acquisition* from the heterogeneous, multiple sources, *data acquisition, feature selection and data pre -processing.*
- **Event Layer**: It deals with the collection of non-semantic data, such as the data which contains the information about running logs and processes, casting attacks where analysers pick the indifferent behaviour from data and generate semantic data. In this layer, the main tasks revolve around *event aggregation, event correlation and event detection, etc*.
- **Alert Layer**: The input point of the alert layer includes anomaly detection results and semantic data. The semantic data include alerts coming from event layer data after analysis and detection the alert layer contains all kinds of alert data with security semantics. In this layer, the research area includes *alert expression, alert clustering and alert fusion, etc.*
- **Context Layer**: As there's a wide range of alert data source we utilise from the alert layer, the consequence of alert correlation is just the preliminary attack steps. It is still a challenge to detect cyber-attack and extract attack context and attack fragments from security data of multi-sourced heterogeneous sources which also demands some sort of manual analysis. In this layer, the research area includes *attack modelling, mining, attack inference, etc.*
- **Scenario Layer**: In this layer, the attack scenario is reconstructed based on attack modelling and attack scenario reconstruction consistent with human cognition. In this layer, the research area includes *threat intelligence expression, attack causality analysis, visual analysis, and attack planning*.

Figure 10 depicts data from multiple sources entering the Sensing layer, where steps such as data acquisition, feature selection, and data pre-processing are performed. Non-semantic data is passed on to the next layer, the Event Layer, where event aggregation, event correlation, and event detection are performed. It is followed by anomaly detection results with semantic data that are absorbed by the Alert Layer, which includes alerts from the Event Layer data and contains all types of alert data. The Context Layer then performs manual analysis, which includes attack modelling, mining, and so on. Finally, the Scenario layer is responsible for reconstructing attacks in order to forecast and prevent them in the future.

Figure 10. Attack Detection Levels



Cybersecurity Challenges in Big Data Environments

For most organizations, Big Data and cybersecurity entails both an opportunity and a risk. As discussed earlier, the term "Big Data," refers to an extremely large data set that is mined and analysed for patterns and behavioural trends. It is characterised by a high density of variety, velocity, and volume. Big Data has created new possibilities with regards to analytics and security solutions for protecting data and preventing future cyber assaults from the perspective of cybersecurity. However, just as Big Data has provided new opportunities for cybersecurity teams, it has also provided cyber criminals with new ways to gain access to huge volumes of sensitive and personal data via advanced technology. So, the organizations have been poised with multiple challenges, while dealing with Big Data. Some of these are enumerated below:

- Big Data Analytics is considered to be a newer technology in the field of cybersecurity. Resultantly, the organisations are hesitant in adapting and applying the technology to shield against the cyber-attacks and breaches.
- With the advances in technology, hackers also hone their skills and can penetrate security using the same technologies. The data breaches are therefore growing more every year. Hackers get around the defences no matter how good the cybersecurity procedures are or how strong the data breach prevention is. They'll find and abuse a loophole that has been overlooked and left unchecked if they can't control the security. To put it another way, there's always a back door.
- Programs and regulations become outdated with time, and one must attempt to upgrade everything on a regular basis. As a result, one of the most pressing concerns today is the lack of oversight of cybersecurity measures. Penetration testing and vulnerability assessments are often overlooked by businesses.
- The organizations sometimes don't educate their personnel about new hazards. Investing in security is a huge overhead for most organisations, especially small ones, due to financial concerns. The consequences of a data breach, on the other hand, are significantly greater than the expenses of investing in good protection or being alert about it. If one truly wants to safeguard the sensitive data, they must keep improving and enhancing security procedures as a top priority every year.
- Inability to interpret data due to lack of talent (e.g., data scientists). Big Data can be mined to draw very useful insights, but organisations aren't capable of allocating a resource person who can make most out of the data and there are not a lot of trained employees who can actually dig in the gold mine of data. A big issue that newcomers face is that they have little or no experience in the field of data science because they are beginners, and firms seek experience because it is necessary for the job. This leads to a dead end.
- While extracting insights from the Big Data the organisations often fail to identify the right objective and end up with insights that are not so helpful for their growth.

- Cybercriminals have more incentive to steal data as online transactions increase. Furthermore, businesses that mine data—for example, customer information, product survey findings, and generic market data—create valuable intellectual property, which is a desirable target in itself.

As the data is coming in with high intensity and is unstructured, one cannot have complete faith in the quality and accuracy of the trends and graphs derived by analysing the data. So, it's always better to be on the safer side and compare parallelly with traditional detection algorithms.

There are enormous privacy concerns with the data that contains personal information of users, which can be misused financially or in other unethical ways and may lead to anonymity breaches. A lot of organisations in the sphere of Big Data do not indulge in the standard security and privacy protection tools. To make sure the user's identity remains intact the data should be kept strictly private regardless of the monetary benefits.

It is expected that with the knowledge of Big Data Analytics, one should be able to predict a crime before it actually takes place, so that the attacks can be prevented. This requires the help of visualisation tools to enable the security analysts to use the past data, patterns, and trends to draw preliminary conclusions about cyber threats and security breaches (de Boer et al., 2019). Sometimes, high-performance infrastructure that facilitates dealing with Big Data in different contexts is also required. In addition, behavioural analytics are sometimes combined with Big Data Analytics for attenuating the attacker's threats properly.

Appraisal of Research Works for Cyber Security using Bid Data Analytics and Deep Learning

In this section we track the progression on work done for cybersecurity using Big Data Analytics and Deep Learning by reviewing in detail some significant research works devoted to the area.

The competitiveness of the Naive Bayesian Networks, given their basic structure and strong independence assumptions, to evaluate an intrusion detection system, using a set of KDD'99 benchmark data has been studied (Lozano et al., n.d.). The study was able to establish that by discretizing continuous domains of specific features, naive Bayesian networks can compete with advanced classification approaches such as decision trees in the intrusion detection field. The two general approaches to Intrusion detection, namely Anomaly detection and misuse detection were taken up for the study.

Bayesian Networks, depict causal relations using direct acyclic networks and communicate with uncertainty in causal relations with conditional probabilities of each node, given its parents. The Naive Bayesian networks are made up of two levels: a root node that represents a session class (normal and different types of attacks) and multiple leaf nodes, each of which has a connection feature. Because it naively assumes a strong independence relation: features are independent in the context of a session class, this method is known as naive Bayes. For discrete features, conditional probabilities were calculated by counting how many times each feature-value pair occurs with each value of the parent node. For the continuous variables, Gaussian distributions and kernel density estimation were discussed. The continuous features were normally treated as though they had a **Gaussian (i.e., normal) probability distribution**, implying that their values remained continuous. As a result, mean and standard deviation for each class value and each continuous characteristic was computed, by in turn computing the probability density function for every value. **Kernel Density Estimation** is a non-parametric density estimate for classification, which is used to generalise the naïve bayes method. This estimate makes no assumptions about the attribute values' distribution and is based on localising the observations close to each target point *ak* using a weighting function or kernel K (*ak*, *ai*) that provides a weight to each observation based on its distance from target point. The Gaussian kernel K is the more popular choice.

Three levels of attack granularities were considered depending on whether dealing with full attacks, separating them into four broad categories, or merely focusing on normal and aberrant behaviour. Before classification, Attacks belonging to the same category were combined after slightly altering the dataset. After classification, the training set was modified by grouping the attacks and each connection was labelled as abnormal or the training set was left as it is, following which each connection was identified as one of the 38 attacks. Features that had ambiguous datatypes were discretized in order to test their incidence on the results. Percent of Correct Classification (PCC) of the examples belonging to the testing was used to evaluate classification efficiency in each of the cases mentioned above. A specific form of the PCC to each class, known as the recall criterion, which shows the percentage of successfully detected connections for each class was applied. Whole attacks demonstrated that naive Bayesian networks implementing the kernel Gaussian estimator were completely consistent with the training set, implying that the latter was coherent, i.e., nearly all training examples with the same attribute values belonged to the same class.

The best strategy for dealing with the four primary attack types was chosen by classing the attacks together. In both the learning and classification phases, naive Bayesian networks based on the kernel estimator outperformed those based on the Gaussian distribution, as in the whole attack example. Gaussian distributions were preferable to kernel Gaussian estimators for classes with low representation in the training set. Cases where all attacks were gathered before classification and where collection occurred after classification were investigated utilising results from both the whole-class and five-class cases. The gap between naive Bayesian networks based on the kernel Gaussian and those that used the Gaussian distribution before gathering was greatly minimized, as the PCC became nearly identical in both techniques and had a very high rate. The analysis of continuous features was improved so that they were handled in a way that was appropriate for their values. In practically all the testing scenarios, the PCC increased (even if only slightly). Furthermore, a thorough analysis of the recall criterion in relation to the five-class example revealed that recall had increased for each of the five categories (normal, DOS, R2L, U2R, and probing).

To conclude, when focussed on five classes and normal/abnormal behaviour, it was found that it is advisable to assemble elementary attacks in their corresponding classes after the classification process rather than before it to avoid producing the worst results in general while handling continuous variables by Gaussian distributions. The kernel density estimator and discretization were found to be the best strategies in all experiments. Deep analysis revealed that in some circumstances, naive Bayesian networks produced results that were equivalent to or slightly better than the winning strategy, which was based on a combination of bagging and boosting decision tree techniques. Different experimental results given in this research supported the conclusion that naive Bayesian networks competed with advanced classification approaches such as decision trees in the intrusion detection field. Globally naïve Bayesian networks, on the other hand, performed slightly worse than these methods. Their construction, on the other hand, was significantly faster in terms of calculation.

A method for a neural network-based intrusion detection system that classifies normal and attack behaviours according to the attack type has been taken up by Moradi, et.al. in (Moradi & Zulkernine, n.d.). The study provided a solution to a multi-class problem in which attack records are separated from regular records and the attack type has been determined along with evaluating the possibility of achieving the same results with this less complicated neural network structure. The work was able to establish that the neural network-based IDS functions as an online classifier for the attack types for which it was trained. It goes offline when it's time to obtain the data needed to evaluate the features. Detecting the type of attack allowed the system to recommend appropriate countermeasures in the event of an assault. The paper was able to successfully establish that there is more work to be done in the field of intrusion detection systems based on artificial neural networks.

Based on an off-line analytic approach, a Multi-Layer Perceptron (MLP) was used for intrusion detection. The number of hidden layers in various neural network structures was evaluated in order to discover the best neural network. During the training phase, an early stopping validation method was used to improve the neural network's generalisation capability. Intrusions were identified in a misuse detection-based Intrusion Detection Systems (IDS) by looking for actions that match known signatures or vulnerabilities. An anomaly detection-based IDS, looks for unusual network traffic to detect intrusions. The unusual traffic pattern can be described as a user's breach of the valid profile built for their regular behaviour, or as a violation of recognised criteria for frequency of events in a connection. Different MLP architectures were investigated in order to create a basic architecture capable of classifying network connection records. The results suggest that even a single layer of hidden neurons in an MLP produced acceptable classification results.

The dataset used in this study contained two different types of attacks, **SYN Flood (Neptune),** a denial-of-service attack that affects every TCP/IP implementation, **Satan**, a probing attack that examines a network of computers automatically seeking information or known vulnerabilities, to test the intrusion detection system's ability to distinguish between them. Three groups of features were included. Group 1 included features that described the commands used in the connection, explaining their properties in defining attack scenarios. Number of file creations, number of root accesses, number of operations on access control files and so on are some examples. Group 2 included features that described the connection specifications and technical aspects. Protocol type, flags, duration, service types, number of data bytes from source to destination, and so on are examples of this group. Group 3 consisted of features that described the last two seconds of connections to the same host. Number of connections with the same destination host and utilising the same service, percentage of links to the current host with a rejection error, percentage of distinct services on the current host, and so on are examples of this group.

The output layer consisted of 3 neurons with the output states as follows: [1 0 0] for normal conditions, [0 1 0] for Neptune attack, [0 0 1] for Satan attack. In execution, the neural network's output occasionally revealed additional patterns, such as [1 1 0], which were considered unimportant. The MLP networks were implemented using MATLABTM Neural Network Toolbox. The training began by inputting training feature vectors as well as the desired outputs to the neural network. All the neural networks that were implemented contained 35 input neurons (equivalent to the feature vector's dimension) and three output neurons (equal to the number of classes). The number of hidden layers and neurons in each were parameters used to optimise the neural network's architecture. For training, the error back-propagation method was utilised. Early stopping method was utilised to solve the overfitting problem by increased generalisation. The training was repeated 200 times with no early stopping validation applied. The course took more than 25 hours to complete. The error dropped to an exceptional level (comparable to zero).

As a result, good classification results were expected. This assumption was validated by the final correct classification rate on the training set, which was very close to 100 percent. When unseen data (test set) was input to the neural network, however, the outcome was unfavourable. The percentage of correct classifications was less than 80 percent. Whereas after applying the Early Stopping validation method, the correct classification rate on the training set fell slightly (98 percent vs. 100 percent in the first trial), as expected. Instead, when unseen data (test set) was supplied to the neural network, the results were far superior to the first experiment, which did not use the early stopping strategy. The correct classification rate was over 90 percent, representing an increase of 11 percent. (From 80 percent in the first experiment). Another benefit of using the early stopping strategy was that the training time was reduced since the number of training epochs was minimized by early stopping. This implementation took less than 5 hours to train and validate, which is an upgrade over the prior experiment's 25-hour training duration. Training was also performed on a two-layer neural network which resulted in a lower classification rate than the best three-layer neural network, but the gap was only 4 percent.

A Machine Learning-based malware detection technique that uses more than 200 features gathered from both static and dynamic analysis of Android apps and provides efficient Deep Learning approaches for preventing malware from spreading and identifying malware with good accuracy has been proposed by Yuan, et.al. in (Yuan et al., 2014). The study established that the Deep Learning technique is particularly well suited to detect Android malware, with a high level of 96 percent accuracy when real-world Android application datasets were used.

Because of its openness, the Android market has been a popular target for malware attacks, posing a severe security and privacy risk to network users. The widespread usage of the Android operating system and development of apps coincided with an all-time high quantity of Android malware, which reached 9,411 sophisticated malware per day in 2018. This meant that fresh malware appeared every ten seconds. Antivirus signature-based and heuristic-based identification approaches only identified malware that had already been discovered, limiting their detection efficiency.

To detect malware, the following steps were followed, firstly, more than 200 features were extracted from each Android app's static and dynamic analyses, and secondly, a Deep Learning technique was used to distinguish malware from normal apps.

Static and dynamic analysis was divided into three categories, mandatory permission, sensitive API, and dynamic behaviour. Static analysis was used to extract required permission and sensitive API, whereas dynamic analysis was used to extract dynamic behaviour. In the case of static analysis, only an Android app's .apk file was needed. The main concentration was on parsing the two files 'AndroidManifest.xml' and 'classes.dex' upon uncompressing the apk file. The permissions an app required were determined by analysing the 'AndroidManifest.xml' file. The sensitive API was accessed by parsing the 'classes.dex' file.

In the instance of dynamic analysis, the 'DroidBox3' was used to run an app's .apk file. DroidBox is a sandbox for Android that has been built on TaintDroid. DroidBox may generate an information log after an app has been run for a set amount of time, which includes the program's dynamic behaviours.

A semi-supervised Deep Learning technique was employed that consisted of two phases: a "unsupervised pre-training phase" and a "supervised back-propagation phase." The Deep Belief Network (DBN) was used in the pre-training phase. The DBN was formed in a hierarchical manner by stacking a number several restricted Boltzmann machines (RBM), with the deep neural network being seen as a latent variable model, which was useful for better describing Android apps. The pre-trained neural network was fine-tuned with labelled values in a supervised way during the back-propagation phase. Following that, the entire Deep Learning model was formed.

The malware sample (250 samples) came from the well-known contagio mobile4, whereas the normal app set came from the list of top 50 apps in the Google Play Store. The two important parameters to be considered when developing a Deep Learning model were the number of layers in DBN and the number of neurons in each layer. In addition, the Deep Learning model was compared to five other commonly used Machine Learning models. The grid search technique was used to maximise the Machine Learning models for the best accuracy. Thus, the modelling results showed that the Deep Learning technique was particularly well suited to detect Android malware, with a high level of 96 percent accuracy when using real-world Android application datasets and the following tool was automated for future malware detection.

The superiority of detection performance obtained by the hybrid detection method over the detection performance obtained by a single DBN and the proposed strategy that improved detection performance while reducing time complexity has been discussed by Li et al. (2015).

The study established numerous Deep Learning models that predicted malicious code behaviour were seen to produce better results than using a surface learning model. The study began by minimizing the dimensionality of data using the Autoencoder Deep Learning algorithm. This method used nonlinear mapping to convert intricate high-dimensional data into low-dimensional codes, lowering the data's dimensionality and identifying the data's major attributes before utilising the DBN learning algorithm that detected fraudulent code. The DBN was made up of multilayer Restricted Boltzmann Machines (RBM) and a BP neural network layer. Then the output layer of the final layer of RBM acted as the input layer of Back Propagation neural network which relied on unsupervised training of each layer of RBM, further conducting supervised training on the BP neural network, and lastly fine-tuned the entire network to produce the ideal hybrid model.

Malicious code detection was divided into two approaches based on the detected position, **Network-based detection methods,** which included honeypot-based approach, deep packet inspection approach, and **Host-based detection methods,** which included check sum-based approach, signature-based approach, heuristic Data Mining approach. For the malicious code detection model performed well, two tasks were accomplished: firstly, identifying the main properties of malicious code data; secondly, developing a high-performing classifier model that accurately distinguished malicious data from regular data.

Autoencoder Dimensionality Reduction achieved the goal of reducing data dimensionality. The Autoencoder structure was divided into encoder and decoder parts. The core of Autoencoder was the code layer, which was a cross section between encoder and decoder that reflected the basic properties of nested high-dimensional data sets and determined the fundamental dimensions of high-dimensional data sets. The compressed vector of input layer termed data dimensionality reduction when the number of hidden layer neurons were less than the number of input layer and output layer neurons.

DBN Deep Learning Structure was a Deep Learning system that comprised of a supervised BP network and an unsupervised multi-layer RBM network. Each layer extracted extremely relevant implicit connections from the front layer's hidden units. The DBN training procedure was broken down into two parts: firstly, each layer of the RBM was trained unsupervisingly; secondly, the entity relation classifier using the BP neural network in the last layer of the DBN was trained supervisingly, using the output vector of the final RBM as the input vector of the BP neural network. Then, a supervised learning process like the standard BP neural network was used to fine tune the entire DBN from the rear to the front. Finally, the trained DBN model was established.

Hybrid Malicious Code Detection involved initialization and input of training samples, followed by digitization and normalisation of the data. Feature mapping was performed utilizing an autoencoder. Network parameters were used to initialise DBN classifier, inputting eigenvector with dimensionality reduction. Layer i=1. The network layer was trained layer by layer using RBM learning rules, saving the outcome along with the weights and biases. i=i+1 was set if i=max <=layer; or if i>max layer, then supervised learning for BP was performed. The test samples were used to test the trained classifier to distinguish between malicious and benign code.

Malicious code data was detected using the KDDCUP'99 dataset. RZL (Remote to Local), UZR (User to Root), Probe, DoS (Denial-of-Service), and Normal data were among the five types included. This research used 10% of the KDDCUP'99 samples as a dataset, totalling 494,021 training data and 311,029 testing data. Each data in the KDDCUP'99 dataset had 41 attributes.

The experimental results revealed that as the number of cycles increases, the suggested method outperforms the single DBN method utilised in the first experiment in terms of detection accuracy. Using autoencoder to enhance data dimension reduction appeared to be beneficial as it improved detection accuracy. Meanwhile, the detection accuracy (TP) was decreasing. Overall, the indicated method given in the research outperformed the single DBN method in terms of prediction accuracy. It adapted to a complicated environment, detected malicious code effectively, and took less time. Utilizing numerous Deep Learning models to predict malicious code behaviour produced better results than using a surface learning model. The proposed strategy improved detection performance while reducing time complexity.

An architecture for the detection of malware using Deep Learning has been proposed by Hardy, et.al. in (Hardy et al., n.d.). The paper is aimed at enhancing overall malware detection performance compared to typical shallow learning methods, by employing the stacked Autoencoders (SAEs) model for malware identification. The SAE model is based on learning the generic malware traits and led to the detection of previously undiscovered malware. The input rested on Windows Application Programming Interface (API) calls retrieved from Portable Executable (PE) files. Malware, such as trojans, backdoors, viruses, worms, and spyware, are malicious software that spread to compromise a system's secrecy, integrity, and functioning. Malware is a severe danger to the security of computers and the Internet, which are indispensable in modern life. Most of the anti-malware software packages employ a signature-based detection mechanism to safeguard legitimate users against attacks. By using techniques like encryption, polymorphism, and obfuscation, malware attackers can readily bypass this mechanism.

For unsupervised feature learning, the stacked Autoencoders (SAEs) model used a greedy layer wise training process, succeeded by supervised parameter fine-tuning (weights and offset vectors). The analysis of Windows API calls was made from the gathered PE files conducted using Deep Learning framework for malware detection (short for DL4MD). The system was divided into two parts: a feature extractor and a Deep Learning-based classifier.

The **Feature Extractor** comprised of two parts: Decompressor and PE Parser. When a PE file had been formerly compressed by a binary compress software or embedded a homemade packer, the Decompressor was used first. Each PE file's Windows API calls were extracted using the PE parser. The API requests were then saved as the PE file signatures. **Deep Learning based Classifier** architecture was based on the Windows API calls and it aimed to do unsupervised feature extraction, supervised fine-tuning, and hence malware detection.

Deep Learning has been showed as a promising framework because of its superior layer wise feature learning models, which achieved equivalent or better efficiency than shallow learning architectures like Naive Bayes (NB), Support Vector Machine

(SVM), Decision Tree (DT), and so on. **Autoencoder** is a neural network that is used to learn efficient coding. It has an input layer, an output layer, and one or perhaps more hidden layers linking them architecturally. Its purpose was to encode features of the input layer into the hidden layer, that were then decoded into the output layer, producing the same output as input layer. As a result, the hidden layer served as a duplicate of the feature space, especially in the case of a narrower hidden layer. The activations layer contained fewer nodes than the input/output layers and was thought of as a compressed version of the final hidden layer.

A Deep Learning architecture with SAEs model is formed by stacking Autoencoders together to build a deep network: the result of one Autoencoder in one layer was used as the input layer of the Autoencoder in the succeeding layer. The Autoencoders were arranged in a hierarchical stack. A classifier was added to the top layer to use the SAEs during malware detection. After that, a greedy layer wise unsupervised learning approach was used.

Greedy layer wise unsupervised learning algorithm pretrained the deep network layer by layer from the bottom up, then fine-tuned the parameters using Back Propagation from the top down to get better results.

The proposed Deep Learning system (DL4MD) also performed well in terms of detection efficiency: each unlabelled data prediction took only about 0.1 second, including feature extraction. As a result, the approach followed was a viable option for proactive malware detection in real-world industrial settings. The experimental results showed that the suggested method outperformed standard shallow learning methods in terms of overall malware detection performance, and that Deep Learning was a suitable architecture for malware detection.

CONCLUSION

This chapter has been written with the intent of familiarizing the reader with the basics of cyber-attacks, cyber security, big data, data analytics and the much-acclaimed new age technologies such as AI, Machine Learning, Data Mining, IOT and Deep Learning. Resultantly, the reader is first introduced to the basic notions of the world of Big Data Analytics, Cyber-attacks and Cybersecurity. Insights into the types of cyber-attacks within the Big Data Analytic frameworks were provided next. An in-depth analysis of how the ever-evolving and gigantically growing cyber-attack threats pose grave problems for organizations and why traditional security landscapes fail has been provided next. The advantages of using Big Data Analytics along with other intelligent technologies were itemized. To provide clarity into the detection of cyber-attacks using big data, a multi-layer data analytics framework for cyber-attack detection, featuring a technique called *Correlation Analysis* was discussed.

High volume, heterogenous, multi-source data and many other challenges for the cybersecurity in big data environments were catalogued next.

Established research in the field of cyber security, big data and deep learning has been covered in substantial detail. These research works have been able to establish that integrated deep learning and big data methods demonstrated far better performances to prevent, detect and respond to cyber-attacks as compared to the traditional cyber security methods. For instance, one of the included works clearly shows deep learning as a promising framework because of its superior layer wise feature learning model that achieved equivalent or better efficiency than shallow learning architectures like Naive Bayes, Support Vector Machine, and Decision Trees. Another included research work also established that numerous deep learning models that predicted malicious code behaviour were seen to produce better results than using surface learning models. The suitability of the deep learning techniques to detect Android malware was established by yet another work.

Looking at the aforementioned results, it would not be out of line to point out that these technologies are now just the today but will also continue to be a part of the future of the cyber security landscape.

REFERENCES

Abeshu, A., & Chilamkurti, N. (2018). Deep learning: The frontier for distributed attack detection in fog-to-things computing. *IEEE Communications Magazine*, *56*(2), 169–175. doi:10.1109/MCOM.2018.1700332

Agerri, R., Artola, X., Beloki, Z., Rigau, G., & Soroa, A. (2015). Big data for Natural Language Processing: A streaming approach. *Knowledge-Based Systems*, *79*, 36–42. doi:10.1016/j.knosys.2014.11.007

Amor, N. B., Benferhat, S., & Elouedi, Z. (2003, July). Naive bayesian networks in intrusion detection systems. *Proc. Workshop on Probabilistic Graphical Models for Classification, 14th European Conference on Machine Learning (ECML) and the 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD).*

Atlam, H. F., Walters, R. J., & Wills, G. B. (2018). Fog computing and the internet of things: A review. *Big Data and Cognitive Computing, 2*(2), 10.

BBC News. (2020). *New Zealand stock exchange halted by cyber-attack*. https://www.bbc.com/news/53918580

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, *27*(10), 1868–1883. doi:10.1111/poms.12838

Cisco. (2020). *Cisco Cybersecurity Reports*. https://www.cisco.com/c/en/us/products/security/cybersecurity-reports.html

Clark Study School. (2017). *Hackers Attack Every 39 Seconds*. https://www.securitymagazine.com/articles/87787-hackers-attack-every-39-seconds

da Costa, K. A., Papa, J. P., Lisboa, C. O., Munoz, R., & de Albuquerque, V. H. C. (2019). Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks*, *151*, 147–157. doi:10.1016/j.comnet.2019.01.023

de Boer, M. H., Bakker, B. J., Boertjes, E., Wilmer, M., Raaijmakers, S., & van der Kleij, R. (2019). Text mining in cybersecurity: Exploring threats and opportunities. *Multimodal Technologies and Interaction*, *3*(3), 62. doi:10.3390/mti3030062

Fraley, J. B., & Cannady, J. (2017, March). The promise of machine learning in cybersecurity. In *SoutheastCon 2017* (pp. 1–6). IEEE. doi:10.1109/SECON.2017.7925283

Garber, L. (2000). Denial-of-service attacks rip the Internet. *Computer*, *33*(04), 12–17. doi:10.1109/MC.2000.839316

Geetha, R., & Thilagam, T. (2021). A review on the effectiveness of machine learning and deep learning algorithms for cyber security. *Archives of Computational Methods in Engineering*, *28*(4), 2861–2879. doi:10.100711831-020-09478-2

Gulati, R. (2003). *The threat of social engineering and your defense against it*. SANS Reading Room.

Gupta, D. (2020). *Marriott Data Breach 2020*. https://www.loginradius.com/blog/start-with-identity/marriott-data-breach-2020/

Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *9*(4), e1306. doi:10.1002/widm.1306

Hardy, W., Chen, L., Hou, S., Ye, Y., & Li, X. (2016). DL4MD: A deep learning framework for intelligent malware detection. In *Proceedings of the International Conference on Data Science (ICDATA)* (p. 61). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Jibilian, I. (2021). *The US is readying sanctions against Russia over the SolarWinds cyber-attack*. https://www.businessinsider.in/tech/news/heres-a-simple-explanation-of-how-the-massive-solarwinds-hack-happened-and-why-its-such-a-big-deal/articleshow/79945993.cms

Kaur, S. (2021). Adding Personal Touches to IoT. In *Big Data Analytics for Internet of Things* (pp. 167–186). Wiley.

Khaliq, S., Tariq, Z. U. A., & Masood, A. (2020, October). Role of User and Entity Behavior Analytics in Detecting Insider Attacks. In *2020 International Conference on Cyber Warfare and Security (ICCWS)* (pp. 1-6). IEEE. 10.1109/ICCWS48432.2020.9292394

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*(1). doi:10.1177/2053951716631130

Li, Y., Ma, R., & Jiao, R. (2015). A Hybrid Malicious Code Detection Method based on Deep Learning. *International Journal of Software Engineering and Its Applications*, *9*, 205–216.

Mahmood, T., & Afzal, U. (2013, December). Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. In *2013 2nd national conference on Information assurance (NCIA)* (pp. 129-134). IEEE.

MetiTalk. (2016). *MeriTalk Cloudera Disruptive Press Release*. https://www.meritalk.com/wp-content/uploads/2016/08/49503-MeriTalk_Cloudera_Disruptive_Press_Release_FINAL_082916.pdf

Moradi, M., & Zulkernine, M. (2004, November). A neural network based system for intrusion detection and classification of attacks. In *Proceedings of the IEEE international conference on advances in intelligent systems-theory and applications* (pp. 15-18). IEEE.

Northeastern University. (2016). *How Much Data is Produced Every Day?* https://www.northeastern.edu/graduate/blog/how-much-data-produced-every-day/

PWC. (2015). *Information Security Breaches Survey*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/432412/bis-15-302-information_security_breaches_survey_2015-full-report.pdf

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 1–10. doi:10.1186/2047-2501-2-3 PMID:25825667

Rengarajan, R., & Babu, S. (2021, March). Anomaly Detection using User Entity Behavior Analytics and Data Visualization. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 842-847). IEEE.

Richardson, R., & North, M. M. (2017). Ransomware: Evolution, mitigation and prevention. *International Management Review*, *13*(1), 10.

Shah, Y., & Sengupta, S. (2020, October). *A survey on Classification of Cyber-attacks on IoT and IIoT devices. In 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 406-413).* IEEE.

Trustwave. (2015). *Trustwave Global Security Report*. http://book.itep.ru/depository/annuals/2015_TrustwaveGlobalSecurityReport.pdf

Twitter Investigation Report. (2020). https://www.dfs.ny.gov/Twitter_Report

Vemuri, V. R. (2005). *Enhancing computer security with smart technology*. Auerbach Publications. doi:10.1201/9780849330452

Vu, L., Bui, C. T., & Nguyen, Q. U. (2017). A Deep Learning Based Method for Handling Imbalanced Problem in Network Traffic Classification. In *Proceedings of the Eighth International Symposium on Information and Communication Technology* (pp. 333–339). Association for Computing Machinery. 10.1145/3155133.3155175

Wikipedia. (2020). *Easyjet Data Breach*. https://en.wikipedia.org/wiki/EasyJet_data_breach

Yuan, Z., Lu, Y., Wang, Z., & Xue, Y. (2014, August). Droid-sec: deep learning in android malware detection. In *Proceedings of the 2014 ACM conference on SIGCOMM* (pp. 371-372). 10.1145/2619239.2631434

KEY TERMS AND DEFINITIONS

Artificial Intelligence: AI in layman's language can be defined as study of how to make computer do things, which at the moment, people do better. It's a way to teach systems how to think as well as act humanly and rationally.

Backpropagation: A short form for backward propagation of errors. It's used to train the chain rule method's neural network. In simple terms, this technique does a backward pass through a network after each feed-forward pass to update the model's parameters depending on weights and biases.

Cyber Analyst: A cyber analyst primarily secures a company's network and systems from cyber-attacks. This comprises of researching impending IT trends, analysing suspicious activity, reporting security breaches, and educating about security precautions.

Cybersecurity: A combination of technologies, processes, and behaviours aimed at preventing attacks, damage, and illegal access to networks, computers, applications, and data.

Data Mining: Finding patterns and other valuable information from huge mountains of data sets is known as Data Mining which can also be referred as knowledge discovery in data (KDD).

Deep Learning: According to an article of *Forbes*, deep learning comes under the umbrella of Machine Learning where Artificial neural networks and various algorithms are inspired by the human brain and learn from enormous volumes of data.

Early Stopping Validation: When using an iterative method to train a learner, this type of regularisation is employed to avoid overfitting.

Intrusion Detection: The process of monitoring and analysing events in a computer system or network for indicators of intrusions, which are described as attempts to circumvent a computer's or network's security systems (compromise the security, integrity, and availability of information resources).

Intrusion Detection System (IDS): When a probable intrusion occurs, a mix of software and hardware, known as IDS, attempts to detect it and raises an alarm.

IoT: The internet of things (IoT) is a massive network of interconnected things and people that all collect and share data. In brief, the internet of things is the concept of connecting any gadget to the internet and to other connected devices.

Machine Learning: Machine learning is a branch of AI, computer science, and statistics that focuses on making most of the data and apply algorithms to mimic the way humans learn, act, and perform with aim of steadily improving accuracy of the predictions/tasks done by the machine.

Multi-Layer Perceptron: A feedforward artificial neural network which produces a collection of outputs based on a set of inputs. Several layers of input nodes are linked as a directed graph between the input and output layers of an MLP.

Predictive Analytics: Predictive analytics is a form of advanced analytics that uses historical data, statistical modelling, data mining techniques, and machine learning to create predictions about future outcomes. It is extensively used to identify dangers and opportunities in businesses.

Ransomware: It is a type of malware where ransom money is demanded after the system is being infiltrated/locked by the attacker. Some simple ransomwares may lock the machine without harming any files on the system whereas more powerful virus uses crypto viral extortion where user's data is encrypted, and ransom amount is demanded to decrypt.

Real-Time Analytics: Real-time analytics is about answering questions, making predictions, understanding relationships, and automating processes as soon as data is generated in real-time.

Chapter 9
Data Analysis of Cognitive, Behavioral, and Emotional Features Having Impact on Student Careers

Gouthami Velakanti
Kakatiya Institute of Technology and Science, India

Anjali Mathur
KL University (Deemed), India

Kalyanapu Srinivas
Kakatiya Institute of Technology and Science, India

ABSTRACT

According to cognitive data analysis theory, career thoughts of students mediate the relationship between career, stress, and the career decision state. A student who opts for higher education and is new to this environment is undergoing stress. The work is focused on identifying behavioral, cognitive, and emotional features of the student, which have an impact on professional careers. So, early recognition of these features is necessary. This chapter presents an analysis by following some statistical approaches of machine learning and mathematical concepts. The study found that an increase in career and life stress is associated with a lower level of decidedness and satisfaction with career choice among genders. The results suggests that counselling is needed to aid students in their career thoughts and decision making.

DOI: 10.4018/978-1-6684-5722-1.ch009

1.INTRODUCTION

A graduate student mental health related problem starts from the beginning of professional life and changing regularly on the basis of their requirements. Behavioral, cognitive and emotional features show impact on student academics and personal life.

Behavioral: The way in which a person behaves in response to a particular context. Cognitive: related to the process of thinking and reasoning.

Emotional: having feelings that are easily excited and openly displayed.

These features can affect student strengths by reducing their quality in performance, confidence levels and negatively impacting relationships with friends and family members. These issues can also have long-term consequences affecting their future employment, earning potential, and overall health. Depression, anxiety, stress, sleeping difficulties are some common characteristics observed in students, under behavioral, cognitive and emotional features. Our work focus on analysis of the students who are pursuing in an new environment undergoing various stress level. The dataset was created by collecting information from students of different engineering colleges and then the dataset is preprocessed which is classified into three categories. 1. Behavior 2. Cognitive 3. Emotional. The work in this paper is continued in the following way. In Python language the proposal is implemented using Statistical formula. The correlation between the features of each category is identified and compared to check the strongly correlated features and weakly correlated features. The variation between girls and boys in terms of behavior, cognitive and emotional features are visualized and observed that weak students are to be counselled. The dataset was created by collecting information from students of different engineering colleges and then the dataset is preprocessed which is classified into three categories:

1. Behavior
2. Cognitive
3. Emotional.

We used python language for implementation. Statistical formulas for standard deviation, variance is applied on preprocessed data and shown the results. We identified the correlation between the features of each category features and shown them using heatmap. For each category of features we identified strongly correlated and weakly correlated features and stated them. The variation between girls and boys in terms of behavior, cognitive and emotional features are visualized and the variation is observed.

2. LITERATURE SURVEY

In (Ahmad Slim, Gregory L. Heileman, Jarred Kozlick and Chaouki T. Abdallah, 2014), the authors propose a framework which uses machine learning algorithms and Bayesian Belief Network (BBN) to predict the performance of students in academic careers. In this BBN model, course is the only parameter considered as node which is treated as discrete value. The framework is used to track the progress of students in order to provide alerts aimed by improving student outcomes. The author considered only the course as parameter. In(Muhammad Fahim Uddin, Jeongkyu Lee,2016) The authors developed a set of algorithms to predict the educational relevance of students talents by considering the personality features. The paper focuses progress of results in Good Fit Students (GFS) algorithms and mathematical constructs. The author developed family of algorithms sequential and associated math modeling and constructs and named them as Predicting Educational Relevance for an Efficient Classification of Talent, Perfect Algorithm Engine (PAE). This consists of three algorithms. 1.Noise Removal and Structured Data Detection(NR and SDD). 2.Good Fit Student (GFS) prediction. 3.Good Fit Job Candidate (GFC) prediction. The work is aimed at predicting only good students.

In (Mushtaq Hussain, Wenhao Zhu, Wu Zhang, Syed Muhammad Raza Abidi, Sadaqat Ali,2018) The author presented a model which identifies student difficulty and predict the students performance. The data is analyzed by Technology Enhanced Learning (TEL) system called Digital electronics Education and Design suite (DEEDS) using machine learning algorithms. The author collected data from B.Sc students of University of Geneva. The data consists of 100 students completing 6 design exercises each of five sessions. The tool developed by author predicts student performance according to status of tasks done in current session. The ANN, SVM, LR, NBC and DT models are developed using MATLAB. Out of all these several training author concluded DT is good but a better algorithm like K-means can also be used. In(Liya Claire,Joy Asha Raj, 2019)The author presented various placement predictor systems which predict the possibility of selecting a company to which the pre-final year student fits for. The author surveyed many papers and stated that the student dataset containing academic and placement details are the rich sources to predict placement chances. This kind of prediction system makes the student to decide which career would suit them best. No new proposal is defined by the author.

In (Rasha Shakir Abdulwahhab, Shaqran Shakir Abdulwahab,2018) The author proposed a model using Linear Analytics(LA) that takes the students electronic records of students data that includes students activity and marks which tells that who is at risk in the career and who is facing problems. The method is tested in terms of accuracy and the results are better compared to the models Dependency Graph (DG) and Prediction by Partial Matching (PMP). The author collected data that is the students transcripts from department of IT at college of Applied Science Oman. The prediction model is built for a particular course and stored in database and predict if student falls into risk or not. Interpretation and finding of CPT+ model, that is grading predicted by CPT+ would confuse students. In(Queen Esther Booker, 2009). In this paper the author propose recommendation system prototype which guides the student in selecting particular course depending on student choice and history of the student as well. The author developed the system in such a way that the student will enter the details of their interest. The author used fabricated data, and used K-Means algorithm for clustering. The interest of the student collected were nearly 12 then after applying the algorithm it resulted with two interests. In (K. Sripath Roy, K. Roopkanth, V. Uday Teja, V. Bhavana, J.Priyanka,2018) According to author it is difficult for a student to asses their capabilities and choose right career. In this research the predicted the student's performance by using different machine learning algorithms. Author collected data from employees of different organizations. Some data from Linked in, some data from college alumni. Around 20 thousand records with 36 columns each is collected and preprocessed and encoded into numerical form using One Hot Encoding. The machine learning algorithm SVM, decision tree and XGBoost are used. The accuracy is considered by these algorithms and the most accuracy generating algorithm is analyzed as better performance and prediction is done.

In (Tajul Rosli Razak, Muhamad Arif Hashim, Noorfaizalfarid Mohd Noor, Iman Hazwam Abd Halim, Nur Fatin Farihin Shamsul,2015) The author states that a student choosing the career in correct path is very difficult and if at that time the decision taken is strong would lead to worst effect. The author developed this recommendation system for Malaysia students. The author used the technique of Fuzzy logic. To recommend student in career selective it is important to built Career Path Selection Recommendation System (CPRS) to UITM students. The system was developed by considering information of students from Bachelor of Science in Information Technology. This system includes two entities i.e. Counselor and Students. The Councilor will provide data of Skill Set required for Career. The information is stored in knowledge base. The students are supposed to answer the career test through Web

– Based Application. The system then generates suitable career for them by using fuzzy logic approach involving fuzzification, fuzzy interference and de fuzzification. In (Huma Samin And Tayyaba Azim, 2019). The recommender system supports the users in decision making by enhancing the ability and thinking power. The author developed a recommender system for academia by constructing a National Database of computer scientist s resumes where professional experiences are shared. As a recommender system is for higher education the Data Set is labeled as Pakscholarscan from which features were extracted. The recommender system takes the pre processed data of faculty resumes and student research proposals then the topic models are generated to learn various themes and author topic relationship. When the distributions are modeled distance metrics are used to find similarities to any given topic. The author collected data which are abstracts of Research from NIPS conference proceedings and also the proceedings from SIGCOMM, KDD, and ICIP. And also, the data consist of faculty resumes of higher education. The data is preprocessed and used as train data to train Latent Dirichlet Allocation (LDA) and Author Topic Model (ATM) models. The performances of these models are evaluated. The author concludes that compared to Author Topic Model ATM, Latent Dirichlet Allocation LDA is better in performance.

In (A. T. M. Shakil Ahamed, Navid Tanzeem Mahmood & Rashedur M Rahman,2017) The author focused on education effectiveness. The author surveyed large number of students from Bangladesh and developed a model based on data balancing, dimensionality reduction and normalization to predict the performance of a student. The author designed a questionnaire by consulting many experts which contains relevant attributes.423 students were surveyed on different factors like psychological, socio-economic, academic. The author used SMOTBOOJ technique for data balancing. For dimensionality reduction principle Component Analysis (PCA), GHA and SOM techniques are used. Artificial Neural network (ANN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) are used for predicting performance of the student. The author concludes that Artificial Neural network (ANN) is best.

3. PROPOSED APPROACH

In higher educational institutes, the students of every semester and every year goals are different and accordingly their mental behavior and attitude varies. For example, a student at schooling their major target is scoring high percentage while coming to UG/PG level the goal changed towards to professional life and getting campus recruitment. The behavior of students at schooling are stress-free but entering into UG/PG may undergo different levels of stress. So, there is a necessity of identifying the behavior features of students who need counselling at right time in right direction. The features identified vary from graduation to post graduation and are classified based on gender.

Our research is carried out by collecting data from the engineering students of first, second, third, final year and post graduate students as shown in fig 3.1. A questionnaire is prepared which consists of the characteristics to identify student nature as well as behavioral, cognitive and emotional features. Data collection is done from the questionnaire that consists of around 87 questions where each student should answer by using the options provided according to the questions. The questions are basically related to certain common conditions that focuses to know feelings, opinions and decision making. The questionnaire is supplied to the students of various engineering colleges by preparing Google form. The features considered and classified are shown below.

The dataset is preprocessed and the features of the students are classification into three categories mainly behavioral, cognitive and emotional. Python language is used to apply all statistical methods to visualize the categorized data. The results shows the variance between boys and girls in the behavior, cognitive and emotional features.

Figure 1. Hierarchy chart of Students in creation of datasets

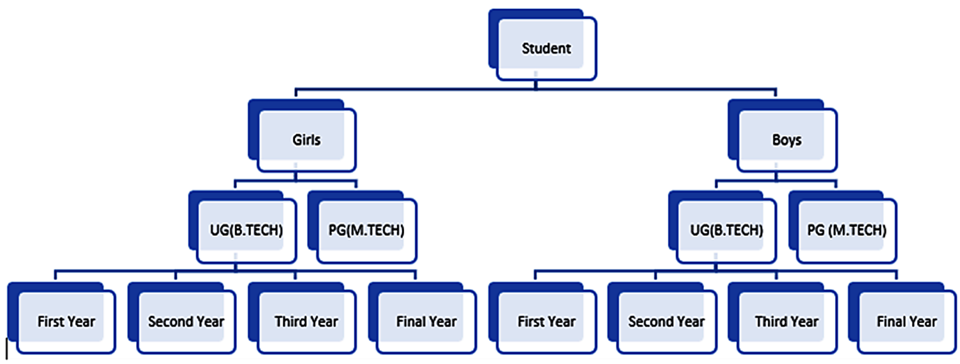


Figure 2. Preprocessing of datasets

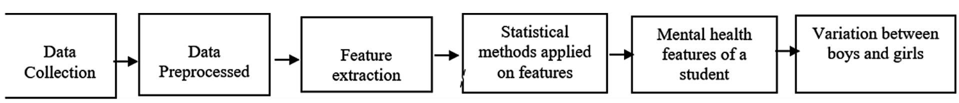
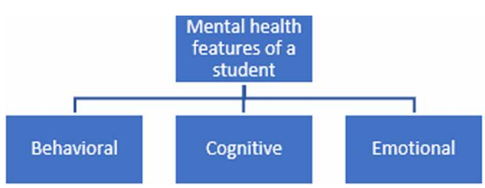


Figure 3. Classification of features



4. METHODOLOGY

From the dataset, to understand the features appropriately visualization of data is presented.

4.1 Data Analysis – A Statistical approach

Standard Deviation: Standard deviation of dataset gives the spread of data. If the standard deviation is low, data spread is low, if standard deviation is high data spread is high. Formula to calculate the Standard deviation is

$$s = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2}{N-1}}$$

Xi is the value of dataset; x is mean of samples, N is the number of observations.

Variance: This is another method to show the spread of dataset. This is simply square of Standard Deviation.

$$s^2 = \frac{\Sigma\left(x - \overline{x}\right)^2}{N-1}$$

Variance is used only for one dimensional data.

$$var\left(X\right) = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(x_i - \overline{x}\right)}{\left(n-1\right)}$$

Variance is calculated for both Boys, Girls, related to behavior, cognitive and emotional features which is represented in tabular form.

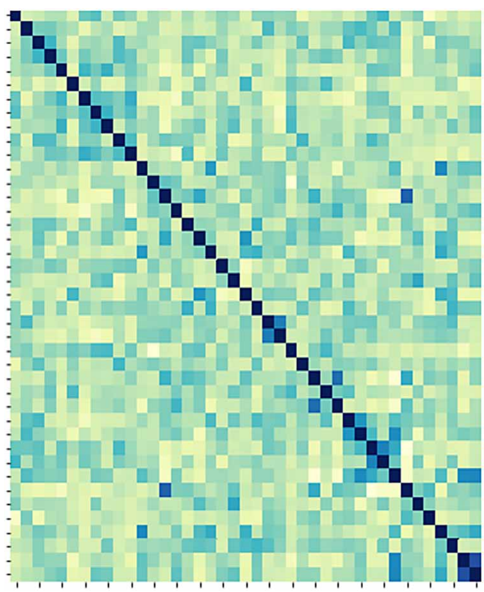
Correlation: Correlation is a statistical technique that is used to determine how one variable is moving in relation with another variable. If two variables are related closely then prediction is easy. For this, heatmaps are used to show the correlation between the features. From the dataset all the features are categorized into three different categories

1.Behavioral 2.Cognitive 3.Emotional

The correlation between features category wise is identified. So, heatmap is generated as follows for each category of features. From each category, identified the strongly correlated and weakly correlated features. From this identification some set of features can be minimized which will be used in the further research. The correlation between the features in behavioral feature category is shown.

Figure 4. Heatmap showing the correlation between behavioral features

Figure 5. The correlation between the features in Cognitive feature category is shown.



4.2 Observations and Results

The main intention is to show the variation between girls and boys in terms of behavioral, cognitive and emotional features. Human psychological features changes accordingly which are considered for our work. Standard deviation is calculated for both Boys, Girls, related to behavior, cognitive and emotional features which is represented in tabular form.

The variance is calculated for both Boys, Girls, related to behavior, cognitive and emotional features which is represented in tabular form. The variation is shown in graphical form.

Variation of behavioral, cognitive and emotional features in boys and girls.

Cognitive

From the above graph it can be stated that y-axis represents the intensity of the cognitive features.0 represent low, 1 represent normal, 2 represent high. From the graph, the identified behavioral features of girls and boys are little different.

Figure 6. Correlation between the features in emotional feature category.

Behavioral

From the above graph it can be stated that y-axis represents the intensity of the behavioral features.0 represent low, 1 represent normal, 2 represent high. From the graph the identified behavioral features of girls and boys are almost same.

Figure 7. Graph showing the variance of cognitive features between girls and boys

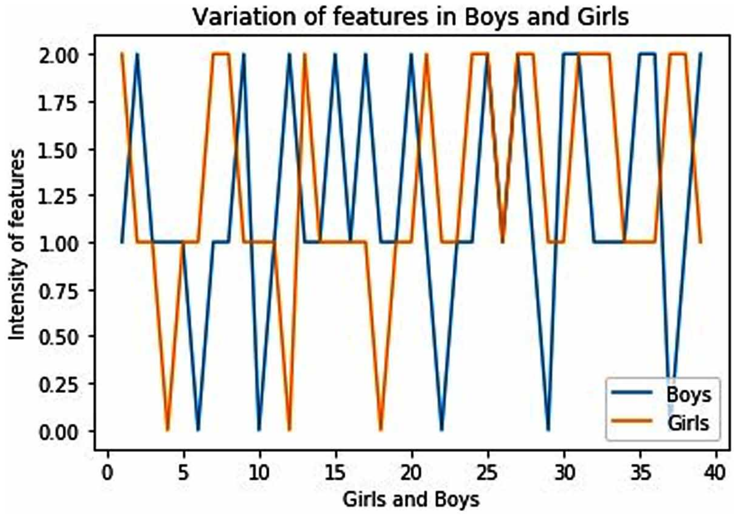
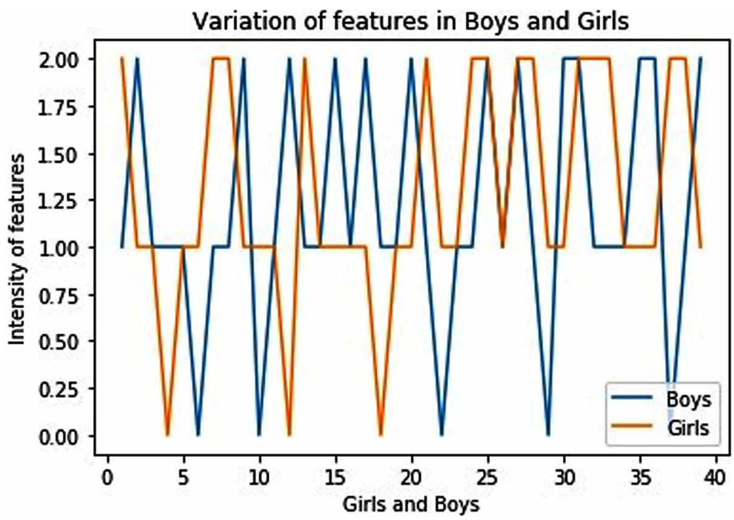


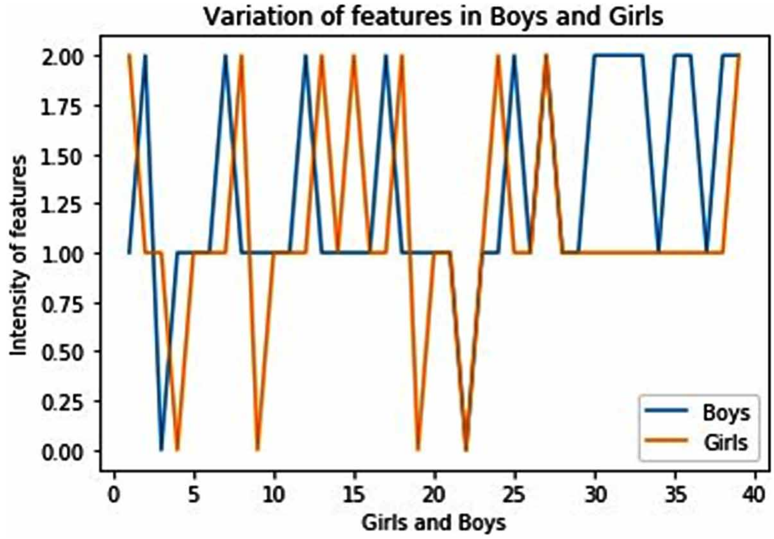
Figure 8. Graph showing the variance of behavioral features between girls and boys



Emotional

From the above graph it can be stated that y-axis represents the intensity of the emotional features.0 represent low, 1 represent normal, 2 represent high. From the graph the identified emotional features of girls are less compared to boys.

Figure 9. Graph showing the variance of emotional features between girls and boys



5. CONCLUSION

From the work carried out, the observations made are there is difference between girls and boys in terms of thinking level, decision making and almost all characteristics. So, we can conclude that the mental health status will never be same for girls and boys though the context is same. We used a small dataset in our work, better results can be found if the dataset size is increased in order to identify the features that are strongly showing impact on the student behavior. In the future work, the students shall be categorized into different classes 1) timid 2) coward 3) courageous and 4) normal based on the important features that shows impact on student attitude and suggest a recommender system for good future.

REFERENCES

Anila, M., & Pradeepini, G. (2018). Least square regression for prediction problems in machine learning using R. *International Journal of Engineering and Technology, 7*(12), 960-962. . doi:10.14419/ijet.v7i3.12305

Aparna, P., & Kishore, P. V. V. (2018). An efficient medical image watermarking technique in E- healthcare application using hybridization of compression and cryptography algorithm. *Journal of Intelligent Systems*, *27*(1), 115–133. doi:10.1515/jisys-2017-0266

Azim, H. S. A. T. (2019). Knowledge Based Recommender System for Academia Using Machine Learning: A Case Study on Higher Education Landscape of Pakistan. *IEEE Access: Practical Innovations, Open Solutions*.

Babu, B. S., Prasanna, P. L., & Vidyullatha, P. (2018). Customer data clustering using density based algorithm. *IACSIT International Journal of Engineering and Technology*, *7*(2), 35–38. doi:10.14419/ijet.v7i2.32.13520

Booker. (2009). A Student Program Recommendation System Prototype. *Issues of Information Technology*.

Claire, L., & Raj, J. A. (2019). A Review on Student Placement Chance Prediction. In *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE.

Hussain, M., Zhu, W., Zhang, W., Syed, M. R. A., & Ali, S. (2018). Using machine learning to predict student difðculties from learning session data. In Artificial Intelligence. Springer.

Razak, T. R., Hashim, M. A., Noor, N. M., Iman, H. A. H., & Nur, F. F. S. (2015). *Career Path Recommendation System for UiTM Perlis Students Using Fuzzy Logic*. IEEE.

Shakil Ahamed, Mahmood, & Rahman. (2017). An intelligent system to predict academic performance based on different factors during adolescence. *Journal of Information and Telecommunication*.

Shakir, R. (2018). Integrating Learning Analytics To Predict Student Performance Behavior. Academic Press.

Slim, A., Heileman, G. L., Kozlick, J., & Abdallah, C. T. (2014). *Predicting Student Success Based on Prior Performance*. IEEE. doi:10.1109/CIDM.2014.7008697

Sripath Roy, K., Roopkanth, K., Uday Teja, V., Bhavana, V., & Priyanka, J. (2018). Student Career Prediction Using Advanced Machine Learning Techniques. *IACSIT International Journal of Engineering and Technology*, *7*(2.20), 26. doi:10.14419/ijet.v7i2.20.11738

Uddin & Lee. (2016). Predicting Good Fit Students by Correlating Relevant Personality Traits with Academic/Career Data. IEEE.

Chapter 10
Big Data Analytics in Industrial IoT and Cybertwin

Rajendran T.
https://orcid.org/0000-0002-0484-2921
Rajalakshmi Institute of Technology, India

Surya S.
Saveetha Engineering College, India

Mohamed Imtiaz N.
HKBK College of Engineering, India

Babu N.
Siddharth Institute of Engineering and Technology, India

ABSTRACT

The internet of things (IoT), big data analytics, artificial intelligence (AI), and cybertwin, as well as other digital technology and designed intelligence have accelerated the 4th industrial revolution known as Industry 4.0. Industry 4.0 applications must construct complicated machine representations from such fundamental pieces, which is a time-consuming, error-prone, and wasteful process that impedes machine and plant mobility. Cybertwin, a comprehensive solution for fast Industry 4.0 application creation, testing, and porting, is proposed in this study. The deployment of cybertwin with IIoT will enhance the efficiency and accuracy of real-time IIoT applications. Further, these huge mixtures of data will be analyzed by using big data analytic tools to produce intensive incident commands, and it is further deeply analyzed to discover various knowledge, which supports redesign and reengineering of the specific process. The cloud computing platform will be utilized to achieve big data analytics effectively.

DOI: 10.4018/978-1-6684-5722-1.ch010

I. INTRODUCTION

The Internet of Things (IoT), Big Data Analytics, Artificial Intelligence (AI), and Cyber twin, as well as other digital technology and designed intelligence, have accelerated the 4th industrial revolution known as industry 4.0. Using high-fidelity and high-value data from machines, workers, and products, Industry 4.0 aims to improve industrial processes. Individual sensors and actuators, rather than complete machines, are supported by industry 4.0 application development on commercial IoT platforms that enable segregated development and runtime environments. As a result, Industry 4.0 applications must construct complicated machine representations from such fundamental pieces, which is a time-consuming, error-prone, and wasteful process that impedes machine and plant mobility.

The IIoT sensors are deployed to gather various information during industrial operations. Every second each sensor produces numerous data that be in mostly in unstructured format and some sensors produce semi-structured data. All these types of data are collected and processed to extract useful information and generate commands for immediate actions that will be helpful to carry out the operations.

Cybertwin, a comprehensive solution for fast Industry 4.0 application creation, testing, and porting, is proposed in this study. The deployment of cybertwin with IIoT will enhance the efficiency and accuracy of real-time IIoT applications. Further, these huge mixtures of data will be analyzed by using Big Data analytic tools to produce intensive incident commands and it is further deep analyzed to discover various knowledge which supports redesign and reengineering the specific process. The Cloud Computing platform will be utilized to achieve big data analytics effectively.

II. IMPORTANCE OF INDUSTRIAL IoT IN INDUSTRY 4.0

The technology is disturbing businesses by setting out new open doors for them to gather information and exactly investigate it. As more ventures are currently utilizing IoT, its degree of complexity empowers them to help proficiency, fulfill the developing needs, and drive better client encounters.

IoT helps fabricate by further developing cycles. The principal, more direct, thought is the means by which IoT can help the business by further developing cycles, upgrading creation productivity and quality, and assisting with bringing more noteworthy adaptability into the business.

Modern IoT can interface machines, apparatuses, and IIoT sensors on the shop floor to give process architects and directors much-required perceivability into creation. For instance, associations can consequently follow parts as they travel through gatherings utilizing sensors, for example, RFID and break radiate.

The Benefits of Industrial IoT

- More prominent energy proficiency.
- Decreased expenses.
- Better quality items.
- Further developed dynamic potential.
- Less gear personal time.

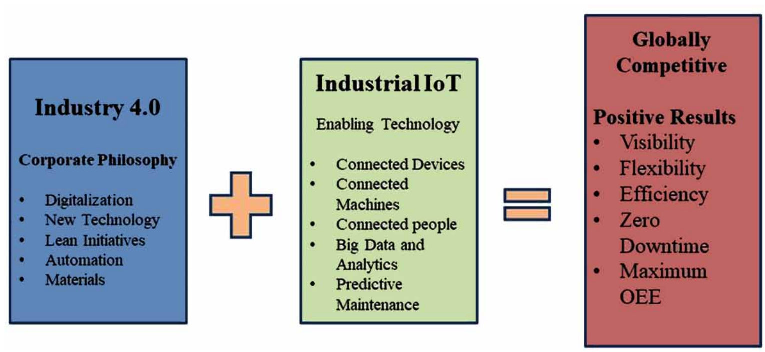
Industry 4.0, additionally alluded to as IIoT or brilliant manufacturing, emphasis actual creation and tasks with shrewd computerized innovation, Artificial Intelligence(AI), and big data to make a more comprehensive and better-associated biological system for organizations that attention to assembling and producing network on the board of Bigdata environment.

Internet of Things (IoT) is a basic part of Industry 4.0. It has broad applications in the checking of creation frameworks in assembling and administration. This innovation opens up more current and imaginative potential outcomes in assembling by working with better execution. IoT's significant capacity is to gather and impart data with the assistance of web-associated machines and gadgets. It is related to special recognizable proof numbers or codes (Adams, 2010); that can be controllable through our everyday utilization of gadgets like cell phones. This innovation's significant parts are programming, equipment with the organization's network for information fight, and assortment. IoT makes problematic advancements in the field of assembling. The need is to comprehend this innovation and how it can help the contemporary creation frameworks. Here, we have concentrated on the capability of this innovation to give better arrangements in Industry 4.0. The significant drivers of IoT for Industry 4.0 are examined. This study deals about how Industry 4.0 provides a envisioned industrial facility. Finally, we have distinguished and represented Bigdata abided with IoT applications to take on Industry 4.0 effectively, and the equivalent is introduced in plain structure. With the appropriate execution of innovation represented in Figure 1 (Asur & Huberman, 2010); ventures notice an improvement in productivity during the assembling of items. Fabricating is finished with lesser expenses and mistakes. Notwithstanding, there is far to receive full rewards for mankind.

III. SPECIFICATIONS OF INDUSTRIAL IoT

The IIoT sensors are deployed to gather various information during industrial operations. Every second each sensor produces numerous data that be in mostly in unstructured format and some sensors produce semi-structured data. All these types of data are collected and processed to extract useful information and generate commands for immediate actions that will be helpful to carry out the operations. Using high-fidelity and high-value data from machines, workers, and products, Industry 4.0 aims to improve industrial processes. Individual sensors and actuators, rather than complete machines, are supported by industry 4.0 application development on commercial IoT platforms that enable segregated development and runtime environments.

Figure 1. Importance of Industrial IoT (Asur & Huberman, 2010).



The modern web of things (IIoT) is the utilization of shrewd sensors and actuators to improve producing and modern processes. Connected sensors and actuators empower organizations to get on failures and issues sooner and set aside time and value, while supporting business insight endeavors.

The modern web of things alludes to interconnected IIoT sensors, instruments, and different gadgets arranged along with PCs' modern applications, including assembling and energy the board. Figure 2 represents the IIoT Architecture depicts the ecosystem of IIoT.

Information Processing, Analytics, Business Application (Cohen et al., 2009; Cuzzocrea et al., 2011); and Automated Process Databases are the critical elements of IIoT. These incorporate IoT Platform and On-premises Server which goes about as the connection point between the IoT door, Edge entryway to the Applications and key elements of IIoT. The last layers are the Sensors, Actuators and Edge Nodes.,

The ANSI/ISA-95 worldwide standard is utilized internationally across all modern areas to foster computerized interfaces among big business and control frameworks. The standard gives a steady phrasing to data and task models that can be utilized to portray data streams for cluster and nonstop cycles.

Figure 2. IIoT_infrastructure (Bakshi, 2012).

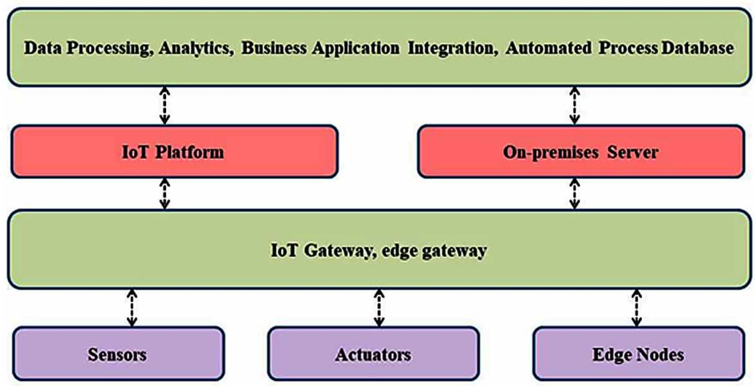
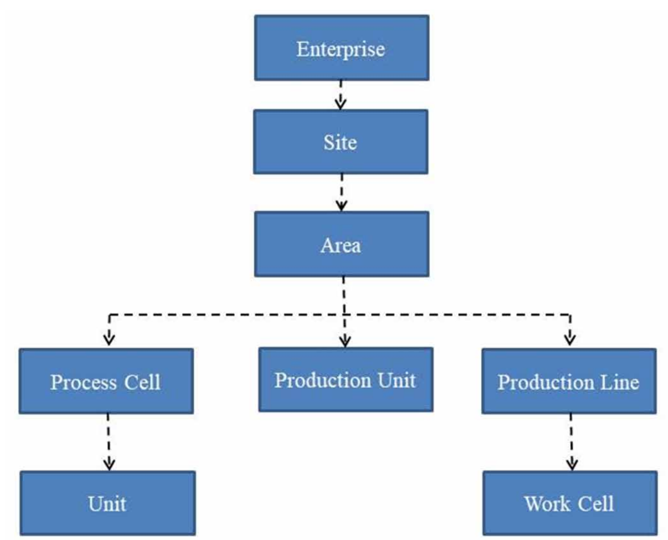


Figure 3 the ISA-95 Equipment model permits you to make a hardware model that matches the ordered progression model depicted in the ISA-95 norm and guides estimations related to resources for things in the I/O model. All the article classes that can be made in the ISA-95 Equipment model have an identity in the ISA-95 ordered progression model and exhibit similar item linkages permitted in the norm. For instance the Enterprise object is the most elevated in the model pecking order and can be comprised of youngster Site objects. This proceeds right down the progressive model to Control Modules (Unit, 2012; Elgendy, 2013).

Figure 3. ISA95 Hierarchical model (Mohamed et al., 2012).



IIoT technologies:
Cloud Computing:

IIoT creates a ton of information. Numerous organizations that utilize IIoT go to distributed computing to store and back up this information. Distributed computing offers various sorts of clouds from which organizations can pick.

Edge Computing

Edge Computing is a type of registering that is done nearby or almost a specific information source, limiting the requirement for information to be handled in a far-off server systen.

Data Mining:

IoT Data mining: The immense organization of gadgets produces another kind of information known as the IoT large information. The greatest test in the present information mining world accompanies a few issues like information stockpiling, the board, protection, security, and handling limits like continuous/streaming information (He et al., 2011).

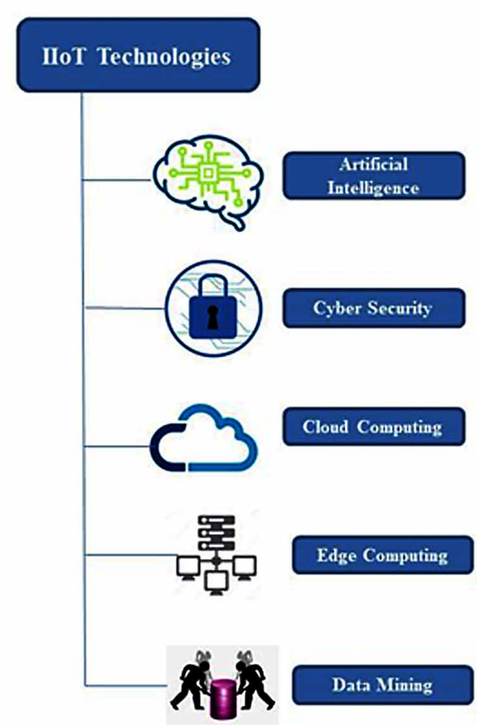
Key Technologies of Industry 4.0:

- Internet of Things (IoT)
- Advanced Robotics
- Artificial Intelligence (AI)
- Cloud Computing & Big Data
- Cyber Security
- Additive Manufacturing
- Augmented Reality (AR) & Virtual Reality(VR)

Concepts on Cybertwin

The concept of Cybertwin arise as a result of the significant rise in the digitalization of the entire globe in Industry 4.0. The primary goal of Cybertwin development is to support Industry 4.0 by making the manufacturing environment more flexible, productive, and resource-efficient. The most important constraint in the decision that has been reached more rapid and progressive is reducing the connection speed to a user's demand, which can only be addressed if the network is speedy enough to handle everything in a single blink. This fast network can benefit from Cyber-assistance. twin's Before going on to the idea of caber-twin, a few essential terminologies must be understood in order to comprehend the concept's goal (Herodotou et al., 2011; Wamba et al., 2015).

Figure 4. IIoT Technologies (Elgendy & Elragal, 2014).



A realistic method (Lee et al., 2011; Manyika et al., 2011); to either genetic element of a computer crimes twin that determines the physical function's response is known as the Cybertwin. In a way, the Cybertwin acts very much like its genuine counterpart. The cyber-core twin's approach incorporates all necessary data and their accompanying linkages. The system's numerous detectors have saved the information in this version, which can be utilized to undertake operations later. As a result, this Cybertwin architecture includes a database that stores the data required for physical twin operations. The Cybertwin creates an algorithm that changes the physical components' behavior.

IV. DEPLOYMENT OF CYBERTWIN IN INDUSTRIAL IOT

To get everything rolling with the sending of the IIoT Platform, clone the vault from the order brief or terminal. Begin the directed organization, the content will gather the necessary data, for example, Azure record, membership, target asset and gathering and application name

One model is the auto business, which involves IIoT gadgets in the assembling system. Modern IIoT sensors gather information regarding soil supplements, dampness and that's just the beginning, empowering ranchers to create an ideal communication. The oil and gas industry likewise utilizes modern IoT gadgets.

The present outstanding modern web of things (IIoT) development makes a significant requirement for vigorous network protection practices and obvious norms that give clients certainty that their associated gadgets will work safely all through their whole lifecycle (Mouthami et al., 2013).

Cybertwin uses the capacities of organizations and serves in various functionalities, by recognizing advanced records of exercises of people and things, from the Internet of Everything (IoE) applications. Cybertwin arises as a promising arrangement alongside cutting edge correspondence organizations, i.e., 6G innovation, in any case, it builds extra difficulties at the edge organizations. Persuaded by the previously mentioned viewpoints, in this work, we present a new cybertwin-driven edge system (Plattner & Zeier, 2012; Russom, 2011); utilizing 6G-empowered innovation with a wise assistance provisioning procedure, for supporting a gigantic size of IoE applications. The proposed procedure disperses the approaching errands from IoE applications utilizing the Deep Reinforcement Learning method in view of their dynamic assistance prerequisites. Other than that, an Artificial Intelligence-driven procedure, i.e., the Support Vector Machines (SVM) classifier model is applied at the edge organization to break down the information and accomplish high exactness. The reproduction results throughout the ongoing monetary datasets exhibit the viability of the proposed administration provisioning methodology and SVM model over the standard calculations as far as different execution measurements. The proposed system diminishes the energy utilization by 15% over the benchmark calculations, while expanding the forecast precision by 12% over the older models.

Structure of Cyber-Twin System

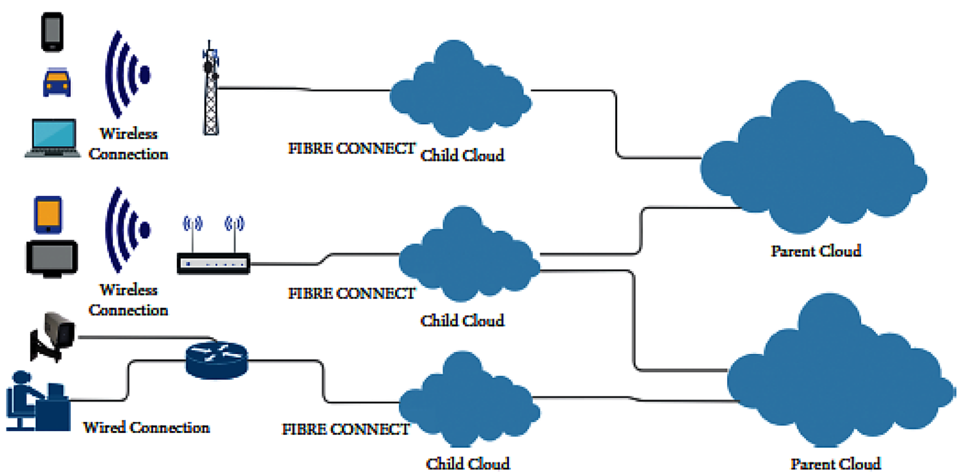
Parent cloud: The network's structure is composed of different skies that are combined to form the master cloud. Any type of link between these groups, such as spacecraft or fiber optics, can be used. The difference between the current cloud and the upcoming sky is that the latter will be utilized for more than just information storage and computation; rather, these parental systems will be used for computation, storing, and effectively sharing resources between different end-users.

Child cloud: With the master sky at the high and the user interaction at the bottom, this arrangement is a three-level hierarchy. The child sky is located between the master and the baby clouds. As a result, the kid cloud is utilized to deliver functionality to the user via the parent cloud's user interface. Apart from that, the master cloud provides the user with additional features than the junior cloud. It can swiftly handle the client requests because it is near to the user experience.

User interface: The user experience has been used to provide connectivity among clouds and customers, and the subscriber refers to the numerous computers on the internet that the authorized consumer is utilizing. If a user decided to discuss assets or access files with another device in the network, they must first send the request to the kid cloud, which is then passed to the parent sky. In order to perform any computation, the user must be hosted in the cloud at all moments through Cybertwin, rather than on the computer.

Cyber-twin– The most important part of the forthcoming network is this. It's next to the child sky, and the connectivity it delivers is different from what a broadband network delivers; as a result, it is capable of enabling a wide range of services not currently supported by traditional network models as shown in Figure 5.

Figure 5. Constitution of Cyber-Twin (Sánchez et al., 2008).



Deployment process

In the instance of computer hackers, numerous ISPs keep track of all of the final user's log files, which a terrorist might readily access. In a peer-to-peer topology, different ISPs maintain track of all the end user's log records, which can be readily accessed by a hacker.

Any object can have a computer hacker, which is a computerized version of it. It's possible that this monster is a human or a real object. The monster has been aided in the physical domain by computer hackers. The end customer must connect to the computer hackers, which then interact with the web (child and parent) to fulfill the final client requests, whereas the peer-to-peer connection needs the end consumer to join back to the cloud in order to access resources. As a result of the cyber-ability twins to understand the needs of the user, the user receives a higher level of service in the original situation (Pignaton et al., 2020).

V. BIG DATA ANALYTICAL TOOLS AND MECHANISMS

The term "Big Data" has successfully been applied to databases that have become so large that utilizing traditional database management tools to interact with them has become more problematic. They are data points that are too enormous to be gathered, stored, managed, and processed in a reasonable length of time using commonly used computational methods and storage solutions. Big data is becoming increasingly large, with individual data sets varying from a few handful terabytes (TB) to several petabytes (PB). As a result, some of the challenges associated with big data include data collection, storage, search, sharing, analytics, and data visualization. Today, businesses are sifting through massive amounts of extremely detailed data in order to uncover previously unknown truths. As a result, big data analytics is the application of advanced analytic techniques to large data sets. Business change is revealed and leveraged through analytics based on huge data samples. However, the more data there is, the more difficult it is to manage. This article will go over the characteristics of big data along with its significance. Naturally, analyzing higher and more complicated data sources that need true or close capabilities can give business outcomes; but, the latest data structures, analyticaltools, andtechniques are required. As a response, the part that follows will go over big data analytics tools and techniques in-depth, starting with big database management systems and proceeding to big data analytic computing. It then moves on to discuss some of the various big data analytics which has exploded in popularity as big data has risen in popularity.

Some of the tools (Yu et al., 2020); have been depicted below

- Talend - used for data integration and management.
- Cassandra - a distributed database used to handle chunks of data.
- Spark - used for real-time processing and analyzing large amounts of data.
- STORM - an open-source real-time computational system.

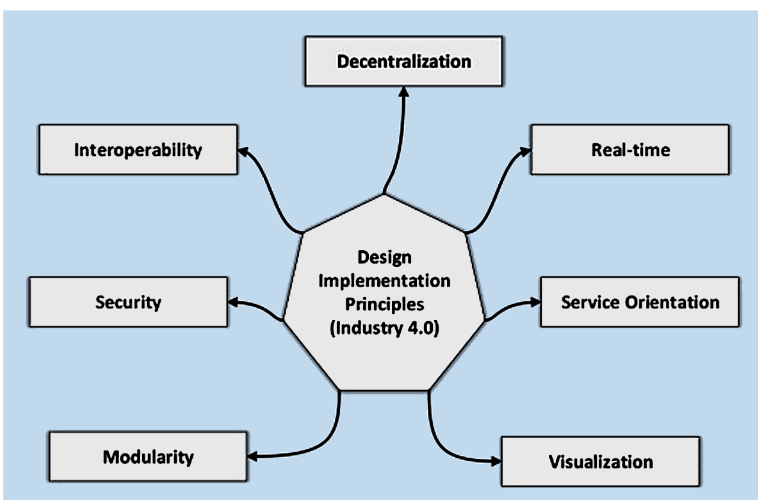
IoT in Big Data analytics helps businesses to extract information to get better business insights. Better business insights help in taking better decisions that result in high ROI. Due to an increase in demand for data storage, companies are switching to big data cloud storage that lowers the implementation cost.

Data integration tools allow businesses to streamline data across a number of big data solutions such as Amazon EMR, Apache Hive, Apache Pig, Apache Spark, Hadoop, MapReduce, MongoDB and Couchbase.

A healthy IoT environment necessitates worldwide standardization, which includes interoperability, flexibility, dependability, and operational effectiveness. Data expansion is accelerated by rapid IoT development. Thousands of connecting sensors are constantly gathering and sending information (environmental statistics, geographical information, astronomical relevant information, logistic information, and so on) to the web for storage and analysis.

Mobile platforms, transportation infrastructure, public buildings, and household appliances are the key data collection devices in the Internet of Things. The flow of massive amounts of data exceeds the capabilities of traditional enterprise IT structures and facilities. The real-time nature of the analysis has a considerable impact on computer capacity (Lim et al., 2020).

Figure 6. Design Implementation (Juneja et al., 2021).

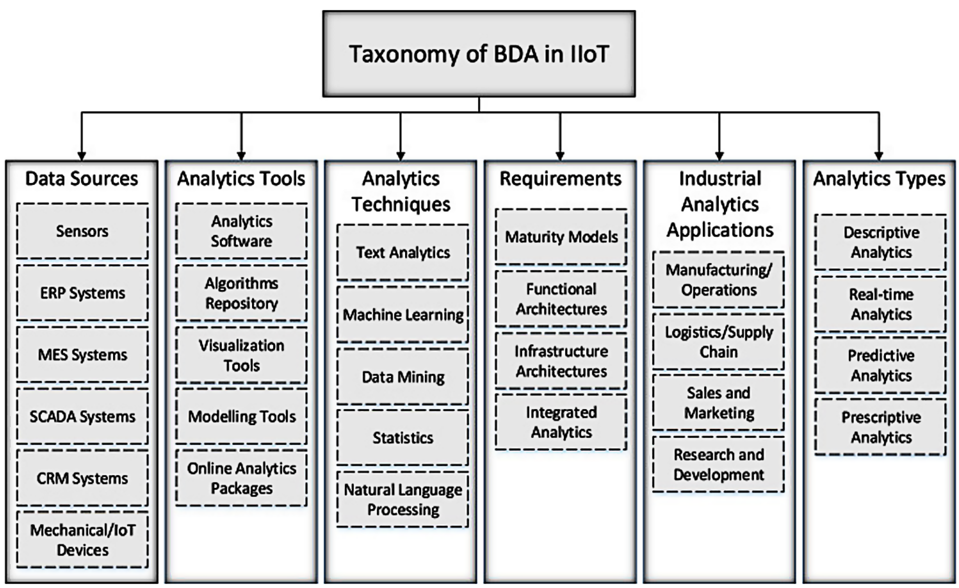


The development of Big Data by IoT has shattered IoT's present data processing capabilities, necessitating the use of data science to improve solutions. The influence of big data analytics on the effectiveness of IoT is critical (Czwick & Anderl, 2020). Figure 6 depicts the design implementation principles with respect to industry standards 4.0.

Taxonomy on Big Data Analytics in IIoT:

Big Data (Bevilacqua et al., 2020); is truly assuming a significant part in the business as well as numerous different associations. With the progression of time, the volume of information is expanding. This expansion will make immense heft of information that needs appropriate apparatuses and methods to deal with its administration and association. Various methods and devices are being utilized to appropriately deal with the administration of information. A definite report of these strategies and devices is required which will assist specialists with effectively recognizing an instrument for their information and take help to effortlessly deal with the information, coordinate the information, and concentrate significant data from it. The proposed study is an undertaking toward summing up and distinguishing the devices and procedures for enormous information utilized in the Industrial Internet of Things. This report will surely help scientists and experts to effortlessly involve the apparatuses and procedures for their needs in a viable manner.

Figure 7. Big Analytics on IIoT (Muhammad et al., 2019).



There is a desire for quicker and more convenient data processing as technology advances and the volume of data coming around through organizations on a regular bases increase. It is no longer required to have a huge volume of information on hand to make fast and efficient decisions. Conventional document management processes and infrastructures are no longer capable of dealing with such massive data sets. As a response, modern big data analytics theories and techniques, as well as the appropriate structures for collecting and manipulating such information, are required. As an end, the rise of big data affects everyone from the information itself and its collecting, to the treatment, to the final determined conclusions.

As a result, Many researchers introduced the Big – Data, Analytics, and Decisions (B-DAD) framework, which integrates big data analytical tools and processes into decision-making. The framework connects various big data storage, administration, and processing tools, as well as analytics tools and methods, visualisation, and assessment tools, to the various stages of the decision-making process. As a response, big data analytics advances can be seen in three aspects: big data storing and structure, predictive analytics computing, and, finally, big statistical results that may be used for information retrieval and better decision making.

This chapter will delve deeper into each topic. However, because big data is indeed constructing as an area of study and research revelations and instruments are being developed all the time, this segment is not extensive of all the potentials and instead focus on delivering a general overview rather than a comprehensive list of all possible improvements and technologies.

VI. RESEARCH DIRECTIONS TOWARDS IIoT AND BIG DATA ANALYTICS

When dealing with big data, one of the first issues that businesses must address is where and how the data will be stored once it has been gathered. Relational databases, data marts, and data warehouses are examples of classic structured data storage and retrieval technologies. Using Export, Transform, Loads (ETL) or Extract, Loading, Transform (ELT) technologies, data is retrieved from operative data storage, converted to satisfy production needs, and then imported into the computer or data warehouse (Agalianos et al., 2020). The information is purified, processed, and classified prior to getting made accessible for data gathering and advanced analytics activities.

Big data, on the other side, needs Modular, Agile, Deeply (MAD) analysis skills that are separate from those needed in a traditional Enterprise Data Warehouse (EDW). For starters, standard EDW procedures forbid the use of advanced analytics unless they have been sanitized and integrated. Because data is so pervasive currently, big data setups must be magnetized to draw in all datasets, regardless of merit. Big data storage should also allow analysts to swiftly develop and modify data, given the growing number of datasets and the intricacy of data analytics. This necessitates a flexible database, with logical and physical contents that can react to rapidly changing data. Moreover, because modern data studies involve complex statistical methods and analysts need to be able to drill deeper into vast databases, a big data warehouse must be deep and operate as a sophisticated algorithmic execution engine.

As a response, a wide range of big data solutions have been employed, including distributed environments and Massively Parallel Processing (MPP) networks for fast accuracy improvement and platform adaptability, as well as non-relational or in-memory libraries. To organize and maintain unprocessed, or non-relational, information, Not Only SQL (NoSQL) and other non-relational computers were developed. NoSQL databases aim for grand scale, master data versatility, and easier application design and implementation. Unlike sql server, NoSQL databases keep data collection and management independent. Instead of writing data management operations in database-specific languages, such databases focus on high-performance scalable data storage and allow data management tasks to be written in the application layer.

In-memory databases, on the other hand, handle data in server memory, eliminating disc input/output (I/O) and enabling real-time database answers. The core database could be kept in silicon-based memory location rather than electromechanical disc drives. This significantly improves functionality and allows for the creation of entirely new programs. In-memory warehouses are also being utilised advanced big data analytics, especially to allow faster accessibility to and rating of algorithms designed. This enables large-scale data availability as well as rapid discovery analytics.

Apache, on the other way, is a big data analytics platform that provides stable, durability, and predictability by adopting the MapReduce methodology (explained in the next chapter) and connecting storage and analysis. Hadoop's two main parts are HDFS for big data storage and MapReduce for big data analytics. The HDFS memory feature offers a robust and efficient distributed database that separates chunks of data and delivers them across cluster members. A replication strategy safeguards data between nodes, assuring accessibility and predictability even if one or both of them fails. The two categories of HDFS networks are Information Nodes and Name Endpoints. Between both the customer and the Data Processor, the Datanode acts as a regulator, guiding the consumer to the Datastore that has the required data.

Map Reduce is a parallel programming technique for big data processing that is inspired by functional languages' "Map" and "Reduce" functions. It's at the heart of Hadoop, and it handles data processing and analytics. According to EMC, the MapReduce paradigm is built on scaling out rather than scaling up, which means adding more computers or resources rather than increasing the power or storage capacity of a single machine. The main premise behind MapReduce is to split down an operation into segments and execute them in parallel to shorten the time required to perform the task.

The first stage of a MapReduce task is to convert input data to a range of core pairs as result. Large computational jobs are divided into smaller tasks using the "Map" function, which distributes them to the relevant key/value pairs (Cuzzocrea, et al., 2011). Unstructured data, such as text, can thus be mapped to a structured key/value pair, with the key being the word in the text and the value being the number of times the word appears. The "Reduce" function uses this output as an input. Reduce next collects and combines this output, merging any values that have the same key value to provide the final result of the computational task. Hadoop's MapReduce function relies on two separate nodes: the Job Tracker and Task Tracker nodes. The Job Tracker nodes are in charge of spreading the mapper and reducer functions to the Task Trackers that are available, as well as keeping track of the results . The Job- Tracker starts the MapReduce job by assigning a piece of an HDFS input file to a map task running on a node. The Scheduler modules, from the other side, are in charge of able to conduct the jobs and reporting back to the Job Monitor with the results. Because meetings are essential is generally done through directory in HDFS, inter-node interaction is decreased.

This section of the report focuses on the fundamental applications of cyber-twin studies, such as healthcare and industrial, as well as smart city development.

The healthcare industry is the most important application area for cyber-twin. Because the adoption of IoT devices is causing the healthcare business to grow at an accelerated pace, the establishment of cyber-twin enables for the rapid exchange of massive amounts of data collected from Connected devices, which switches and routers cannot do with the same precision and speed. Second, a computer hackers can be used to imitate pharmaceutical impacts on human health, as well as the formulation and management of medical therapy. Cyber-twin can also be used in AI-enabled medicine for decision- and forecast, especially when physicians are having trouble making medical treatment recommendations for any patient. The digital twin is capable of making decisions based on real-life scenarios.

The development of intelligent features can result in industrial processes that are highly automated and optimised. As a result, big data analytics and optimization will rely heavily on AI, leading to extremely effective industrial processes. Big data from various online and offline, as well as incoming and outbound processes, will be integrated and ingested by future IIoT systems. High-dimensional, multimillion variable datasets will be produced as a result of the integration of customer and company data. Such large datasets can be optimised and analysed using AI approaches (Wang et al., 2016).

Applications that use fog may produce real-time radio and network data that can give users a better analysis experience. This converted data (huge IoT data) not only responds quickly but also creates new opportunities. Fog computing's big data analytics make it possible for devices to make intelligent and imprudent decisions without involving a human.

The Blockchain uses big data to build a storage network and databases that can integrate large amounts of data to produce new data sets. This could create brand-new business models that are highly pertinent to significant data segments like apps for digital currencies and identities. Both technologies allow users to create extensive data collections without being constrained by a data center or offline storage. Big data uses data to be stored and processed as data and improves the stability and dependability of data collecting. To preserve their assets and satisfy their privacy requirements, some firms must take data security into account before using Blockchain (Saldamli et al., 2020).

SDN will significantly impact big data applications as a critical networking paradigm. Big data acquisition, transmission, storage, and processing can be made much easier by several beneficial aspects, such as the separation of the control and data planes, logically centralized control, a global perspective of the network, and the capacity to profile the network. For instance, big data processing typically takes place in cloud data centers. By dynamically distributing resources in data centers to various big data applications to meet their service level agreements (SLAs), SDN-based data centers can perform better than traditional data centers (Han et al., 2014).

The industrial industry is looking for a system that can deliver high connections while keeping track of items. Cyber-twin is assisting these applications in the industrial business. Innovative technologies have the inteligneсe systems to detect the functioning of devices in the manufacturing company in live environment and predict the console's lifespan and future growth, which is crucial for determining product quality and durability. By setting up a simulation platform, the cyber-twin may quickly test things.

Smart buildings are rapidly spreading as a result of the proliferation of IoT devices; individuals are becoming nearer to each other for the exchange of knowledge, and digital twins are expected to join all of these groups in this situation. The cyber-twin can help with smart city planning, as well as design intelligent systems, smart intelligent traffic, smart agriculture and sensible livestock, and so on, using AI approaches.

Challenges of Cyber-Twin Machinery

Computer hackers now clearly needs the usage of various technologies at the same time, offering numerous obstacles. As a result, the first step in addressing these concerns is to recognize them. This section will go through some of the challenges that digitalization has to deal with, and they must be taken into account when building the computer security system.

(1) The computer hackers design requires connectivity that is interoperable with future technologies like AI and IoT, which will work in concert with the information security to guarantee that the entire system runs smoothly. The cyber-twin will fail to achieve its aim without a compatible and well-defined infrastructure, resulting in a system malfunction.
(2) The data required for cyber-twin is the second issue that arises. The essential data should be of excellent quality with no gaps in between. The performance of the cyber-twin is harmed by interrupted and low-quality data. To guarantee that the computer security architecture (Upadhyay et al., 2021); runs well, some interpretation of the data should be carried out to make sure that the information is perfect in every way and contains no errors.
(3) Another major flaw in the cybertwin architecture is the protection of the confidential documents it obtains from the numerous IoT devices it connects. As a response, multiple authorization and security measures should be established at the IoT system to prevent hackers and so safeguard the cyber-twin information in order to cope with cyber-twin data protection.
(4) Faces major challenges, the next problem is that there is no common design for the growth of cyber-twins, and if the solution is designed without one, there will be no standardization throughout, which will cause information flow to be misled, likely to result in an inefficient system ineffectual of reaching its objectives.
(5) The third issue with digital manufacturing technology is that it demands more care, and owing to a lack of resources, users are seldom well trained nor experienced in managing cyber-twins.

REFERENCES

Adams, N. M. (2010). Perspectives on data mining. *International Journal of Market Research*, *52*(1), 11–19. doi:10.2501/S147078531020103X

Agalianos, K., Ponis, S. T., Aretoulaki, E., Plakas, G., & Efthymiou, O. (2020). Discrete event simulation and digital twins: Review and challenges for logistics. *Procedia Manufacturing*, *51*, 1636–1641. doi:10.1016/j.promfg.2020.10.228

Asur, S., & Huberman, B. A. (2010, August). *Predicting the future with social media. In 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology* (Vol. 1). IEEE.

Bakshi, K. (2012, March). *Considerations for big data: Architecture and approach. In 2012 IEEE aerospace conference*. IEEE.

Bevilacqua, M., Bottani, E., Ciarapica, F. E., Costantino, F., Di Donato, L., Ferraro, A., Mazzuto, G., Monteriù, A., Nardini, G., Ortenzi, M., Paroncini, M., Pirozzi, M., Prist, M., Quatrini, E., Tronci, M., & Vignali, G. (2020). Digital twin reference model development to prevent operators' risk in process plants. *Sustainability*, *12*(3), 1088. doi:10.3390u12031088

Cohen, J., Dolan, B., Dunlap, M., Hellerstein, J. M., & Welton, C. (2009). MAD skills: New analysis practices for big data. *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases*, *2*(2), 1481–1492. doi:10.14778/1687553.1687576

Cuzzocrea, A., Song, I. Y., & Davis, K. C. (2011, October). Analytics over large-scale multidimensional data: the big data revolution! In *Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP* (pp. 101-104). 10.1145/2064676.2064695

Czwick, C., & Anderl, R. (2020). Cyber-physical twins-definition, conception and benefit. *Procedia CIRP*, *90*, 584–588. doi:10.1016/j.procir.2020.01.070

Elgendy, N. (2013). *Big Data Analytics in Support of the Decision Making Process [MSc Thesis]*. German University in Cairo.

Elgendy, N., & Elragal, A. (2014, July). Big data analytics: a literature review paper. In Industrial conference on data mining (pp. 214-227). Springer. doi:10.1007/978-3-319-08976-8_16

Han, Y., Seo, S. S., Li, J., Hyun, J., Yoo, J. H., & Hong, J. W. K. (2014, September). Software defined networking-based traffic engineering for data center networks. In *The 16th Asia-Pacific Network Operations and Management Symposium* (pp. 1-6). IEEE. 10.1109/APNOMS.2014.6996601

He, Y., Lee, R., Huai, Y., Shao, Z., Jain, N., Zhang, X., & Xu, Z. (2011, April). RCFile: A fast and space-efficient data placement structure in MapReduce-based warehouse systems. In *2011 IEEE 27th International Conference on Data Engineering* (pp. 1199-1208). IEEE. 10.1109/ICDE.2011.5767933

Herodotou, H., Lim, H., Luo, G., Borisov, N., Dong, L., Cetin, F. B., & Babu, S. (2011, January). Starfish: A Self-tuning System for Big Data Analytics. In Cidr (Vol. 11, No. 2011, pp. 261-272). Academic Press.

Juneja, S., Gahlan, M., Dhiman, G., & Kautish, S. (2021). Futuristic cyber-twin architecture for 6G technology to support internet of everything. *Scientific Programming*, *2021*, 2021. doi:10.1155/2021/9101782

Lee, R., Luo, T., Huai, Y., Wang, F., He, Y., & Zhang, X. (2011, June). Ysmart: Yet another sql-to-mapreduce translator. In *2011 31st International Conference on Distributed Computing Systems* (pp. 25-36). IEEE.

Lim, K. Y. H., Zheng, P., & Chen, C. H. (2020). A state-of-the-art survey of Digital Twin: Techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, *31*(6), 1313–1337. doi:10.100710845-019-01512-w

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Mohamed, S., Ismail, O., & Hogan, O. (2012). *Data equity: Unlocking the value of big data*. Centre for Economics and Business Research.

Mouthami, K., Devi, K. N., & Bhaskaran, V. M. (2013, February). Sentiment analysis and classification based on textual reviews. In 2013 international conference on Information communication and embedded systems (ICICES) (pp. 271-276). IEEE. doi:10.1109/ICICES.2013.6508366

Pignaton de Freitas, E., Olszewska, J. I., Carbonera, J. L., Fiorini, S. R., Khamis, A., Ragavan, S. V., Barreto, M. E., Prestes, E., Habib, M. K., Redfield, S., Chibani, A., Goncalves, P., Bermejo-Alonso, J., Sanz, R., Tosello, E., Olivares-Alarcos, A., Konzen, A. A., Quintas, J., & Li, H. (2020). Ontological concepts for information sharing in cloud robotics. *Journal of Ambient Intelligence and Humanized Computing*, 1–12. doi:10.100712652-020-02150-4

Plattner, H., & Zeier, A. (2012). *In-memory data management: technology and applications*. Springer Science & Business Media. doi:10.1007/978-3-642-29575-1

Russom, P. (2011). *Big Data Analytics*. TDWI Best Practices Report, Fourth Quarter, 2011.

Saldamli, G., & Razavi, A. (2020, October). Surveillance Missions Deployment on the Edge by Combining Swarm Robotics and Blockchain. In *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)* (pp. 106-112). IEEE. 10.1109/MCNA50957.2020.9264274

Sánchez, D., Martín-Bautista, M. J., Blanco, I., & de la Torre, C. J. (2008, December). Text knowledge mining: an alternative to text data mining. In *2008 IEEE International Conference on Data Mining Workshops* (pp. 664-672). IEEE. 10.1109/ICDMW.2008.57

Unit, E. I. (2012). The deciding factor: Big data & decision making. *Capgemini Reports*, 1-24.

Upadhyay, H. K., Juneja, S., Maggu, S., Dhingra, G., & Juneja, A. (2021). Multi-criteria analysis of social isolation barriers amid COVID-19 using fuzzy AHP. *World Journal of Engineering*.

ur Rehmana, Yaqoobb, Salahc, Imrand, Jayaramane, & Pereraf. (2019). The role of big data analytics in industrial Internet of Things. *Future Generation Computer Systems, 99,* 247-259.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. doi:10.1016/j.ijpe.2014.12.031

Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination. *Computer Networks*, *101*, 158–168. doi:10.1016/j.comnet.2015.12.017

Yu, Q., Ren, J., Zhou, H., & Zhang, W. (2020, March). *A cybertwin based network architecture for 6G. In 2020 2nd 6G Wireless Summit (6G SUMMIT)*. IEEE.

Chapter 11
The APT Cyber Warriors With TTP Weapons to Battle:
An Review on IoT and Cyber Twin

Diana Arulkumar
Kalasalingam Academy of Research and Education, India

Kartheeban K.
Kalasalingam Academy of Research and Education, India

Arulkumaran G.
https://orcid.org/0000-0002-5166-3037
Bule Hora University, Ethiopia

ABSTRACT

Due to the blooming of Industrial 4.0 such as internet of things (IoT), cloud computing, industrial IoT (IIoT), and artificial intelligence (AI), with their innovative ideas and opportunities, the cyber attacker's modus operandi against the cyber defense triage is incredible. The genre of advanced persistent threat (APT) actors/group are equipped with sophisticated and substantial resources of tools, techniques, and procedure (TTP) at a breakneck pace. The IoT gadgets such as sensors, intelligent devices, and various rapidly emerging resources with energy, memory, and processing power are exponentially prone to multiple vulnerabilities. The nature of IIoT prompts heterogenous and rapid changes ranging the vulnerabilities from simple to complex attacks. APT menace follows the covert TTPs to target the asset of any organization like the government, military, or financial industry.

DOI: 10.4018/978-1-6684-5722-1.ch011

INTRODUCTION

In this pandemic covid season, the cyber warriors use disruptive cyber weapons as inexorable tide of the cybercrimes, data breaches, industrial espionage, and budding ruin of national infrastructure. The cyber incident reports leaves footprint impression and covering the tracks of new pernicious threats and drowning in tides of new risks. The challenges are due to sheer lack of knowledge about the new tools launched in the market and their consequence, the allocation of unsophisticated budget in mitigating the advanced cyber threats and unreliable attack vectors (one click, water holing, drive by downloads etc.). The data sources can be collected from external feeds, mining hacker discussion forums and academic researchers and also looping with security professional and their experiences.

In cyber era, millions of devices are connected through the internet, which is inevitability prone to vulnerable attacks by the perpetrators. The cyber actors whose are prone to espionage /sabotage different sectors, such as, industrial, military, economic, technical and intellectual property, financial extortion, and political manipulation. As there are numerous existences of cyber threats or zero-day threats, one among is Advanced Persistence Threat (APT). The APT are pernicious, highly – sophisticated, well-organized, with full spectrum proficiency of TTPs (Tactics, Techniques and Procedures or TTPs) and exploit target or IT networks of an organizations and cover their track and persist for long term endurance of access. WEBC2 backdoor family, that targets millions of computers to steal banking information and other credentials. The Three main operating systems are the sources which defines the structure of the cyber threats are Microsoft, Apple, and Linux. Many APT threat actors are capable of generating the variants of APT with Cyber threat intelligence (CTI). The CTI could be collected in 4 ways such as operational, tactical, technical and strategical.

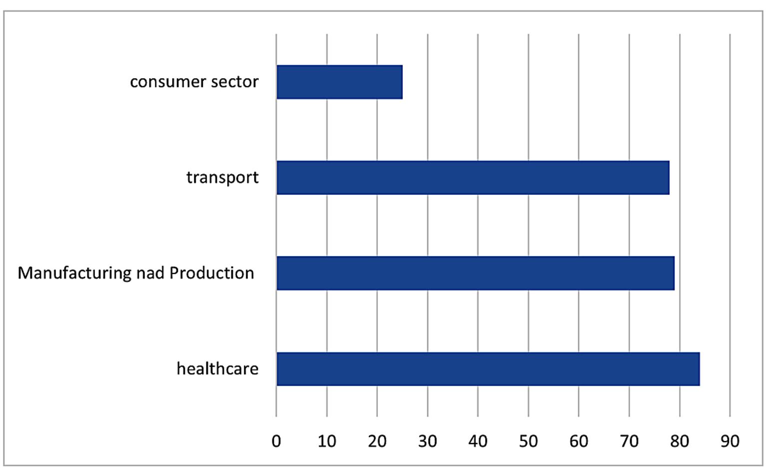
The Internet of Things (IoT)is the most ubiquitous from Consumer (smart homes),Machine to Machine(mobile fitness devices, smart cities, smart factories, and the smart grid) and Industry the Industrial Internet of Things (IIoT) (smart agriculture).Among that IIoT is adopted and enabled by the cheap cost of affordable sensors, actuators, processors and its availability facilitates the real time data access possibility and from that the data can be analysis to predict the future events.

BACKGROUND

In order to categorize the identity of attackers, in 2006 APT Phrase is framed by U.S. Airforce Analysts. The characteristics of an APT attackers are well skilled and persistent, equipped with sophisticated resources and targeted. The APT attackers launch an attack in multi stages. The APT is multi stage model. Quintero-Bonilla

.et.al,2020 says, that the APT life cycle model consists of three-stages with Initial compromise (IC), Lateral movement, command &control(C2C), intrusion kill chain (IKC) is a four-stage model Information Collection, Intrusion phase, Lateral expansion, Information theft phase, 4 Stages Initial Compromise, C&C, Lateral movement, Attack achievement. This model called attack chain which comprises five Stages such as Reconnaissance, Incursion, Discovery, Capture, Ex-filtration. 5 Stages Delivery, Exploit, Installation, C&C, Actions. Attackers once after run a malware and exploit the zero -day vulnerability, access the network through the compromised computer to achieve the default goals. This life cycle based on the intrusion kill chain model which consists of 6Stages like Reconnaissance, weaponization, Delivery, Initial intrusion, C&C, Lateral movement, Data ex-filtration. The Lockheed Martin company designed a life cycle called cyber kill chain CKC, to understand the attackers TTP, they proposed 7 Stages Research, Preparation, Intrusion, conquering network, hiding presence, gathering data, Maintaining access. (Formerly Mandiant) the FireEye, after done penetrated testing of the APT1 campaign, it concluded with 8 Stages Initial recons, Initial, compromise, establish foothold, Escalate, privileges, Internal recon, move laterally, Maintain presence, Complete mission. The ATT and CK Focuses on the tactics based on the cyber threat actor who wants to accomplishes strategic goal and it classifies into 11 Stages such as Initial access, Persistence, Privilege Escalation, Discovery, Lateral movement, Collection, Exfiltration Stages executed in parallel: Execution, Defence evasion, Credential access, and Command &Control.

Figure 1. A survey from 2018 to 2021 of cyber threats Challenges on IIoT



Attacker's Malware Concealment of Strategies'

Fig.2 states the malware generation techniques followed by the attackers to target the assets which will be either categorize based on its essentiality. The stack of pyramid explains the pain need to be taken to launch an attack to target the defender. The adversaries used TTP for different attack pattern, In order to analyze the Advanced Persistent Threat(APT),the below contents are the basic ideas about the cyber threat.

TTP Profiling

In order to profile any threat actor behavior, the term Tactics, Techniques, and Procedures (TTP) is proliferated for analyzing APT's complexity of life-cycle and predict the vulnerability in early stages. Maria Markstedte,2017 The tactics is defined that the process involved from initial stage till the end. It is simple for the adversary to use the already available tools for certain stages of the process instead being developing the new tools. To be proactive, it is necessary to understand what type of adversary deals with us and what target is focused throughout all the stages of tactics from initial to last stage of covering tracks of the campaign James Spiteri et.al, 2021, Markstedte,2 017 Fig.3 the multi stages of APT and its TTP application.

Figure 2. Overview of Malware Generation to target the assets

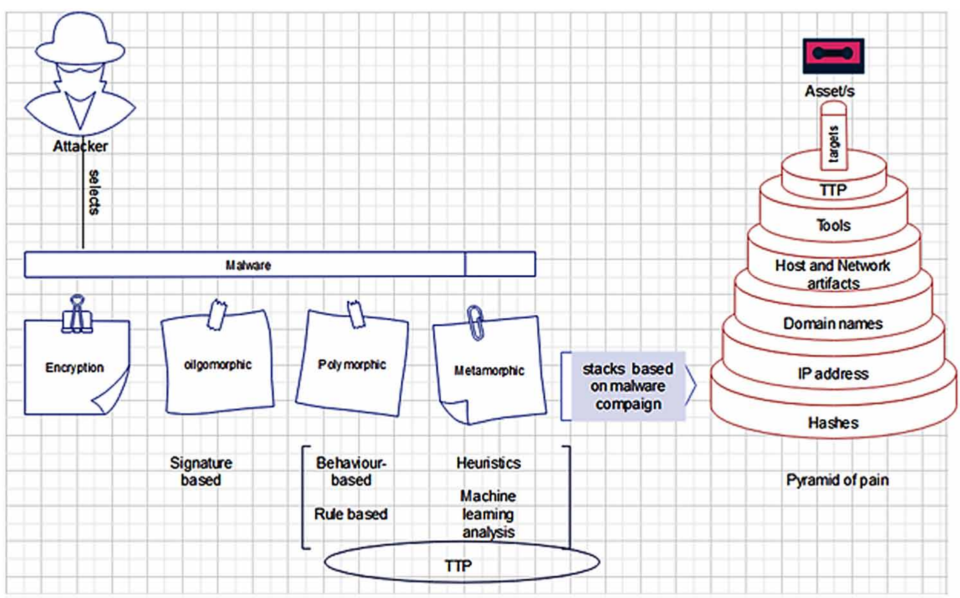
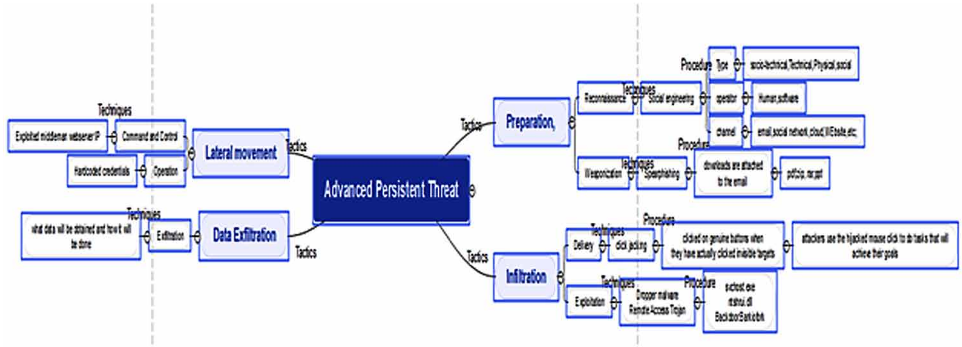


Figure 3. APT's Tactics Technique Procedure (TTP).



Tactics

The techniques can be states that techniques which are used in each process of launching a campaign. Procedures are the commands used by the threat actor to understand the adversary focusing target within the infrastructure

1. Analyze campaign's early stages.

From the past campaign which had launched by the Cyber threat actors and their cause of stealing the information are analyzed and the data related to the event are collected and stored in database which would be useful to predict current and future campaign from the attacker

2. Examine campaign's infrastructure and artifacts.

The sophisticated Threat actor can develop new tools, zero -day vulnerabilities and including high rated concealment of obsure. James Spiteri et.al, 2021 explains If an APT target the specific blind spot in the organization, after the last stage of exfiltrating the information as big chunk in either few rounds or at once can cover their track silent cleanup /wipeout completely

3. Some prefer re-compromising infrastructure, but not all.

As an indicator to be persistent for long period threat actor can deploy in last stages of campaign.

Technique

1. Understanding techniques in various /every stage of the attack

The various techniques are to facilitate the entire campaign process to move forward and hide the exfiltrating the informationAlshamrani,i et.al, 2019 states that. Based on the type of threat actors employ either they use existing or develop their own unique techniques and differ from other threat actors. In the early stages they use software tools to deploy their techniques for ex: social engineering, once they are successful to breach the target with initial information, they proceed with technological tools to gain control and escalate privileges.

2. Employing both technological and non-technological tools.

The adversary selects their own Protocol to obfuscate their exfiltrate data using encrypted and network technology over the network. Radwarel, 2019 states that finally, the either wipeout of covering the tracks or long-term persistence can be worked with set of software tools. The attribution may reply on social tricks with the intentional use of artifacts or public claim of fake. Since a wide range of techniques are available, the crucial part of investigation to forging the evidence of the digital worlds is misleading.

Command/Procedure

A step by step of set of actions which is used by the threat actors to execute the process of every stages APT life cycle. Based on APT actors and their purposes procedures will be varies. To orchestrates the success rate in every stage of APT lifecycle, Radwarel, 2019 states that Advanced threat actors use well -tailored procedures.

Cyber Threat Actors

The Threat actors are categorized into Nation-state or government-sponsored groups, Organized crime, Hacktivists, Insider threat, Opportunists, (Script kiddies or amateur criminals), Grey hat hackers or security researchers. By considering the following criteria the threat actors are called to be known with their profiles. In Table 1 defines the sophisticated APT threat actorsBlog, R, 2020 states that proliferate Zero-Day vulnerabilities, newly developed tools and apply heavy obfuscation while a less sophisticated actor relies on public exploits and open-source tools.Tollefson, 2021 states that

Table 1. The characteristics of Cyber Threat Actors

CTA	levels	Funding	Proficiency	Timeframe	Aggressiveness	Focus	Motivation	Goal
Nation-state or government-sponsored groups	Sophisticated	Well	Integrated-	*Enduring*	Sophisticated	Targeted attacks	economic, political or military proportion	Information for espionage
Organized crime	Sophisticated	Well	Advanced	Sustained or enduring	Significant	Targeted attacks	High Profits	personally identifiable information (PII) or to steal high-value digital resources
Hacktivists	Midlevel sophisticated	Self - funded	Developed	sustained	Moderate	High profile attacks to create headline	Political	draw attention to their cause
Insider threat	low to mid-level of sophistication	self-funded (low funding	Augmented	Sustained or Episodic	some expertise	expose or damage the company / steal data	revenge or money	go undetected
Opportunists (Script kiddies or amateur criminals)	low to mid-level of sophistication	self-funded (low funding	Acquired	*Episodic*	Self -Expertise	fame or money	headlines	cause disorder
Grey hat hackers or security researchers	mid-level of sophistication	funded by research company	---	--	--	money and prestige	industry kudos	find a zero-day

System Design and Implementation From ICS to IIOT

Industrial Control System

Staff, E states that Industrial control system (ICS) is a general term that encompasses several types of control systems, including supervisory control and data acquisition (SCADA) systems, distributed control systems (DCS), and other control system configurations such as Programmable Logic Controllers (PLC) often found in the industrial sectors and critical infrastructures. An ICS consists of combinations of control components (e.g., electrical, mechanical, hydraulic, pneumatic) that act together to achieve an industrial objective (e.g., manufacturing, transportation of matter or energy). The part of the system primarily concerned with producing the output is referred to as the process. The control part of the system includes the specification of the desired output or performance. Control can be fully automated or may include a human in the loop. Systems can be configured to operate open-loop, closed-loop, and manual mode. In open-loop control systems the output is controlled by established settings. In closed-loop control systems, the output has an effect on the input in such a way as to maintain the desired objective. In manual mode the system is controlled completely by humans. The part of the system primarily concerned with maintaining conformance with specifications is referred to as the controller (or control). A typical ICS may contain numerous control loops, Human Machine Interfaces (HMIs), and remote diagnostics and maintenance tools built using an array of network protocols. ICS control industrial processes are typically used in electrical, water and wastewater, oil and natural gas, chemical, transportation, pharmaceutical, pulp and paper, food and beverage, and discrete manufacturing (e.g., automotive, aerospace, and durable goods) industries.

In this era, Trend micro and Wikipedia states the integration of OT and IT as Human -Computer Interface (HMI) is used to control the ICS (Industrial control system) in various critical infrastructures like water treatment plant, power station, pharmaceutical, drugs production and many more. In this pandemic covid season, used one of prominent attack model by the adversaries by implementing the Advanced Persistent threat (APT) Tactics and Techniques Procedure (TTP) APT cyber kill chain of MITRE(ATT&CK). The legacy ICS systems are add-on with IT are fixing their targets as cyber threat actors who use the TTP are remain passive and undetectable even by the antivirus the threat actors. Since the cost of Protecting the assets is very expensive many organizations are not spending much interest and budget for it. Once after the execution of the payload, (i.e the damages in hardware or exfiltration of data) sometimes it will be identified after some days, months or even years.

Basically the four layer Such as IOT device consist of System like smart city, smart hospital called Physical Layer, Networking connectivity may be wired or wireless like WIFI and Ethernet LAN combinedly called as Network Layer, Data Processing, Management and storage such as cloud computing employing is called processing/ Infrastructure layer and last layer is the application layer where controlling, monitoring, analytical services, etc. The functionality of the architecture is designed for Data collection and control and optimization.

1. Physical Layer: Sensors, actuators and Microcontrollers like esp32,esp8266
2. Networking Layer: Can be used wired or wireless
3. Processing/Infrastructure Layer: data processing, management and storage
4. Application Layer: Controlling, monitoring, safety etc.

The Fig.4 Portrait's the working model of ICS Architecture and its functional components, and also the challenges of attacks for their respective layer. The forward flow from application layer to physical layer can be used for Control and optimization. The reverse of it is used for Data collection.

From the Fig 5. Islamet.al portraits that The assets of the world are Physical, Digital and Cyber which are prone to attack with their associative vulnerabilities. The assets are tended to targeted by the adversaries for various reasons such as political, financial, fame, eagerness etc.in table 2 layers are described and mapped with its associated attacks.

Figure 4. A general I-IoT Frame work of developed (SCADA) the attack types

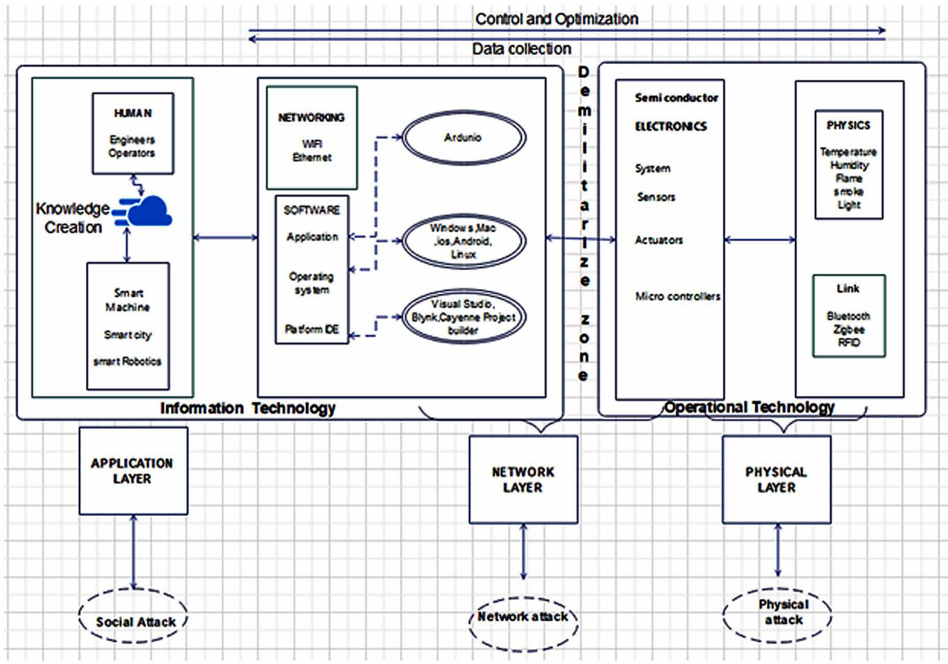


Figure 5. Interconnectivity of attacks with related to the Assets

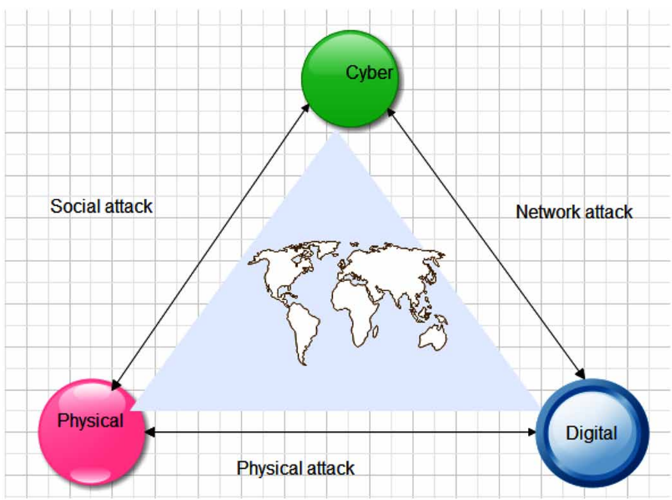


Table 2. IIoT Layer and its associated attacks

	Attacks
1. Application Layer	Attacks at this level include a set of malicious actions like interception and deception, but also more advanced types such as adversarial attacks.
a. Human	DoS, MItM
b. Cloud	Cloud Malware Injection, Authentication attacks
2. Infrastructure Layer	Attacks targeted at this level exploit known or unknown vulnerabilities of these services and enter malicious code where the application expects legitimate data from the user in order to gain access with administrator privileges.
a. Software	
i. Application software	Phishing, DNS poisoning, Remote code execution, Web application attacks
ii. Operating system	
iii. Platform IDE	
3. Network Layer	IP packet creation techniques with false attributes such as the source address, in order to disguise the identity of the sender of the packet, encouraging the recipient to think that it came from a legitimate network user.
a. Wired	Passive or Active Eavesdropping Man-in-the-Middle (MitM) packet sniffing, data sniffing,Malware,
b. wireless	
4. Physical Layer	Attacks aimed at this level require an excellent knowledge of the design of the IIoT system, and access to the specifications of active devices, engineering plans, and detailed information about their installation and operational functionality Attacks at this level aim at preventing legitimate communication between the two levels and controlling the flow of communication
a. IoT devices	Reverse engineering, Malware Eaves dropping, Brute-force attacks
b. Micro controllers	

Hypothesis of APT Attacker

From Grooby, Set.al,2019 portraits that, The following 3 hypothesis of an attack pattern are considered for an APT adversary group attacker.

target(A_{adv}, B_{tact}, C_{asset})

Table 3. Different Attack pattern

α_i	β_i	δ_i	P_{tn}
1	1	1	P1
few	same	few	P2
n	n	all	P3

Hypothesis one: when an adversary target AND an asset AND with one tactics

$\alpha_i \in A_{adv}$, i=1,2,…m; $\beta_j \in B_{tact}$, j =1,2,…,n; $\delta_i \in C_{asset}$, i=1,2,…1

$$\tau\left(\alpha_i, \beta_j, \delta_i\right) = \left[\left[\alpha_i \cap \beta_j \cap \delta_i\right] \ldots\ldots\ldots\ldots \left[\alpha_m \cap \beta_j \cap \delta_i\right]\right] \quad (1)$$

A_{adv}, B_{tact}, C_{asset} are the profiles of adversary, tactics repository, asset Inventory
If: Sim(ai,bj)≥ϕ, and Sim(bj,ai)≥ϕ; ϕ is the threshold, and 0<ϕ

Hypothesis two: when few adversary targets AND select a number of assets AND with same tactics α_i, α_t, β_j, δ

$\alpha_i \in A_{adv}$, i=1,2,…m; $\alpha_t \in A_{adv}$, t=1,2,…m; i≠t;

$\beta_j \in B_{tact}$, j =1,2,…,n $\delta_i \in C_{asset}$, i=1,2,...1

$$\tau\left(\alpha_i, \alpha_t, \beta_{j,}\right) = \left[\begin{bmatrix} \alpha_i \cap \beta \cap \delta_i & \alpha_i \cap \beta \cap \delta_i \\ \alpha_t \cap \beta \cap \delta_i & \alpha_t \cap \beta \cap \delta_i \end{bmatrix} \ldots\ldots\ldots\ldots \begin{bmatrix} \alpha_m \cap \beta \cap \delta_i & \alpha_m \cap \beta \cap \delta_i \\ \alpha_m \cap \beta \cap \delta_i & \alpha_m \cap \beta n \cap \delta_i \end{bmatrix}\right] \quad (2)$$

A and B are two data sets.
If: Sim(α_i β_j)≥ϕ and Sim(α_t, β_j)≥ϕ; ϕ is the threshold, and 0<ϕ

Hypothesis three: when the n-adversary targets AND all of the assets AND with n-tactics. There are entities α_i, α_t, β_j, β_k, β_r, δ_i;

$\alpha_i \in A_{adv}$, i=1,2,…m; $\alpha_t \in A_{adv}$, t=1,2,…m; i≠t;

$\beta_j \in B_{tact}$, j =1,2,…,n; $\beta_k \in B_{tact}$, k =1,2,…,n;

$\beta_r \in B_{tact}$, r =1,2,…,n*;* $j \neq k \neq r$; $\delta_a \in C_{asset}$, i=1,2,…p $\delta_b \in C_{asset}$, i=1,2,…0 $\delta_c \in C_{asset}$, i=1,2,…q

$$= \left[\begin{bmatrix} \alpha_i \cap \beta_j \cap \delta_a & \alpha_i \cap \beta_k \cap \delta_a & \alpha_i \cap \beta_r \cap \delta_a \\ \alpha_t \cap \beta_{j,} \cap \delta_b & \alpha_t \cap \beta_k \cap \delta_b & \alpha_t \cap \beta_r \cap \delta_b \end{bmatrix} \ldots\ldots\ldots\ldots\ldots \right] \quad (3)$$

A and B are two data sets.

If: $\text{Sim}(\alpha_i, \beta_j) \geq \phi$, and $\text{Sim}(\alpha_i, \beta_k) \geq \phi$, and $\text{Sim}(\alpha_i, \beta_k) \geq \phi$, and $\text{Sim}(\alpha_i, \beta_r) \geq \phi$; ϕ is the threshold and $0 < \phi$.

Fig.6 explains about the Multi stages of tactics and techniques and Procedure in an hierarchical order based on goals and subgoals to target the assets of ICS and IIoT by adopted of different taxonomy like Lockheed's Cyber kill chain phases.The mind map is used to APT's TTP to target ICS/IIoT layers.

Security Layer of IIoT

To monitor the several entity in industrial process, the Supervisory Control and Data Acquisition (SCADA) system is the core component of I-IoTGrooby, Set.al,2019 portraits that. A few APT campaigns of I-IoT are Stuxnet, Shamoon malware, Sauron, Copy Kittens, Volatile Cedar and ShellCrew .Since the evolution renders the variety of IoT devices in the industry based on the novel technologies. Javed, S et.al,2019 portraits that the spectacular focus industries is about the security and Vulnerability of IoT. To provide a scope, Fig.7 DL/ML can be applied as a solution. Hussain et.al, 2021, Alotaibi, et.al 2020,Radware, 2019 states that proposed various publicly available datasets for Industrial control system and IoT such as ISCX2012, CTU-13,POT, N_BaIoT. Hussain, Z et.al,2021 portraits that Alotaibi, Bl,2020 portraits that and Power gridAlotaibi, Bl,2020 portraits that are estimated the accuracy with deep learning algorithm mainly CNN and LSTM, ResNet

CONCLUSION

APT is an catastrophic disaster with its highly sophisticated to target an enterprise networks IIoT As the topics are discussed related to the digital twin and I- IoT analytics are carried out with various algorithm and application like ardunio, platform. io, visio, etc. In future furthermore, the I-IoT for Blockchain security using Deep Learning could be adopted in industrial application for security and functionality.

Figure 6. Mitre's ATT&CK ICS TTP for APT's Cyber kill chain phases

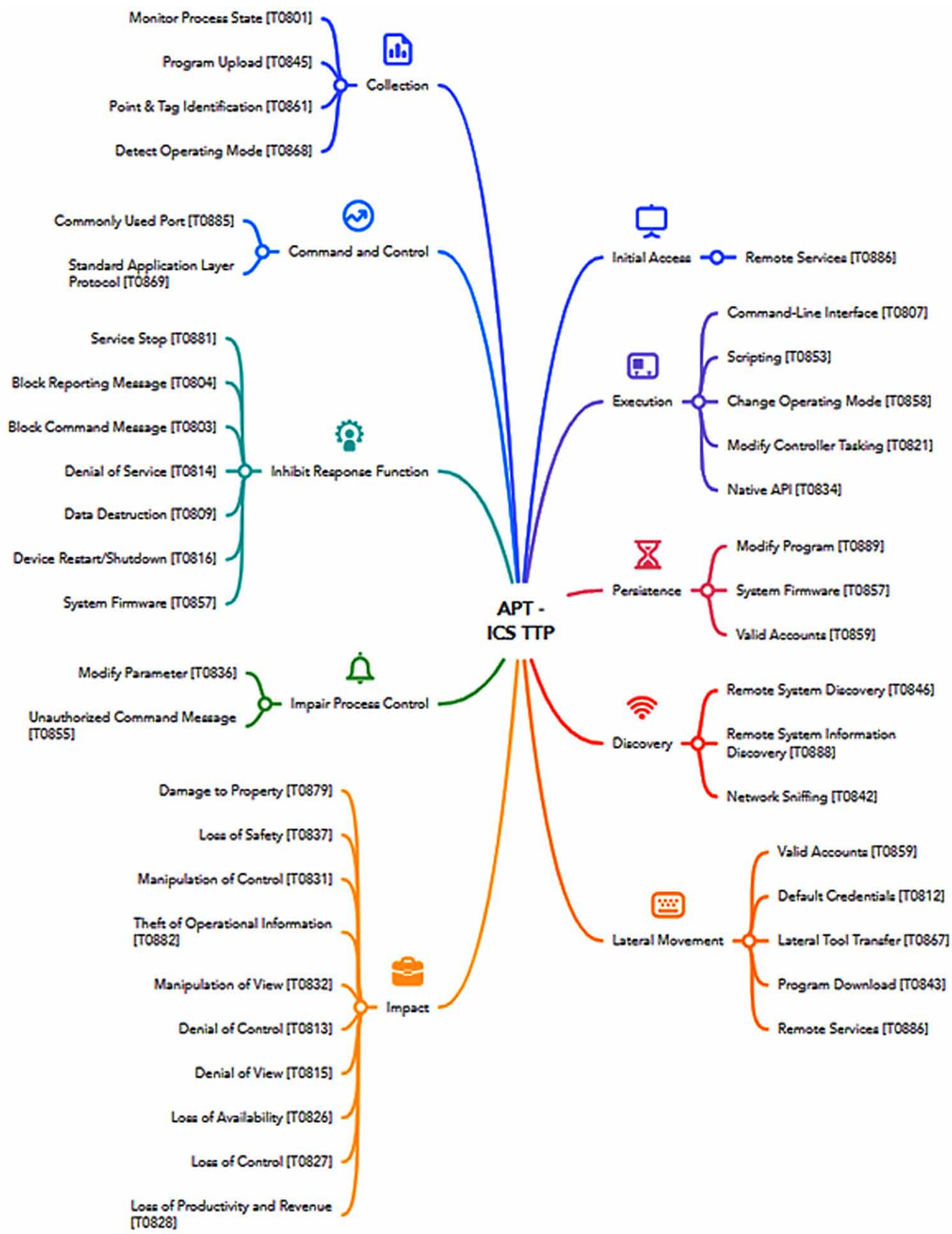
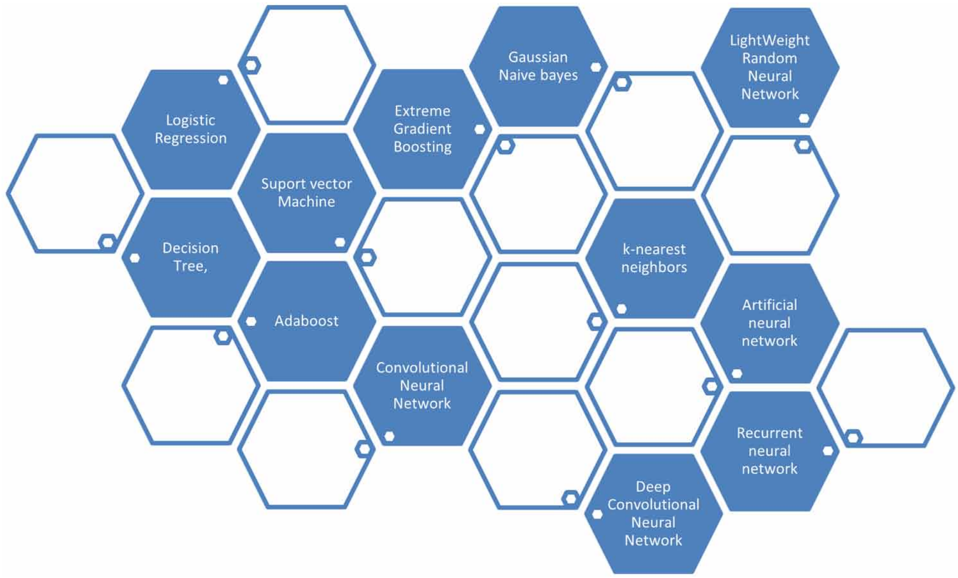


Figure 7. Data analytics can be applied with Deep learning and machine learning algorithms



REFERENCES

Alotaibi, B., & Alotaibi, M. (2020). A stacked deep learning approach for IoT cyber attack detection. Journal of Sensors.

Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. IEEE Communicatio*ns Surveys and Tutorials, 21(2), 1851–187*7. *d*oi:10.1109/COMST.2019.2891891

Blog, R. (2020, April 2). 7 Types of Cyb*er Threat Actors And Their Damage. https://www.*redlegg.com/blog/cyber-threat-actor-types

da Rocha, B. C., de Melo, L. P., & de Sousa, R. T., Jr. (2021). A Study on APT in IoT Networks. In ICE-B (pp. 160-164). Academic Press.

Grooby, S., Dargahi, T., & Dehghantanha, A. (2019). Protecting IoT and ICS platforms against advanced persistent threat actors: analysis of APT1, silent chollima and molerats. In Handbook *of big data and IoT security (pp. 22*5–255). Springer. doi:10.1007/978-3-030-10543-3_10

Hussain, Z., Akhunzada, A., Iqbal, J., Bibi, I., & Gani, A. (2021). Secure IIoT-Enabled Industry 4.0. Susta*inability, 13(*22*),* 12384. doi:10.3390u132212384

Islam, M. R., & Aktheruzzaman, K. M. (2020). An analysis of cybersecurity attacks against internet of things and security solutions. Jo*urnal of Computer and Communications,* 8(*4*), 11–25. doi:10.4236/jcc.2020.84002

Javed, S. H., Ahmad, M. B., Asif, M., Almotiri, S. H., Masood, K., & Ghamdi, M. A. A. (2022). An Intelligent System to Detect Advanced Persistent Threats in Industrial Internet of Things (I-IoT). *Electronics (Basel)*, *11*(5), 742. doi:10.3390/electronics11050742

Karia. (2021, August 31). *Advanced Persistent Threat (APT) Groups*. Cyber Sophia. https://cybersophia.net/cyber-threat-intel/advanced-persistent-threat-apt-groups/

Markstedte. (2017, May 20). *Tactics, Techniques, and Procedures (TTPs).* Azeria-Labs. https://azeria-labs.com/tactics-techniques-and-procedures-ttps/

Pedreira, V., Barros, D., & Pinto, P. (2021). A review of attacks, vulnerabilities, and defenses in industry 4.0 with new challenges on data sovereignty ahead. *Sensors (Basel)*, *21*(15), 5189. doi:10.339021155189 PMID:34372425

Quintero-Bonilla, S., & Martín del Rey, A. (2020). A new proposal on the advanced persistent threat: A survey. *Applied Sciences (Basel, Switzerland)*, *10*(11), 3874. doi:10.3390/app10113874

Radware. (2019, December 8). *Tactics, Techniques and Procedures*. www.radware.com

Spiteri. (2021, June 14). *Adversary emulation with Prelude Operator and Elastic Security*. https://www.elastic.co/blog/adversary-emulation-with-prelude-operator-and-elastic-security

Staff, E. (2020, August 4). *Overview of Industrial Control Systems*. Inst Tools. https://instrumentationtools.com/overview-of-industrial-control-systems/

Tollefson, R. (2021, March 26). *Security+: How to explain threat actor types and attributes* [updated 2021]. Infosec Resources. https://resources.infosecinstitute.com/certification /security-how-to-explain-threat-actor-types-and-attributes/

Compilation of References

Abeshu, A., & Chilamkurti, N. (2018). Deep learning: The frontier for distributed attack detection in fog-to-things computing. *IEEE Communications Magazine*, *56*(2), 169–175. doi:10.1109/MCOM.2018.1700332

Adams, N. M. (2010). Perspectives on data mining. *International Journal of Market Research*, *52*(1), 11–19. doi:10.2501/S147078531020103X

Adamson, G., Wang, L., & Moore, P. (2017). Feature-based control and information framework for adaptive and distributed manufacturing in cyber physical systems. *Journal of Manufacturing Systems*, *43*, 305–315. doi:10.1016/j.jmsy.2016.12.003

Adhikari, A., & Adhikari, J. (2015). Advances in Knowledge Discovery in Databases. Springer Publications.

Adhikari, M., Munusamy, A., Kumar, N., & Srirama, S. N. (2021). Cybertwin-driven resource provisioning for IoE applications at 6G-enabled edge networks. *IEEE Transactions on Industrial Informatics*, *18*(7), 4850–4858. doi:10.1109/TII.2021.3096672

Aditya Kusupati, M. S. (2019). *Fastgrnn: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network.* arXiv:1901.02358

Agalianos, K., Ponis, S. T., Aretoulaki, E., Plakas, G., & Efthymiou, O. (2020). Discrete event simulation and digital twins: Review and challenges for logistics. *Procedia Manufacturing*, *51*, 1636–1641. doi:10.1016/j.promfg.2020.10.228

Agerri, R., Artola, X., Beloki, Z., Rigau, G., & Soroa, A. (2015). Big data for Natural Language Processing: A streaming approach. *Knowledge-Based Systems*, *79*, 36–42. doi:10.1016/j.knosys.2014.11.007

Alam, K. M., & El Saddik, A. (2017a). C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access: Practical Innovations, Open Solutions*, *5*, 2050–2062. doi:10.1109/ACCESS.2017.2657006

Alkan, B., & Harrison, R. (2019). A virtual engineering-based approach to verify structural complexity of component-based automation systems in early design phase. *Journal of Manufacturing Systems*, *53*, 18–31. doi:10.1016/j.jmsy.2019.09.001

Alkinani, M. H., Almazroi, A. A., Adhikari, M., & Menon, V. G. (2022). Artificial Intelligence-Empowered Logistic Traffic Management System Using Empirical Intelligent XGBoost Technique in Vehicular Edge Networks. *IEEE Transactions on Intelligent Transportation Systems*, 1–10. doi:10.1109/TITS.2022.3145403

Alotaibi, B., & Alotaibi, M. (2020). A stacked deep learning approach for IoT cyber attack detection. *Journal of Sensors*.

Alrawais, A., Alhothaily, A., Hu, C., & Cheng, X. (2017). Fog computing for the Internet of Things: Security and privacy issues. *IEEE Internet Computing*, *21*(2), 34–42. doi:10.1109/MIC.2017.37

Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys and Tutorials*, *21*(2), 1851–1877. doi:10.1109/COMST.2019.2891891

Amin, S. U., Agarwal, K., & Beg, R. (2013, April). Genetic neural network based data mining in prediction of heart disease using risk factors. *IEEE Conference on Information & Communication Technologies*, 1227-1231.

Amor, N. B., Benferhat, S., & Elouedi, Z. (2003, July). Naive bayesian networks in intrusion detection systems. *Proc. Workshop on Probabilistic Graphical Models for Classification, 14th European Conference on Machine Learning (ECML) and the 7th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD).*

Anbarasi, M., Anupriya, E., & Iyengar, N. C. S. N. (2010). Enhanced prediction of heart disease with feature subset selection using genetic algorithm. *International Journal of Engineering Science and Technology*, *2*(10), 5370–5376.

Andreev, S., Balandin, S., & Koucheryavy, Y. (2022). Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer International Publishing.

Anila, M., & Pradeepini, G. (2018). Least square regression for prediction problems in machine learning using R. *International Journal of Engineering and Technology, 7*(12), 960-962. . doi:10.14419/ijet.v7i3.12305

Anooj, P. K. (2012). Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University-Computer and Information Sciences.*, *24*(1), 27–40.

Ansari, X. S. (2016). Edge IoT: Mobile edge computing for the Internet of Things. *IEEE Communications Magazine*, *54*(12), 22–29. doi:10.1109/MCOM.2016.1600492CM

Anscombe, F. J. (1973). Graphs in statistic analysis. *The American Statistician*.

Aparna, P., & Kishore, P. V. V. (2018). An efficient medical image watermarking technique in E-healthcare application using hybridization of compression and cryptography algorithm. *Journal of Intelligent Systems*, *27*(1), 115–133. doi:10.1515/jisys-2017-0266

Apple. (2017). *Deep Learning for Siri's Voice: On-Device Deep Mixture Density Networks for Hybrid Unit Selection Synthesis*. Retrieved from https://machinelearning.apple.com/2017/08/06/sirivoices.html

Ashton, K. (2009). That "Internet of things" thing. *RFiD J.*, 97–114.

Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2015, June). Big data in healthcare: challenges and opportunities. In *2015 International Conference on Cloud Technologies and Applications (CloudTech)* (pp. 1-7). IEEE.

Assistant, A. (2017). *Hey Siri: An On-Device DNN-Powered Voice Trigger for Apple's Personal*. Retrieved from https://machinelearning.apple.com/2017/10/01/heysiri.html

Asur, S., & Huberman, B. A. (2010, August). *Predicting the future with social media. In 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology* (Vol. 1). IEEE.

Atlam, H. F., Walters, R. J., & Wills, G. B. (2018). Fog computing and the internet of things: A review. *Big Data and Cognitive Computing, 2*(2), 10.

Aujla, N. K. (2018). Optimal decision making for big data processing at edge-cloud environment: An SDN perspective. *IEEE Trans. Ind. Informat., 14*(2), 778–789.

Azim, H. S. A. T. (2019). Knowledge Based Recommender System for Academia Using Machine Learning: A Case Study on Higher Education Landscape of Pakistan. *IEEE Access: Practical Innovations, Open Solutions*.

Babu, B. S., Prasanna, P. L., & Vidyullatha, P. (2018). Customer data clustering using density based algorithm. *IACSIT International Journal of Engineering and Technology*, *7*(2), 35–38. doi:10.14419/ijet.v7i2.32.13520

Bakshi, K. (2012, March). *Considerations for big data: Architecture and approach. In 2012 IEEE aerospace conference*. IEEE.

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data. John Wiley and Sons*.

Barone, D., Yu, E., Won, J., Jiang, L., & Mylopoulos, J. (2010, November). Enterprise modeling for business intelligence. In *IFIP Working Conference on the Practice of Enterprise Modeling* (pp. 31-45). Springer. 10.1007/978-3-642-16782-9_3

BBC News. (2020). *New Zealand stock exchange halted by cyber-attack*. https://www.bbc.com/news/53918580

Becvar, P. M. (2017). Mobile edge computing: A survey on architecture and computation offlfloading. *IEEE Commun. Surveys Tuts.*, *19*(3), 1628–1656. doi:10.1109/COMST.2017.2682318

Berry, M. A., & Linoff, G. S. (2000). Mastering data mining: The art and science of customer relationship management. *Industrial Management & Data Systems*, *100*(5), 245–246. doi:10.1108/imds.2000.100.5.245.2

Bevilacqua, M., Bottani, E., Ciarapica, F. E., Costantino, F., Di Donato, L., Ferraro, A., Mazzuto, G., Monteriù, A., Nardini, G., Ortenzi, M., Paroncini, M., Pirozzi, M., Prist, M., Quatrini, E., Tronci, M., & Vignali, G. (2020). Digital twin reference model development to prevent operators' risk in process plants. *Sustainability*, *12*(3), 1088. doi:10.3390u12031088

Bhatla, N., & Jyoti, K. (2012). A Novel Approach for heart disease diagnosis using Data Mining and Fuzzy logic. *International Journal of Computers and Applications*, *54*(17).

Biron, J. F. (2016). *Foundational Elements of an IoT Solution*. O'Reilly Media Inc.

Blog, R. (2020, April 2). *7 Types of Cyber Threat Actors And Their Damage*. https://www.redlegg.com/blog/cyber-threat-actor-types

Booker. (2009). A Student Program Recommendation System Prototype. *Issues of Information Technology*.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent, *Proc. COMPSTAT*, 177–186. 10.1007/978-3-7908-2604-3_16

Box, G. E., Hunter, W. H., & Hunter, S. (1978). *Statistics for experimenters* (Vol. 664). John Wiley and Sons.

Branch, L. I. (2017). Pilot study on application of MEC local shunting service. *Modern Inf. Technol.*, *1*(3), 65–67.

Chadha, R., &Mayank, S. (2016). Prediction of heart disease using data mining techniques. *CSI Transactions on ICT, 4*(2-4), 193-198.

Chakravarti, I. M., Laha, R. G., & Roy, J. (1967). Handbook of methods of applied statistics. Wiley Series in Probability and Mathematical Statistics (USA) eng.

Chauvin, Y., & Rumelhart, D. E. (2013). *Backpropagation: Theory, Architectures, and Applications*. Psychology Press. doi:10.4324/9780203763247

Chen, . (2019). NeuroPilot: A cross-platform framework for edge-AI. *2019 IEEE International Conference on Artiicial Intelligence Circuits and Systems (AICAS)*, 167-170.

Chen, L., Xu, G., Zhang, S., Yan, W., & Wu, Q. (2020). Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks. *Journal of Manufacturing Systems*, *54*, 1–11. doi:10.1016/j.jmsy.2019.11.008

Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. (2017). Disease prediction by machine learning over big data from healthcare communities. *IEEE Access : Practical Innovations, Open Solutions*, *5*, 8869–8879.

Chen, W. D. (2012). Integration of workflflow partitioning and resource provisioning. *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID 2012)*, 764–768. 10.1109/CCGrid.2012.57

Chen, W.-J., Chang, C.-C., & Le, T. H. N. (2010). High payload steganography mechanism using hybrid edge detector. *Expert Systems with Applications*, *37*(4), 3292–3301. doi:10.1016/j.eswa.2009.09.050

Chien-Chun Hung, G. A. (2018). Videoedge: Processing camera streams using hierarchical clusters. *IEEE/ACM Symposium on Edge Computing (SEC)*, 115-131. 10.1109/SEC.2018.00016

Chiou, S. F., Liao, I.-E., & Hwang, M.-S. (2011, February). A capacity-enhanced reversible data hiding scheme based on SMVQ. *Imaging Science Journal*, *59*(1), 17–24. doi:10.1179/136821910X12750339175943

Chiou, S. F., Lu, Y. C., Liao, I.-E., & Hwang, M.-S. (2013). An efficient reversible data hiding scheme based on SMVQ. *Imaging Science Journal*, *61*(6), 467–474. doi:10.1179/1743131X12Y.0000000035

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, *27*(10), 1868–1883. doi:10.1111/poms.12838

Cisco. (2020). *Cisco Cybersecurity Reports*. https://www.cisco.com/c/en/us/products/security/cybersecurity-reports.html

Claire, L., & Raj, J. A. (2019). A Review on Student Placement Chance Prediction. In *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE.

Clark Study School. (2017). *Hackers Attack Every 39 Seconds*. https://www.securitymagazine.com/articles/87787-hackers-attack-every-39-seconds

Cohen, J., Dolan, B., Dunlap, M., Hellerstein, J. M., & Welton, C. (2009). MAD skills: New analysis practices for big data. *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases*, *2*(2), 1481–1492. doi:10.14778/1687553.1687576

Collobert, J. W. (2011). Natural language processing (almost) from scratch. *Learn. Res.*, *12*, 2493–2537.

Cox, P. (2011). *Mobile cloud computing devices, trends, issues, and the enabling technologies, developerWorks*. IBM.

Cuzzocrea, A., Song, I. Y., & Davis, K. C. (2011, October). Analytics over large-scale multidimensional data: the big data revolution! In *Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP* (pp. 101-104). 10.1145/2064676.2064695

Czwick, C., & Anderl, R. (2020). Cyber-physical twins-definition, conception and benefit. *Procedia CIRP*, *90*, 584–588. doi:10.1016/j.procir.2020.01.070

da Costa, K. A., Papa, J. P., Lisboa, C. O., Munoz, R., & de Albuquerque, V. H. C. (2019). Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks*, *151*, 147–157. doi:10.1016/j.comnet.2019.01.023

da Rocha, B. C., de Melo, L. P., & de Sousa, R. T., Jr. (2021). A Study on APT in IoT Networks. In ICE-B (pp. 160-164). Academic Press.

Dash, S. P., Joshi, S., Satapathy, S. C., Shandilya, S. K., & Panda, G. (2022). A cybertwin-based 6G cooperative IoE communication network: Secrecy outage analysis. *IEEE Transactions on Industrial Informatics*, *18*(7), 4922–4932. doi:10.1109/TII.2021.3140125

de Boer, M. H., Bakker, B. J., Boertjes, E., Wilmer, M., Raaijmakers, S., & van der Kleij, R. (2019). Text mining in cybersecurity: Exploring threats and opportunities. *Multimodal Technologies and Interaction*, *3*(3), 62. doi:10.3390/mti3030062

Deren, L., Wenbo, Y., & Zhenfeng, S. (2021). Smart city based on digital twins. *Computers & Urban Society*, *1*, 4.

Dhiman, G., Nagar, A., Vimal, S., & Rho, S. (2022). Guest Editorial: Cybertwin-Driven 6G for Internet of Everything: Architectures, Challenges, and Industrial Applications. *IEEE Transactions on Industrial Informatics*, *18*(7), 4846–4849. doi:10.1109/TII.2022.3151914

Ding, C. T., J. N. (2019). Edge computing: Applications, state-of-the-art and challenges. *Zte Technol.*, *25*(3), 2–7.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). John Wiley and Sons.

El-Bialy, R., Salamay, M. A., Karam, O. H., & Khalifa, M. E. (2015). Feature analysis of coronary artery heart disease data sets. *Procedia Computer Science*, *65*, 459–468.

Elgendy, N., & Elragal, A. (2014, July). Big data analytics: a literature review paper. In Industrial conference on data mining (pp. 214-227). Springer. doi:10.1007/978-3-319-08976-8_16

Elgendy, N. (2013). *Big Data Analytics in Support of the Decision Making Process [MSc Thesis]*. German University in Cairo.

Escamilla-Ambrosio, A. R.-M.-A.-B.-R. (2017). Distributing Computing in the Internet of Things: Cloud, Fog and Edge Computing Overview. *NEO*, 87-115.

Fraley, J. B., & Cannady, J. (2017, March). The promise of machine learning in cybersecurity. In *SoutheastCon 2017* (pp. 1–6). IEEE. doi:10.1109/SECON.2017.7925283

G̈unther, J., Pilarski, P. M., Helfrich, G., Shen, H., & Diepold, K. (2016). Intelligent laser welding through representation, prediction, and control learning: An architecture with deep neural networks and reinforcement learning. *Mechatronics, 34*, 1–11.

Garber, L. (2000). Denial-of-service attacks rip the Internet. *Computer*, *33*(04), 12–17. doi:10.1109/MC.2000.839316

Garcia-Lopez, P. M. (2015). Edge-centric computing: vision and challenges. *ACM SIGCOMM Comput. Commun. Rev.*, 37–42.

Geetha, R., & Thilagam, T. (2021). A review on the effectiveness of machine learning and deep learning algorithms for cyber security. *Archives of Computational Methods in Engineering*, *28*(4), 2861–2879. doi:10.100711831-020-09478-2

Ghafarian, T. J., & Javadi, B. (2015). Cloud-aware data intensive workflflow scheduling on volunteer computing systems. *Future Generation Computer Systems*, *51*, 87–97. doi:10.1016/j.future.2014.11.007

Ghosh, Shaw, Islam, & Piuri. (2022). *AI and IoT for Smart City Applications.* Springer.

Giorgini, P., Rizzi, S., & Garzetti, M. (2005, November). Goal-oriented requirement analysis for data warehouse design. In *Proceedings of the 8th ACM international workshop on Data warehousing and OLAP* (pp. 47-56). 10.1145/1097002.1097011

Glaessgen, E., & Stargel, D. (2012). The digital twin paradigm for future nasa and us air force vehicles. *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, 1818.

Goodfellow, I. (2014). Generative adversarial nets. *in Proc. Advances in Neural Information Processing Systems*, 2672–2680.

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper, 1*(2014), 1–7.

Grooby, S., Dargahi, T., & Dehghantanha, A. (2019). Protecting IoT and ICS platforms against advanced persistent threat actors: analysis of APT1, silent chollima and molerats. In *Handbook of big data and IoT security* (pp. 225–255). Springer. doi:10.1007/978-3-030-10543-3_10

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, *21*(1), 27–58. doi:10.1214/aoms/1177729885

Gubbi, J. B., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660. doi:10.1016/j.future.2013.01.010

Guillaume Lample, M. B. (2016). *Neural architectures for named entity recognition.* arXiv:1603.01360.

Gulati, R. (2003). *The threat of social engineering and your defense against it*. SANS Reading Room.

Gupta, D. (2020). *Marriott Data Breach 2020.* https://www.loginradius.com/blog/start-with-identity/marriott-data-breach-2020/

Gupta, S., & Godavarti, R. (2020). Iot data management using cloud computing and big data technologies. *International Journal of Software Innovation*, *8*(4), 50–58. doi:10.4018/IJSI.2020100104

Han, Y., Seo, S. S., Li, J., Hyun, J., Yoo, J. H., & Hong, J. W. K. (2014, September). Software defined networking-based traffic engineering for data center networks. In *The 16th Asia-Pacific Network Operations and Management Symposium* (pp. 1-6). IEEE. 10.1109/APNOMS.2014.6996601

Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *9*(4), e1306. doi:10.1002/widm.1306

Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). *A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms*. Mobile Information Systems.

Hardy, W., Chen, L., Hou, S., Ye, Y., & Li, X. (2016). DL4MD: A deep learning framework for intelligent malware detection. In *Proceedings of the International Conference on Data Science (ICDATA)* (p. 61). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Hasibur Rahman, R. R. (2019). The role of mobile edge computing towards assisting IoT with distributed intelligence: a smartliving perspective. Mobile Solutions and Their Usefulness in Everyday Life, 33-45.

Hass, R. (2019). *What's Powering Artiicial Intelligence?* ARM.

He, Y., Lee, R., Huai, Y., Shao, Z., Jain, N., Zhang, X., & Xu, Z. (2011, April). RCFile: A fast and space-efficient data placement structure in MapReduce-based warehouse systems. In *2011 IEEE 27th International Conference on Data Engineering* (pp. 1199-1208). IEEE. 10.1109/ICDE.2011.5767933

Heintz, B. A. C. (2015). Optimizing grouped aggregation in geo-distributed streaming analytics. *Proc. ACM HPDC*, 133–144. 10.1145/2749246.2749276

He, K. X. Z. (2016). Deep residual learning for image recognition. *Proc. IEEE CVPR*, 770–778.

Herodotou, H., Lim, H., Luo, G., Borisov, N., Dong, L., Cetin, F. B., & Babu, S. (2011, January). Starfish: A Self-tuning System for Big Data Analytics. In Cidr (Vol. 11, No. 2011, pp. 261-272). Academic Press.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735 PMID:9377276

Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., & Mylopoulos, J. (2014). Strategic business modeling: Representation and reasoning. *Software & Systems Modeling*, *13*(3), 1015–1041. doi:10.100710270-012-0290-8

Hu, L., Liu, Z., Hu, W., Wang, Y., Tan, J., & Wu, F. (2020). Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network. *Journal of Manufacturing Systems*, *55*, 1–14. doi:10.1016/j.jmsy.2020.02.004

Hussain, M., Zhu, W., Zhang, W., Syed, M. R. A., & Ali, S. (2018). Using machine learning to predict student difðculties from learning session data. In Artificial Intelligence. Springer.

Hussain, Z., Akhunzada, A., Iqbal, J., Bibi, I., & Gani, A. (2021). Secure IIoT-Enabled Industry 4.0. *Sustainability*, *13*(22), 12384. doi:10.3390u132212384

Ilayaraja, M., & Meyyappan, T. (2015). Efficient data mining method to predict the risk of heart diseases through frequent itemsets. *Procedia Computer Science*, *70*, 586–592.

Internet Edge Solution Overview. (2010). Academic Press.

Islam, M. R., & Aktheruzzaman, K. M. (2020). An analysis of cybersecurity attacks against internet of things and security solutions. *Journal of Computer and Communications*, *8*(4), 11–25. doi:10.4236/jcc.2020.84002

Jain, V., Kumar, B., & Gupta, A. (2022). Cybertwin-driven resource allocation using deep reinforcement learning in 6G-enabled edge environment. *Journal of King Saud University-Computer and Information Sciences*.

Jain, D. K., Tyagi, S. K. S., Neelakandan, S., Prakash, M., & Natrayan, L. (2021). Metaheuristic optimization-based resource allocation technique for cybertwin-driven 6G on IoE environment. *IEEE Transactions on Industrial Informatics*, *18*(7), 4884–4892. doi:10.1109/TII.2021.3138915

Javed, S. H., Ahmad, M. B., Asif, M., Almotiri, S. H., Masood, K., & Ghamdi, M. A. A. (2022). An Intelligent System to Detect Advanced Persistent Threats in Industrial Internet of Things (I-IoT). *Electronics (Basel)*, *11*(5), 742. doi:10.3390/electronics11050742

Jeans, D. (2019). *Related's Hudson Yards: Smart City or Surveillance City?* Retrieved from https://therealdeal.com/2019/03/15/hudsonyards-smart-city-or-surveillance-city/

Jiang, Salama, & King. (2017). A Public-Key Approach of Selective Encryption for Images. International Journal of Network Security, 19(1), 118-126.

Ji, B., Wang, Y., Song, K., Li, C., Wen, H., Menon, V. G., & Mumtaz, S. (2021). A survey of computational intelligence for 6G: Key technologies, applications and trends. *IEEE Transactions on Industrial Informatics*, *17*(10), 7145–7154. doi:10.1109/TII.2021.3052531

Jibilian, I. (2021). *The US is readying sanctions against Russia over the SolarWinds cyber-attack*. https://www.businessinsider.in/tech/news/heres-a-simple-explanation-of-how-the-massive-solarwinds-hack-happened-and-why-its-such-a-big-deal/articleshow/79945993.cms

Juneja, S., Gahlan, M., Dhiman, G., & Kautish, S. (2021). Futuristic cyber-twin architecture for 6G technology to support internet of everything. *Scientific Programming*, *2021*, 2021. doi:10.1155/2021/9101782

Kang, J., Yu, R., Huang, X., & Zhang, Y. (2018). Privacy-preserved pseudonym scheme for fog computing supported Internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *19*(8), 2627–2637. doi:10.1109/TITS.2017.2764095

Karia. (2021, August 31). *Advanced Persistent Threat (APT) Groups*. Cyber Sophia. https://cybersophia.net/cyber-threat-intel/advanced-persistent-threat-apt-groups/

Kaur, S. (2021). Adding Personal Touches to IoT. In *Big Data Analytics for Internet of Things* (pp. 167–186). Wiley.

Khaliq, S., Tariq, Z. U. A., & Masood, A. (2020, October). Role of User and Entity Behavior Analytics in Detecting Insider Attacks. In *2020 International Conference on Cyber Warfare and Security (ICCWS)* (pp. 1-6). IEEE. 10.1109/ICCWS48432.2020.9292394

Khan, L. U., Tun, Y. K., Alsenwi, M., Imran, M., Han, Z., & Hong, C. S. (2021). A dispersed federated learning framework for 6G-enabled autonomous driving cars. arXiv preprint arXiv:2105.09641.

Khan, L. U., Han, Z., Niyato, D., & Hong, C. S. (2021). Socially-aware-clustering-enabled federated learning for edge networks. *IEEE eTransactions on Network and Service Management*, *18*(3), 2641–2658. doi:10.1109/TNSM.2021.3090446

Khan, L. U., Yaqoob, I., Imran, M., Han, Z., & Hong, C. S. (2020). 6G wireless systems: A vision, architectural elements, and future directions. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 147029–147044. doi:10.1109/ACCESS.2020.3015289

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*(1). doi:10.1177/2053951716631130

Kochunas, B., & Huan, X. (2021). Digital twin concepts with uncertainty for nuclear power applications. *Energies*, *14*(14), 4235. doi:10.3390/en14144235

Kotis, K. K. (2012). Semantic interoperability on the web of things: the semantic smart gateway framework. *Proceedings of the IEEE Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, 630–635. 10.1109/CISIS.2012.200

Krizhevsky, A. I. S. (2012). Imagenet classification with deep convolutional neural networks, *Proc. NIPS*, 1097–1105.

Kucukoglu, I., Atici-Ulusu, H., Gunduz, T., & Tokcalar, O. (2018). Application of the artificial neural network method to detect defective assembling processes by using a wearable technology. *Journal of Manufacturing Systems*, *49*, 163–171. doi:10.1016/j.jmsy.2018.10.001

Kumar, R., Srivastava, R., & Balas, V. E. (2019). Recent Trends and Advances in Artificial Intelligence and Internet of Things. Springer International Publishing.

Lai, Z.-H., Tao, W., Leu, M. C., & Yin, Z. (2020). Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing. *Journal of Manufacturing Systems*, *55*, 69–81. doi:10.1016/j.jmsy.2020.02.010

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 21–32.

Lee, R., Luo, T., Huai, Y., Wang, F., He, Y., & Zhang, X. (2011, June). Ysmart: Yet another sql-to-mapreduce translator. In *2011 31st International Conference on Distributed Computing Systems* (pp. 25-36). IEEE.

Le, Y., & Cun, Y. B. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539 PMID:26017442

Lim, K. Y. H., Zheng, P., & Chen, C. H. (2020). A state-of-the-art survey of Digital Twin: Techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, *31*(6), 1313–1337. doi:10.100710845-019-01512-w

Liu, C., Vengayil, H., Lu, Y., & Xu, X. (2019). A cyber-physical machine tools platform using opc ua and mtconnect. *Journal of Manufacturing Systems*, *51*, 61–74. doi:10.1016/j.jmsy.2019.04.006

Liu, W. (2016). SSD: Single shot multibox detector. *Proc. Eur. Conf. Comput. Vis.*, 21–37.

Li, Y., Ma, R., & Jiao, R. (2015). A Hybrid Malicious Code Detection Method based on Deep Learning. *International Journal of Software Engineering and Its Applications*, *9*, 205–216.

Lu & Leng. (2017). Reversible Dual-Image-Based Hiding Scheme Using Block Folding Technique. *Information Technology And Its Applications*.

Lu, F., Hu, J., Yang, L. T., Tang, Z., Li, P., Shi, Z., & Jin, H. (2020). Energyeffificient traffific offlfloading for mobile users in two-tier heterogeneous wireless networks. *Future Generation Computer Systems*, *105*, 855–863. doi:10.1016/j.future.2017.08.008

Lu, H. C., Chu, Y. P., & Hwang, M. S. (2006). A new steganographic method of the pixel-value differencing. *The Journal of Imaging Science and Technology*, *50*(5), 424–426. doi:10.2352/J.ImagingSci.Technol.(2006)50:5(424)

Lu, T. C., Tseng, C. Y., & Wu, J. H. (2015). Dual imaging-based reversible hiding technique using LSB matching. *Signal Processing*, *108*, 77–89. doi:10.1016/j.sigpro.2014.08.022

Lu, Y., & Asghar, M. R. (2020). Semantic communications between distributed cyber-physical systems towards collaborative automation for smart manufacturing. *Journal of Manufacturing Systems*, *55*, 348–359. doi:10.1016/j.jmsy.2020.05.001

Mahmood, T., & Afzal, U. (2013, December). Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. In *2013 2nd national conference on Information assurance (NCIA)* (pp. 129-134). IEEE.

Maimon, O., & Rokach, L. (2005). Data Mining and Knowledge Discovery Handbook. *Springer*.

Manoharan & RajKumar. (2016). Pixel Value Differencing Method Based on CMYK Colour Model. *Int. J. of Electronics and Information Engineering, 5*(1), 37-46.

Manoharan, S. (2016). Pixel value differencing method based on CMYK colour model. *International Journal of Electronics and Information Engineering*, *5*(1), 37–46.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Mao, C. Y. (2017). A survey on mobile edge computing: The communication perspective. *IEEE Commun. Surveys Tuts., 19*(4), 2322–2358.

Mao, H., Yao, S., Tang, T., Li, B., Yao, J., & Wang, Y. (2018). Towards real-time object detection on embedded systems. *IEEE Transactions on Emerging Topics in Computing*, *6*(3), 417–431. doi:10.1109/TETC.2016.2593643

Markstedte. (2017, May 20). *Tactics, Techniques, and Procedures (TTPs).* Azeria-Labs. https://azeria-labs.com/tactics-techniques-and-procedures-ttps/

Maroto. (2019). *Industrial IoT - Edge Computing Vendors Overview.* Independently Published.

Mayani, M. G., Svendsen, M., & Oedegaard, S. (2018). Drilling digital twin success stories the last 10 years. In *SPE Norway One Day Seminar*. OnePetro. 10.2118/191336-MS

Mazhelis, O. W. (2013). *Internet-of-things market, value networks, and business models: State of the art report*. University of Jyvaskyla.

MetiTalk. (2016). *MeriTalk Cloudera Disruptive Press Release*. https://www.meritalk.com/wp-content/uploads/2016/08/49503-MeriTalk_Cloudera_Disruptive_Press_Release_FINAL_082916.pdf

Meurisch, C. S. (2015). Upgrading wireless home routers for enabling large-scale deployment of cloudlets. *International Conference on Mobile Computing, Applications, and Services*, 12–29. 10.1007/978-3-319-29003-4_2

Mitchell, T. M., & Mitchell, T. M. (1997). Article. *Machine Learning*, *1*(9).

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. doi:10.1038/nature14236 PMID:25719670

Moghaddam, M., Cadavid, M. N., Kenley, C. R., & Deshmukh, A. V. (2018). Reference architectures for smart manufacturing: A critical review. *Journal of Manufacturing Systems*, *49*, 215–225. doi:10.1016/j.jmsy.2018.10.006

Mohamed, S., Ismail, O., & Hogan, O. (2012). *Data equity: Unlocking the value of big data*. Centre for Economics and Business Research.

Mohr, J. (2018). *Digital twins for the oil and gas industry. Hashplay*. Tech. Rep.

Moradi, M., & Zulkernine, M. (2004, November). A neural network based system for intrusion detection and classification of attacks. In *Proceedings of the IEEE international conference on advances in intelligent systems-theory and applications* (pp. 15-18). IEEE.

Mouradian, D. N. (2018). A comprehensive survey on fog computing: State-of-theArt and research challenges. *IEEE Commun. Surveys Tuts., 20*(1), 416–464.

Mouthami, K., Devi, K. N., & Bhaskaran, V. M. (2013, February). Sentiment analysis and classification based on textual reviews. In 2013 international conference on Information communication and embedded systems (ICICES) (pp. 271-276). IEEE. doi:10.1109/ICICES.2013.6508366

Muhammad Sharjeel Zareen, S. T. (2019). Artiicial Intelligence/ Machine Learning in IoT for Authentication and Authorization of Edge Devices. *2019 International Conference on Applied and Engineering Mathematics (ICAEM)*, 220-224. 10.1109/ICAEM.2019.8853780

Munusamy, A., Adhikari, M., Balasubramanian, V., Khan, M. A., Menon, V. G., Rawat, D., & Srirama, S. N. (2021). Service deployment strategy for predictive analysis of FinTech IoT applications in edge networks. *IEEE Internet of Things Journal*, 1. doi:10.1109/JIOT.2021.3078148

Nadembega, A. A. S. (2016). Mobility prediction modelbased service migration procedure for follow me cloud to support QoS and QoE. *Proc. IEEE Int. Conf. Commun. (ICC)*, 1–6.

Nadhan, D., Mayani, M. G., & Rommetveit, R. (2018). Drilling with digital twins. In *IADC/SPE Asia pacific drilling technology conference and exhibition*. OnePetro.

Nagavci, D., Hamiti, M., & Selimi, B. (2018). Review of Prediction of Disease Trends using Big Data Analytics. *International Journal of Advanced Computer Science and Applications.*, *9*, 8.

Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, *40*(4), 1086–1093.

Napoleone, A., Macchi, M., & Pozzetti, A. (2020). A review on the characteristics of cyber-physical systems for the future smart factories. *Journal of Manufacturing Systems*, *54*, 305–335. doi:10.1016/j.jmsy.2020.01.007

Negri, E., Berardi, S., Fumagalli, L., & Macchi, M. (2020). Mes-integrated digital twin frameworks. *Journal of Manufacturing Systems*, *56*, 58–71. doi:10.1016/j.jmsy.2020.05.007

Northeastern University. (2016). *How Much Data is Produced Every Day?* https://www.northeastern.edu/graduate/blog/how-much-data-produced-every-day/

OpenFog reference architecture for fog computing. (2017). *OpenFog Consortium*.

Padhi, P. K., & Charrua-Santos, F. (2021). 6G enabled industrial internet of everything: Towards a theoretical framework. *Applied System Innovation*, *4*(1), 11. doi:10.3390/asi4010011

Padiya, T., Bhise, M., & Rajkotiya, P. (2015). Data management for internet of things. In *2015 IEEE Region 10 Symposium* (pp. 62–65). IEEE. 10.1109/TENSYMP.2015.26

Pankajavalli, P. B., & Karthick, G. S. (2019). Incorporating the Internet of Things in Healthcare Applications and Wearable Devices. IGI Global.

Park, H.-S., & Tran, N.-H. (2012). An autonomous manufacturing system based on swarm of cognitive agents. *Journal of Manufacturing Systems*, *31*(3), 337–348. doi:10.1016/j.jmsy.2012.05.002

Parnas, D. L., & Madey, J. (1995). Functional documents for computer systems. *Science of Computer Programming*, *25*(1), 41–61. doi:10.1016/0167-6423(95)96871-J

Parrott, A., & Warshaw, L. (2017). *Industry 4.0 and the digital twin*. Deloitte Insights.

Patierno, P. (2016). Hybrid IoT: On fog computing, gateways, and protocol translation. *DZone/IoT Zone.*

Pedreira, V., Barros, D., & Pinto, P. (2021). A review of attacks, vulnerabilities, and defenses in industry 4.0 with new challenges on data sovereignty ahead. *Sensors (Basel)*, *21*(15), 5189. doi:10.339021155189 PMID:34372425

Piateski, G., & Frawley, W. (1991). Knowledge Discovery in Databases. AAAI Press.

Pignaton de Freitas, E., Olszewska, J. I., Carbonera, J. L., Fiorini, S. R., Khamis, A., Ragavan, S. V., Barreto, M. E., Prestes, E., Habib, M. K., Redfield, S., Chibani, A., Goncalves, P., Bermejo-Alonso, J., Sanz, R., Tosello, E., Olivares-Alarcos, A., Konzen, A. A., Quintas, J., & Li, H. (2020). Ontological concepts for information sharing in cloud robotics. *Journal of Ambient Intelligence and Humanized Computing*, 1–12. doi:10.100712652-020-02150-4

Plattner, H., & Zeier, A. (2012). *In-memory data management: technology and applications*. Springer Science & Business Media. doi:10.1007/978-3-642-29575-1

Poddar, T. (2018). Digital twin bridging intelligence among man, machine and environment. In *Offshore Technology Conference Asia*. OnePetro.

Prices, C. O. (2017). *Year historical chart.* http://www.macrotrends.net/1369/crude-oil-price-history-chart

Pu, Q. (2015). Low latency geo-distributed data analytics. *Proc. ACM SIGCOMM*, 421–434.

PWC. (2015). *Information Security Breaches Survey*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/432412/bis-15-302-information_security_breaches_survey_2015-full-report.pdf

Qadir, Z., Le, K. N., Saeed, N., & Munawar, H. S. (2022). *Towards 6G Internet of Things: Recent advances, use cases, and open challenges*. ICT Express.

Quintero-Bonilla, S., & Martín del Rey, A. (2020). A new proposal on the advanced persistent threat: A survey. *Applied Sciences (Basel, Switzerland)*, *10*(11), 3874. doi:10.3390/app10113874

Radware. (2019, December 8). *Tactics, Techniques and Procedures*. www.radware.com

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 1–10. doi:10.1186/2047-2501-2-3 PMID:25825667

Ramezani, P., Lyu, B., & Jamalipour, A. (2022). Toward RIS-Enhanced Integrated Terrestrial/Non-Terrestrial Connectivity in 6G-Enabled IoE Era. arXiv preprint arXiv:2203.11312.

Ram, S., Zhang, W., Williams, M., & Pengetnze, Y. (2015). Predicting asthma-related emergency department visits using big data. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1216–1223.

Razak, T. R., Hashim, M. A., Noor, N. M., Iman, H. A. H., & Nur, F. F. S. (2015). *Career Path Recommendation System for UiTM Perlis Students Using Fuzzy Logic*. IEEE.

Redmon, J. S. D. (2016). You only look once: Unified, real-time object detection. *in Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, 779–788.

Reinders, J. (2007). *Intel Threading Building Blocks: Outfifitting C++ for Multi-core Processor Parallelism*. O'Reilly.

Rengarajan, R., & Babu, S. (2021, March). Anomaly Detection using User Entity Behavior Analytics and Data Visualization. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 842-847). IEEE.

Ren, L., Sun, Y., Cui, J., & Zhang, L. (2018). Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. *Journal of Manufacturing Systems*, *48*, 71–77. doi:10.1016/j.jmsy.2018.04.008

Richardson, R., & North, M. M. (2017). Ransomware: Evolution, mitigation and prevention. *International Management Review*, *13*(1), 10.

Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, *78*, 680–698. doi:10.1016/j.future.2016.11.009

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). G. E. (n.d.). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. doi:10.1038/323533a0

Russom, P. (2011). *Big Data Analytics*. TDWI Best Practices Report, Fourth Quarter, 2011.

S´anchez, F., & Hartlieb, P. (2020). Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation. *Mining, Metallurgy & Exploration, 37*(5), 1385–1399.

Sahoo, P. K., Mohapatra, S. K., & Wu, S. L. (2016). Analyzing healthcare big data with prediction for future health condition. *IEEE Access : Practical Innovations, Open Solutions*, *4*, 9786–9799.

Saini, G., Ashok, P., van Oort, E., & Isbell, M. R. (2018). Accelerating well construction using a digital twin demonstrated on unconventional well data in north america. In *Unconventional Resources Technology Conference*, *Houston, Texas*, 23-25 July 2018 (pp. 3264–3276). Society of Exploration Geophysicists, American Association of Petroleum. 10.15530/urtec-2018-2902186

Saldamli, G., & Razavi, A. (2020, October). Surveillance Missions Deployment on the Edge by Combining Swarm Robotics and Blockchain. In *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)* (pp. 106-112). IEEE. 10.1109/MCNA50957.2020.9264274

Sánchez, D., Martín-Bautista, M. J., Blanco, I., & de la Torre, C. J. (2008, December). Text knowledge mining: an alternative to text data mining. In *2008 IEEE International Conference on Data Mining Workshops* (pp. 664-672). IEEE. 10.1109/ICDMW.2008.57

Satyanarayanan, M. (2017). The emergence of edge computing. *Computer*, *50*(1), 30–39. doi:10.1109/MC.2017.9

Seraphin, B., & Calo, M. T. (2017). Edge computing architecture for applying AI to IoT. *2017 IEEE International Conference on Big Data (Big Data)*, 3012-3016.

Shah, Y., & Sengupta, S. (2020, October). *A survey on Classification of Cyber-attacks on IoT and IIoT devices. In 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 406-413)*. IEEE.

Shakil Ahamed, Mahmood, & Rahman. (2017). An intelligent system to predict academic performance based on different factors during adolescence. *Journal of Information and Telecommunication*.

Shakir, R. (2018). Integrating Learning Analytics To Predict Student Performance Behavior. Academic Press.

Sharma, P., Knezevic, D., Huynh, P., & Malinowski, G. (2018). Rb-fea based digital twin for structural integrity assessment of offshore structures. In *Offshore Technology Conference*. OnePetro. 10.4043/29005-MS

Shi, H. S. (2017). Edge computing-an emerging computing model for the Internet of everything era. *J. Comput. Res. Develop.*, *54*(5), 907–924.

Shi, X. Z. (2019). Edge computing: State-of-the-art and future directions. *J. Comput. Res. Develop.*, *56*(1), 1–21. doi:10.1360/crad20070101

Shiu, Chen, & Hong. (2019). *Reversible Data Hiding in Permutation-based Encrypted Images with Strong Privacy.* Academic Press.

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637–646. doi:10.1109/JIOT.2016.2579198

Singh, P., Singh, S., & Pandi-Jain, G. S. (2018). Effective heart disease prediction system using data mining techniques. *International Journal of Nanomedicine*, *13*, 121.

Slim, A., Heileman, G. L., Kozlick, J., & Abdallah, C. T. (2014). *Predicting Student Success Based on Prior Performance*. IEEE. doi:10.1109/CIDM.2014.7008697

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods, eight edition*. Iowa state University press.

Spiteri. (2021, June 14). *Adversary emulation with Prelude Operator and Elastic Security*. https://www.elastic.co/blog/adversary-emulation-with-prelude-operator-and-elastic-security

Sripath Roy, K., Roopkanth, K., Uday Teja, V., Bhavana, V., & Priyanka, J. (2018). Student Career Prediction Using Advanced Machine Learning Techniques. *IACSIT International Journal of Engineering and Technology*, *7*(2.20), 26. doi:10.14419/ijet.v7i2.20.11738

Staff, E. (2020, August 4). *Overview of Industrial Control Systems*. Inst Tools. https://instrumentationtools.com/overview-of-industrial-control-systems/

Supakkul, S., & Chung, L. (2009, March). Extending problem frames to deal with stakeholder problems: An agent-and goal-oriented approach. In *Proceedings of the 2009 ACM symposium on Applied Computing* (pp. 389-394). ACM.

Svozil, D., Kvasnicka, V., & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, *39*(1), 43–62. doi:10.1016/S0169-7439(97)00061-0

Swain, A. K. (2016). Mining big data to support decision making in healthcare. *Journal of Information Technology Case and Application Research*, *18*(3), 141–154.

Szegedy, C. (2015). Going deeper with convolutions. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, 1–9.

Tang, W. J. (2014). Data-aware resource scheduling for multicloud workflflows: a fifine-grained simulation approach. *2014 IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom)*, 887–892.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018a). Digital twin-driven product design, manufacturing and service with big data. *International Journal of Advanced Manufacturing Technology*, *94*(9), 3563–3576. doi:10.100700170-017-0233-1

Tao, F., Liu, W., Zhang, M., & Hu, T. (2019). Five-dimension digital twin model and its ten applications. *Computer Integrated Manufacturing Systems*, *25*, 1–18.

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. (2018b). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405–2415. doi:10.1109/TII.2018.2873186

The Internet of Things reference model. (2014). *Internet of Things World Forum,* 1–12.

Tian, J. (2003). Reversible data embedding using a difference expansion. *IEEE Transactions on Circuits System Video Technology*, *13*(8), 890–896. doi:10.1109/TCSVT.2003.815962

Tollefson, R. (2021, March 26). *Security+: How to explain threat actor types and attributes* [updated 2021]. Infosec Resources. https://resources.infosecinstitute.com/ certification /security-how-to-explain-threat-actor-types-and-attributes/

Tom Young, D. H. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine, 13*, 55-75.

Trustwave. (2015). *Trustwave Global Security Report*. http://book.itep.ru/depository/annuals/2015_TrustwaveGlobalSecurityReport.pdf

Tuptuk, N., & Hailes, S. (2018). Security of smart manufacturing systems. *Journal of Manufacturing Systems*, *47*, 93–106. doi:10.1016/j.jmsy.2018.04.007

Twitter Investigation Report. (2020). https://www.dfs.ny.gov/Twitter_Report

Uckelmann, D. H. (2011). An architectural approach towards the future Internet of things. Architecting the Internet of Things, 1–24.

Uddin & Lee. (2016). Predicting Good Fit Students by Correlating Relevant Personality Traits with Academic/Career Data. IEEE.

Uhlemann, T. H.-J., Lehmann, C., & Steinhilper, R. (2017). The digital twin: Realizing the cyber-physical production system for industry 4.0. *Procedia CIRP*, *61*, 335–340. doi:10.1016/j.procir.2016.11.152

Unit, E. I. (2012). The deciding factor: Big data & decision making. *Capgemini Reports*, 1-24.

Upadhyay, H. K., Juneja, S., Maggu, S., Dhingra, G., & Juneja, A. (2021). Multi-criteria analysis of social isolation barriers amid COVID-19 using fuzzy AHP. *World Journal of Engineering*.

ur Rehmana, Yaqoobb, Salahc, Imrand, Jayaramane, & Pereraf. (2019). The role of big data analytics in industrial Internet of Things. *Future Generation Computer Systems, 99,* 247-259.

Uyar, K., & İlhan, A. (2017). Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks. *Procedia Computer Science*, *120*, 588–593.

Varghese, B. W. (2017). Feasibility of fog computing. *arXiv*.

Varghese, B. W. (2016). Challenges and opportunities in edge computing. *IEEE International Conference on Smart Cloud (SmartCloud)*, 20–26. 10.1109/SmartCloud.2016.18

Velliangiri, S., Anbarasu, V., Karthikeyan, P., & Anandaraj, S. P. (2021). Intelligent Personal Health Monitoring and Guidance Using Long Short-Term Memory. *Journal of Mobile Multimedia,* 349-372.

Velliangiri, S., Karthikeyan, P., Joseph, I. T., & Kumar, S. A. (2019, December). Investigation of deep learning schemes in medical application. In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 87-92). IEEE.

Vemuri, V. R. (2005). *Enhancing computer security with smart technology*. Auerbach Publications. doi:10.1201/9780849330452

Viceconti, M., Hunter, P., & Hose, R. (2015). Big data, big knowledge: Big data for personalized healthcare. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1209–1215.

Vivekanandan, T., & Iyengar, N. C. S. N. (2017). Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Computers in Biology and Medicine*, *90*, 125–136.

Vongsingthong, S., & Smanchat, S. (2015). A review of data management in internet of things. *Asia-Pacific Journal of Science and Technology*, *20*(2), 215–240.

Vu, L., Bui, C. T., & Nguyen, Q. U. (2017). A Deep Learning Based Method for Handling Imbalanced Problem in Network Traffic Classification. In *Proceedings of the Eighth International Symposium on Information and Communication Technology* (pp. 333–339). Association for Computing Machinery. 10.1145/3155133.3155175

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. doi:10.1016/j.ijpe.2014.12.031

Wanasinghe, T. R., Wroblewski, L., Petersen, B. K., Gosine, R. G., James, L. A., De Silva, O., Mann, G. K., & Warrian, P. J. (2020). Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 104175–104197. doi:10.1109/ACCESS.2020.2998723

Wang, D. Z. (2017). Mobile network edge computing and caching technology. *Railway Comput. Appl.*, *26*(8), 51–54.

Wang, R., Li, M., Peng, L., Hu, Y., Hassan, M. M., & Alelaiwi, A. (2020). Cognitive multi-agent empowering mobile edge computing for resource caching and collaboration. *Future Generation Computer Systems*, *102*, 66–74. doi:10.1016/j.future.2019.08.001

Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination. *Computer Networks*, *101*, 158–168. doi:10.1016/j.comnet.2015.12.017

Wang, W. O. (2013). Joint deep learning for pedestrian detection. *Proceedings - IEEE International Conference on Computer Vision*, 2056–2063.

Wang, X. H. (2018). Edge computing technology: Development and countermeasures. *Chin. J. Eng. Sci.*, *20*(2), 20. doi:10.1016/j.jes.2017.11.002

Wang, Y. L., Shen, J. J., & Hwang, M. S. (2017). An improved dual image-based reversible hiding technique using LSB matching. *International Journal of Network Security*, *19*(5), 858–862.

Watson, I. (2016). *The power of analytics at the edge*. IBM.

Weiser, M. G., Gold, R., & Brown, J. S. (1999). The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal*, *38*(4), 693–696. doi:10.1147j.384.0693

Werbos, P. J. (1990). Backpropagation through time: What it does and how to do it. *Proceedings of the IEEE*, *78*(10), 1550–1560. doi:10.1109/5.58337

Wikipedia. (2020). *Easyjet Data Breach*. https://en.wikipedia.org/wiki/EasyJet_data_breach

Wu, N. I., & Hwang, M. S. (2017). Development of a data hiding scheme based on combination theory for lowering the visual noise in binary images. *Displays*, *49*, 116–123. doi:10.1016/j.displa.2017.07.009

Yin, D. K. (2011). A data-aware workflflow scheduling algorithm for heterogeneous distributed systems. *International Conference on High Performance Computing and Simulation (HPCS) IEEE*, 114–120.

Yonghui Wu, M. S. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv*.

Yu, E., Giorgini, P., Maiden, N. A., & Mylopoulos, J. (2011). *Social Modeling for Requirements Engineering: An Introduction*. Academic Press.

Yuan, Z., Lu, Y., Wang, Z., & Xue, Y. (2014, August). Droid-sec: deep learning in android malware detection. In *Proceedings of the 2014 ACM conference on SIGCOMM* (pp. 371-372). 10.1145/2619239.2631434

Yu, L. D. (2014). Deep learning: Methods and applications. *Found. Trends Signal Process.*, *7*(3–4), 197–387.

Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020a). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, *55*, 317–324. doi:10.1016/j.jmsy.2020.03.009

Yu, Q., Ren, J., Zhou, H., & Zhang, W. (2020, March). *A cybertwin based network architecture for 6G. In 2020 2nd 6G Wireless Summit (6G SUMMIT)*. IEEE.

Zhang. (2004). *Knowledge Discovery In Multiple Databases.* Springer.

Zhang, T. A. C. (2015). The design and implementation of a wireless video surveillance system. *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 426-438. 10.1145/2789168.2790123

Zhao, Zhang, & Yao. (2017). DeepSense: A uniied deep learning framework for time-series mobile sensing data processing. *Proc. 26th Int. Conf. World Wide Web*, 351-360.

Zha, Z. M., F. L. (2018). Edge computing: Platforms; Applications and challenges. *J. Comput. Res. Develop.*, *55*(2), 327–337.

Zhenqiu Huang, K.-J. L.-L.-S. (2018). Building edge intelligence for online activity recognition in service-oriented IoT systems. *Future Generation Computer Systems*, *87*, 557–567. doi:10.1016/j.future.2018.03.003

Zhidchenko, V., Malysheva, I., Handroos, H., & Kovartsev, A. (2018). Faster than real-time simulation of mobile crane dynamics using digital twin concept. *Journal of Physics: Conference Series*, *1096*, 012071. doi:10.1088/1742-6596/1096/1/012071

Zhou, H. (2013). *The Internet of Things in the Cloud: A Middleware Perspective*. CRC Press.

Zhou, M., Yan, J., & Feng, D. (2019). Digital twin framework and its application to power grid online analysis. *CSEE Journal of Power and Energy Systems*, *5*(3), 391–398.

Zisserman, K. S. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv*.

Zobaa, A. F., & Bihl, T. J. (2018). Big Data Analytics in Future Power Systems. CRC Press.

Related Readings

To continue IGI Global's long-standing tradition of advancing innovation through emerging research, please find below a compiled list of recommended IGI Global book chapters and journal articles in the area of the data analytics and the internet of things as it applies to the field of digital twin. These related readings will provide additional information and guidance to further enrich your knowledge and assist you with your own research.

A., S., & Gunasekaran G.,. (2019). A Novel Approach of Load Balancing and Task Scheduling Using Ant Colony Optimization Algorithm. *International Journal of Software Innovation (IJSI), 7*(2), 9-20. doi:10.4018/IJSI.2019040102

A., S., K. S., B., Shanmugam, B., & Pitchaimuthu, S. (2020). Deep Learning Techniques for Biomedical Image Analysis in Healthcare. In Suresh, A., Udendhran, R., & Vimal, S. (Ed.), *Deep Neural Networks for Multimodal Imaging and Biomedical Applications* (pp. 31-46). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3591-2.ch003

Abouzid, H., & Chakkor, O. (2020). Autoencoders in Deep Neural Network Architecture for Real Work Applications. In J. Zbitou, C. I. Pruncu, & A. Errkik (Eds.), *Handbook of Research on Recent Developments in Electrical and Mechanical Engineering* (pp. 214–236). IGI Global., doi:10.4018/978-1-7998-0117-7.ch007

Abtahi, A., & Abdi, F. (2018). Designing Software as a Service in Cloud Computing Using Quality Function Deployment. [IJEIS]. *International Journal of Enterprise Information Systems*, *14*(4), 16–27. doi:10.4018/IJEIS.2018100102

Adegun, A. A., Ogundokun, R. O., Adebiyi, M. O., & Asani, E. O. (2020). CAD-Based Machine Learning Project for Reducing Human-Factor-Related Errors in Medical Image Analysis. In S. Misra & A. Adewumi (Eds.), *Handbook of Research on the Role of Human Factors in IT Project Management* (pp. 164–172). IGI Global., doi:10.4018/978-1-7998-1279-1.ch011

Afolayan, J. O., Ogundokun, R. O., Afolabi, A. G., & Adegun, A. A. (2020). Artificial Intelligence, Cloud Librarianship, and Infopreneurship Initiatives for Inclusiveness. In A. Tella (Ed.), *Handbook of Research on Digital Devices for Inclusivity and Engagement in Libraries* (pp. 45–69). IGI Global., doi:10.4018/978-1-5225-9034-7.ch003

Afzal, S., & Kavitha, G. (2020). A Hybrid Multiple Parallel Queuing Model to Enhance QoS in Cloud Computing. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *12*(1), 18–34. doi:10.4018/IJGHPC.2020010102

Agarwal, M., & Srivastava, G. M. (2019). A PSO Algorithm Based Task Scheduling in Cloud Computing. [IJAMC]. *International Journal of Applied Metaheuristic Computing*, *10*(4), 1–17. doi:10.4018/IJAMC.2019100101

Agarwal, P., Hassan, S. I., Mustafa, S. K., & Ahmad, J. (2020). An Effective Diagnostic Model for Personalized Healthcare Using Deep Learning Techniques. In R. Wason, D. Goyal, V. Jain, S. Balamurugan, & A. Baliyan (Eds.), *Applications of Deep Learning and Big IoT on Personalized Healthcare Services* (pp. 70–88). IGI Global., doi:10.4018/978-1-7998-2101-4.ch005

Ahuja, S. P., Czarnecki, E., & Willison, S. (2020). Multi-Factor Performance Comparison of Amazon Web Services Elastic Compute Cluster and Google Cloud Platform Compute Engine. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(3), 1–16. doi:10.4018/IJCAC.2020070101

Akbar, M., Suaib, M., Husain, M. S., & Shukla, S. (2020). A Compendium of Cloud Forensics. In M. Husain & M. Khan (Eds.), *Critical Concepts, Standards, and Techniques in Cyber Forensics* (pp. 215–227). IGI Global., doi:10.4018/978-1-7998-1558-7.ch012

Al-Dmour, A., Al-Dmour, R. H., Al-Dmour, H. H., & Ahmadamin, E. B. (2021). The Effect of Big Data Analytic Capabilities Upon Bank Performance via FinTech Innovation. [IJISSS]. *International Journal of Information Systems in the Service Sector*, *13*(4), 62–87. doi:10.4018/IJISSS.2021100104

Al-Gharibi, M., Warren, M., & Yeoh, W. (2020). Risks of Critical Infrastructure Adoption of Cloud Computing by Government. [IJCWT]. *International Journal of Cyber Warfare & Terrorism*, *10*(3), 47–58. doi:10.4018/IJCWT.2020070104

Al-Momani, A. M., Mahmoud, M. A., & Ahmad, M. S. (2019). A Review of Factors Influencing Customer Acceptance of Internet of Things Services. [IJISSS]. *International Journal of Information Systems in the Service Sector*, *11*(1), 54–67. doi:10.4018/IJISSS.2019010104

Al Qasem, O., & Akour, M. (2019). Software Fault Prediction Using Deep Learning Algorithms. [IJOSSP]. *International Journal of Open Source Software and Processes*, *10*(4), 1–19. doi:10.4018/IJOSSP.2019100101

Alam, T. (2020). IoT-Fog-Blockchain Framework. [IJFC]. *International Journal of Fog Computing*, *3*(2), 1–20. doi:10.4018/IJFC.2020070101

Alcan, P. (2020). Internet of Things (IoT) in Healthcare Systems. In E. Koç (Ed.), *Internet of Things (IoT) Applications for Enterprise Productivity* (pp. 62–90). IGI Global., doi:10.4018/978-1-7998-3175-4.ch003

Alghatrifi, I., & Al Musawi, A. S. (2020). Emerging Technologies and Educational Requirements in Engineering Education for the Fourth Industrial Revolution. In SerdarAsan, Ş., & Işıklı, E. (Ed.), Engineering Education Trends in the Digital Era (pp. 26-52). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2562-3.ch002

Ali, M., & Edghiem, F. (2021). Sustainable Business and Collaboration Driven by Big Data Analytics Amidst the Emergence of the Remote Work Culture. In M. Ali (Ed.), *Remote Work and Sustainable Changes for the Future of Global Business* (pp. 15–32). IGI Global., doi:10.4018/978-1-7998-7513-0.ch002

Ali, M., Wood-Harper, T., & Ramlogan, R. (2020). A Framework Strategy to Overcome Trust Issues on Cloud Computing Adoption in Higher Education. In B. B. Gupta (Ed.), *Modern Principles, Practices, and Algorithms for Cloud Security* (pp. 162–183). IGI Global., doi:10.4018/978-1-7998-1082-7.ch008

Ali, M. B. (2019). Multiple Perspective of Cloud Computing Adoption Determinants in Higher Education a Systematic Review. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(3), 89–109. doi:10.4018/IJCAC.2019070106

Ali, M. B. (2021). Internet of Things (IoT) to Foster Communication and Information Sharing. In M. B. Ali & T. Wood-Harper (Eds.), *Fostering Communication and Learning With Underutilized Technologies in Higher Education* (pp. 1–20). IGI Global., doi:10.4018/978-1-7998-4846-2.ch001

Ali, M. B. (2021). Multi-Perspectives of Cloud Computing Service Adoption Quality and Risks in Higher Education. In Khosrow-Pour D.B.A., M. (Eds.), Handbook of Research on Modern Educational Technologies, Applications, and Management (pp. 1-19). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3476-2.ch001

Alismaili, S. Z., Li, M., Shen, J., Huang, P., He, Q., & Zhan, W. (2020). Organisational-Level Assessment of Cloud Computing Adoption. [JGIM]. *Journal of Global Information Management*, *28*(2), 73–89. doi:10.4018/JGIM.2020040104

Aliyu, M., Murali, M., Gital, A. Y., & Boukari, S. (2020). Efficient Metaheuristic Population-Based and Deterministic Algorithm for Resource Provisioning Using Ant Colony Optimization and Spanning Tree. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(2), 1–21. doi:10.4018/IJCAC.2020040101

Alkhalil, A., Abdallah, M. A., Alogali, A., & Aljaloud, A. (2021). Applying Big Data Analytics in Higher Education. [IJICTE]. *International Journal of Information and Communication Technology Education*, *17*(3), 29–51. doi:10.4018/IJICTE.20210701.oa3

Almeida, F., & Lourenço, J. (2020). Privacy and Security Challenges in the Internet of Things. In Khosrow-Pour D.B.A., M. (Ed.), Encyclopedia of Criminal Activities and the Deep Web (pp. 749-762). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-9715-5.ch051

AlMubarak, H. A., Stanley, J., Guo, P., Long, R., Antani, S., Thoma, G., Zuna, R., Frazier, S., & Stoecker, W. (2019). A Hybrid Deep Learning and Handcrafted Feature Approach for Cervical Cancer Digital Histology Image Classification. [IJHISI]. *International Journal of Healthcare Information Systems and Informatics*, *14*(2), 66–87. doi:10.4018/IJHISI.2019040105

Almulihi, A. H., Alharithi, F. S., Mechti, S., Alroobaea, R., & Rubaiee, S. (2021). A Software for Thorax Images Analysis Based on Deep Learning. [IJOSSP]. *International Journal of Open Source Software and Processes*, *12*(1), 60–71. doi:10.4018/IJOSSP.2021010104

Alsadi, A. K., Alaskar, T. H., & Mezghani, K. (2021). Adoption of Big Data Analytics in Supply Chain Management. [IJISSCM]. *International Journal of Information Systems and Supply Chain Management*, *14*(2), 88–107. doi:10.4018/IJISSCM.2021040105

Althobaiti, O. S. (2020). IoT Limitations and Concerns Relative to 5G Architecture. In J. J. Thomas, U. Fiore, G. P. Lechuga, V. Kharchenko, & P. Vasant (Eds.), *Handbook of Research on Smart Technology Models for Business and Industry* (pp. 45–69). IGI Global., doi:10.4018/978-1-7998-3645-2.ch003

Ambekar, K. A., & Kamatchi, R. (2020). An Unified Secured Cloud System for the Education Sector of India. In G. Cornetta, A. Touhafi, & G. Muntean (Eds.), *Social, Legal, and Ethical Implications of IoT, Cloud, and Edge Computing Technologies* (pp. 69–102). IGI Global., doi:10.4018/978-1-7998-3817-3.ch004

Anand, A., & Muthusamy, A. (2020). Data Security and Privacy-Preserving in Cloud Computing Paradigm. In S. Gochhait, D. T. Shou, & S. Fazalbhoy (Eds.), *Cloud Computing Applications and Techniques for E-Commerce* (pp. 99–133). IGI Global., doi:10.4018/978-1-7998-1294-4.ch006

Andročec, D., Novak, M., & Oreški, D. (2018). Using Semantic Web for Internet of Things Interoperability. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *14*(4), 147–171. doi:10.4018/IJSWIS.2018100108

Angeles, R. (2019). Internet of Things (IOT)-Enabled Product Monitoring at Steadyserv. [JCIT]. *Journal of Cases on Information Technology*, *21*(4), 27–45. doi:10.4018/JCIT.2019100103

Annansingh, F. (2021). Using Big Data Analytics to Assist a Smart City to Prevent Cyber Security Threats. In F. Annansingh (Ed.), *Examining the Socio-Technical Impact of Smart Cities* (pp. 107–124). IGI Global., doi:10.4018/978-1-7998-5326-8.ch005

Ansari, S., Aslam, T., Poncela, J., Otero, P., & Ansari, A. (2020). Internet of Things-Based Healthcare Applications. In B. S. Chowdhry, F. K. Shaikh, & N. A. Mahoto (Eds.), *IoT Architectures, Models, and Platforms for Smart City Applications* (pp. 1–28). IGI Global., doi:10.4018/978-1-7998-1253-1.ch001

Anshul,, & Kumar, R. (2021). Deep Learning Techniques in Perception of Cancer Diagnosis. In Raut, R., & Mihovska, A. D. (Ed.), *Examining the Impact of Deep Learning and IoT on Multi-Industry Applications* (pp. 1-20). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-7511-6.ch001

Anunciação, P. F., Dinis, V. M., & Esteves, F. M. (2021). Industrial Maintenance Entering the Industry 4.0 Era. In N. Geada & P. Anunciação (Eds.), *Reviving Businesses With New Organizational Change Management Strategies* (pp. 229–250). IGI Global., doi:10.4018/978-1-7998-7452-2.ch013

Anwarul, S., & Joshi, D. (2020). Deep Learning With TensorFlow. In M. Mahrishi, K. K. Hiran, G. Meena, & P. Sharma (Eds.), *Machine Learning and Deep Learning in Real-Time Applications* (pp. 96–120). IGI Global., doi:10.4018/978-1-7998-3095-5.ch004

Aouat, A., Deba, E. A., Benyamina, A. E., & Benhamamouch, D. (2020). Deployment in Cloud Computing. [IJDST]. *International Journal of Distributed Systems and Technologies*, *11*(1), 27–37. doi:10.4018/IJDST.2020010103

Appati, J. K., Denwar, I. W., Owusu, E., & Soli, M. A. (2021). Construction of an Ensemble Scheme for Stock Price Prediction Using Deep Learning Techniques. [IJIIT]. *International Journal of Intelligent Information Technologies*, *17*(2), 72–95. doi:10.4018/IJIIT.2021040104

Aridoss, M., Dhasarathan, C., Dumka, A., & Loganathan, J. (2020). DUICM Deep Underwater Image Classification Mobdel using Convolutional Neural Networks. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *12*(3), 88–100. doi:10.4018/IJGHPC.2020070106

Arora, S., & Bhatia, M. P. (2020). Biometrics for Forensic Identification in Web Applications and Social Platforms Using Deep Learning. In K. Sharma, M. Makino, G. Shrivastava, & B. Agarwal (Eds.), *Forensic Investigations and Risk Management in Mobile and Wireless Communications* (pp. 80–113). IGI Global., doi:10.4018/978-1-5225-9554-0.ch004

Arya, K. V., & Gore, R. (2019). Internet of Things Using Software-Defined Network and Cognitive Radio Network. In Bagwari, A., Bagwari, J., & Tomar, G. (Ed.), Sensing Techniques for Next Generation Cognitive Radio Networks (pp. 312-328). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-5354-0.ch017

Asmaa, A., El Abbassia, D., Hassan, B. A., & Djilali, B. (2019). Model-Based Application Deployment on Cloud Computing. [IJDST]. *International Journal of Distributed Systems and Technologies*, *10*(2), 110–127. doi:10.4018/IJDST.2019040106

Awad, W., & Al-Noaimi, D. N. (2019). A Framework for Improving Knowledge Management Using Cloud-Based Business Intelligence. In Y. A. Albastaki, A. I. Al-Alawi, & S. Abdulrahman Al-Bassam (Eds.), *Handbook of Research on Implementing Knowledge Management Strategy in the Public Sector* (pp. 322–345). IGI Global., doi:10.4018/978-1-5225-9639-4.ch018

Ayoobkhan, M. U. D., Y., A., J., Easwaran, B., & R., T. (2021). Smart Connected Digital Products and IoT Platform With the Digital Twin. In Vasant, P., Weber, G., & Punurai, W. (Ed.), Research Advancements in Smart Technology, Optimization, and Renewable Energy (pp. 330-350). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3970-5.ch016

Bajaj, A., Sharma, T., & Sangwan, O. P. (2020). Information Retrieval in Conjunction With Deep Learning. In A. Solanki, S. Kumar, & A. Nayyar (Eds.), *Handbook of Research on Emerging Trends and Applications of Machine Learning* (pp. 300–311). IGI Global., doi:10.4018/978-1-5225-9643-1.ch014

Bakri Hassan, M., Sayed Ali Ahmed, E., & Saeed, R. A. (2021). Machine Learning for Industrial IoT Systems. In J. Zhao & V. Kumar (Eds.), *Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies* (pp. 336–358). IGI Global., doi:10.4018/978-1-7998-6870-5.ch023

Balios, D. (2021). The Impact of Big Data on Accounting and Auditing. [IJCFA]. *International Journal of Corporate Finance and Accounting*, *8*(1), 1–14. doi:10.4018/IJCFA.2021010101

Banik, D., & Bhattacharjee, D. (2021). Mitigating Data Imbalance Issues in Medical Image Analysis. In D. P. Rana & R. G. Mehta (Eds.), *Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance* (pp. 66–89). IGI Global., doi:10.4018/978-1-7998-7371-6.ch004

Bansal, A., Ahirwar, M. K., & Shukla, P. K. (2019). Assessment on Different Classification Algorithms Used in Internet of Things Applications. [IJOCI]. *International Journal of Organizational and Collective Intelligence*, *9*(1), 1–11. doi:10.4018/IJOCI.2019010101

Bansal, A., Shukla, P. K., & Ahirwar, M. K. (2019). Opinion on Different Classification Algorithms Used in Internet of Things Environment for Large Data Set. [IJOCI]. *International Journal of Organizational and Collective Intelligence*, *9*(1), 51–60. doi:10.4018/IJOCI.2019010104

Bansal, R., & Singh, V. K. (2020). Proposed Technique for Efficient Cloud Computing Model in Effective Digital Training Towards Sustainable Livelihoods for Unemployed Youths. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(4), 13–27. doi:10.4018/IJCAC.2020100102

Barrett, D. (2021). Forensic Acquisition Methods for Cloud Computing Environments. In Khosrow-Pour D.B.A., M. (Eds.), Encyclopedia of Information Science and Technology, Fifth Edition (pp. 462-472). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3479-3.ch033

Bashorun, M. T., Omopupa, K. T., & Dahiru, G. (2020). Cloud Computing and Academic Libraries in Nigeria. In A. Tella (Ed.), *Handbook of Research on Digital Devices for Inclusivity and Engagement in Libraries* (pp. 113–134). IGI Global., doi:10.4018/978-1-5225-9034-7.ch006

Baumeister, F., Barbosa, M. W., & Gomes, R. R. (2020). What Is Required to Be a Data Scientist? [IJHCITP]. *International Journal of Human Capital and Information Technology Professionals*, *11*(4), 21–40. doi:10.4018/IJHCITP.2020100102

Bedi, P., Goyal, S. B., Kumar, J., & Kumar, S. (2021). Application of Image Processing for Autism Spectrum Disorder. In S. Kautish & G. Dhiman (Eds.), *Artificial Intelligence for Accurate Analysis and Detection of Autism Spectrum Disorder* (pp. 1–24). IGI Global., doi:10.4018/978-1-7998-7460-7.ch001

Benadda, M., & Belalem, G. (2020). Improving Road Safety for Driver Malaise and Sleepiness Behind the Wheel Using Vehicular Cloud Computing and Body Area Networks. [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, *12*(4), 19–41. doi:10.4018/IJSSCI.2020100102

Benhammouda, M., & Malki, M. (2019). A GPU Based Approach for Solving the Workflow Scheduling Problem. [IJIRR]. *International Journal of Information Retrieval Research*, *9*(4), 1–12. doi:10.4018/IJIRR.2019100101

Berger, J. I., & Kungu, K. (2019). Cultural Markers and Their Impact on Teaching in Higher Education. In J. Keengwe & K. Kungu (Eds.), *Handbook of Research on Cross-Cultural Online Learning in Higher Education* (pp. 22–41). IGI Global., doi:10.4018/978-1-5225-8286-1.ch002

Bhajantri, L. B., & Baluragi, P. M. (2020). Context Aware Data Perception in Cognitive Internet of Things - Cognitive Agent Approach. [IJHIoT]. *International Journal of Hyperconnectivity and the Internet of Things*, *4*(2), 1–24. doi:10.4018/IJHIoT.2020070101

Bhandari, G. P., & Gupta, R. (2020). Fault Prediction in SOA-Based Systems Using Deep Learning Techniques. [IJWSR]. *International Journal of Web Services Research*, *17*(3), 1–19. doi:10.4018/IJWSR.2020070101

Bhargavi, K. (2021). Deep Learning Architectures and Tools. In K. Senthilnathan, B. Shanmugam, D. Goyal, I. Annapoorani, & R. Samikannu (Eds.), *Deep Learning Applications and Intelligent Decision Making in Engineering* (pp. 55–75). IGI Global., doi:10.4018/978-1-7998-2108-3.ch002

Bhavsar, S. A., Pandit, B. Y., & Modi, K. J. (2019). Social Internet of Things. In D. J. Mala (Ed.), *Integrating the Internet of Things Into Software Engineering Practices* (pp. 199–218). IGI Global., doi:10.4018/978-1-5225-7790-4.ch010

Boobalan, M. P. (2019). Deep Clustering. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 164–179). IGI Global., doi:10.4018/978-1-5225-7862-8.ch010

Boufenara, M. N., Boufaida, M., & Berkane, M. L. (2020). A Prediction Approach Based on Self-Training and Deep Learning for Biological Data. [IJOCI]. *International Journal of Organizational and Collective Intelligence*, *10*(4), 50–64. doi:10.4018/IJOCI.2020100104

C., C., & V., S. (2021). Big Data IoT Analytics for Smart Cities With Cloud Computing Technique. In Velayutham, S. (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 104-126). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch007

Cai, X. (2020). Agricultural Environment Information Management. [IJAEIS]. *International Journal of Agricultural and Environmental Information Systems*, *11*(3), 48–60. doi:10.4018/IJAEIS.2020070104

Castaneda, G., Morris, P., & Khoshgoftaar, T. M. (2019). Maxout Networks for Visual Recognition. [IJMDEM]. *International Journal of Multimedia Data Engineering and Management*, *10*(4), 1–25. doi:10.4018/IJMDEM.2019100101

Chahal, H., Bhasin, A., & Kaveri, P. R. (2019). QoS Based Efficient Resource Allocation and Scheduling in Cloud Computing. [IJTHI]. *International Journal of Technology and Human Interaction*, *15*(4), 13–29. doi:10.4018/IJTHI.2019100102

Chakraborty, S., & Mali, K. (2020). An Overview of Biomedical Image Analysis From the Deep Learning Perspective. In Chakraborty, S., & Mali, K. (Ed.), Applications of Advanced Machine Intelligence in Computer Vision and Object Recognition (pp. 197-218). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2736-8.ch008

Chander, B. (2020). Deep Learning Network. In S., S., & M., J. (Eds.), Neural Networks for Natural Language Processing (pp. 1-30). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-1159-6.ch001

Chandrakar, M. K., & Mishra, A. (2020). Brain Tumor Detection Using Multipath Convolution Neural Network (CNN). [IJCVIP]. *International Journal of Computer Vision and Image Processing*, *10*(4), 43–53. doi:10.4018/IJCVIP.2020100103

Chang, Y., Hsu, P., Shiau, W., & Hsu, M. (2019). An Empirical Study on Factors Affecting Switching Intention to Cloud Enterprise Resource Planning. [JGIM]. *Journal of Global Information Management*, *27*(4), 46–69. doi:10.4018/JGIM.2019100103

Chaubey, N. K., & Jayanthi, P. (2020). Disease Diagnosis and Treatment Using Deep Learning Algorithms for the Healthcare System. In R. Wason, D. Goyal, V. Jain, S. Balamurugan, & A. Baliyan (Eds.), *Applications of Deep Learning and Big IoT on Personalized Healthcare Services* (pp. 99–114). IGI Global., doi:10.4018/978-1-7998-2101-4.ch007

Chaudhary, A., Chouhan, K. S., Gajrani, J., & Sharma, B. (2020). Deep Learning With PyTorch. In M. Mahrishi, K. K. Hiran, G. Meena, & P. Sharma (Eds.), *Machine Learning and Deep Learning in Real-Time Applications* (pp. 61–95). IGI Global., doi:10.4018/978-1-7998-3095-5.ch003

Chauhan, D., & Jain, J. K. (2021). IoT and Blockchain in Indian Perspective. In H. Patel & G. S. Thakur (Eds.), *Blockchain Applications in IoT Security* (pp. 186–202). IGI Global., doi:10.4018/978-1-7998-2414-5.ch011

Chauhan, P., Mandoria, H. L., Negi, A., & Rajput, R. S. (2021). Plant Diseases Concept in Smart Agriculture Using Deep Learning. In A. K. Gupta, D. Goyal, V. Singh, & H. Sharma (Eds.), *Smart Agricultural Services Using Deep Learning, Big Data, and IoT* (pp. 139–153). IGI Global., doi:10.4018/978-1-7998-5003-8.ch008

Chawla, N., Kumar, D., & Sharma, D. K. (2020). Improving Cost for Data Migration in Cloud Computing Using Genetic Algorithm. [IJSI]. *International Journal of Software Innovation*, *8*(3), 69–81. doi:10.4018/IJSI.2020070105

Chen, Y., Duan, L., & Zhang, W. (2020). Effect of User Involvement in Supply Chain Cloud Innovation. [JGIM]. *Journal of Global Information Management*, *28*(1), 23–38. doi:10.4018/JGIM.2020010102

Chiappetta, A. (2019). Survey on Industrial Internet of Things (IoT) Threats and Security. In B. Gupta & D. P. Agrawal (Eds.), *Handbook of Research on Cloud Computing and Big Data Applications in IoT* (pp. 330–366). IGI Global., doi:10.4018/978-1-5225-8407-0.ch016

Chitra, P., & Abirami, S. (2020). Leveraging Fog Computing and Deep Learning for Building a Secure Individual Health-Based Decision Support System to Evade Air Pollution. In R. C. Joshi & B. B. Gupta (Eds.), *Security, Privacy, and Forensics Issues in Big Data* (pp. 380–406). IGI Global., doi:10.4018/978-1-5225-9742-1.ch017

Chowdhury, A., Karmakar, G., & Kamruzzaman, J. (2019). The Co-Evolution of Cloud and IoT Applications. In S. Singh & R. Mohan Sharma (Eds.), *Handbook of Research on the IoT, Cloud Computing, and Wireless Network Optimization* (pp. 213–234). IGI Global., doi:10.4018/978-1-5225-7335-7.ch011

Chowdhury, M. M., & Amin, S. H. (2021). Forecasting Sales and Return Products for Retail Corporations and Bridging Among Them. In A. Taghipour (Ed.), *Demand Forecasting and Order Planning in Supply Chains and Humanitarian Logistics* (pp. 250–281). IGI Global., doi:10.4018/978-1-7998-3805-0.ch009

Correia, E. (2021). Systems, Services, Solutions of the Public Cloud. In Khosrow-Pour D.B.A., M. (Eds.), Encyclopedia of Organizational Knowledge, Administration, and Technology (pp. 644-652). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3473-1.ch047

Dad, D., & Belalem, G. (2020). Efficient Strategies of VMs Scheduling Based on Physicals Resources and Temperature Thresholds. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(3), 81–95. doi:10.4018/IJCAC.2020070105

Daglarli, E. (2021). Explainable Artificial Intelligence (xAI) Approaches and Deep Meta-Learning Models for Cyber-Physical Systems. In A. K. Luhach & A. Elçi (Eds.), *Artificial Intelligence Paradigms for Smart Cyber-Physical Systems* (pp. 42–67). IGI Global., doi:10.4018/978-1-7998-5101-1.ch003

Dar, G. M., Sharma, A., & Singh, P. (2021). Deep Learning Models for Detection and Diagnosis of Alzheimer's Disease. In M. Roy & L. R. Gupta (Eds.), *Machine Learning and Data Analytics for Predicting, Managing, and Monitoring Disease* (pp. 140–149). IGI Global., doi:10.4018/978-1-7998-7188-0.ch011

Das, A., & Mohanty, M. N. (2021). An Useful Review on Optical Character Recognition for Smart Era Generation. In A. K. Tyagi (Ed.), *Multimedia and Sensory Input for Augmented, Mixed, and Virtual Reality* (pp. 1–41). IGI Global., doi:10.4018/978-1-7998-4703-8.ch001

Das, M. S., Govardhan, A., & Lakshmi, D. V. (2019). Cost Minimization Through Load Balancing and Effective Resource Utilization in Cloud-Based Web Services. [IJNCR]. *International Journal of Natural Computing Research*, *8*(2), 51–74. doi:10.4018/IJNCR.2019040103

Dash, A. K., & Mohapatra, P. (2021). A Survey on Prematurity Detection of Diabetic Retinopathy Based on Fundus Images Using Deep Learning Techniques. In S. Saxena & S. Paul (Eds.), *Deep Learning Applications in Medical Imaging* (pp. 140–155). IGI Global., doi:10.4018/978-1-7998-5071-7.ch006

Dede, G., Fragiadakis, G., Michalakelis, C., & Kamalakis, T. (2019). Brokering Intelligence as a Service for the Internet of Things. [IJTD]. *International Journal of Technology Diffusion*, *10*(2), 18–33. doi:10.4018/IJTD.2019040102

Demir, A. K., & Alam, S. (2021). Advancing Artificial Intelligence-Enabled Cybersecurity for the Internet of Things. In K. Sandhu (Ed.), *Handbook of Research on Advancing Cybersecurity for Digital Transformation* (pp. 118–143). IGI Global., doi:10.4018/978-1-7998-6975-7.ch007

Demirci, S., Şahin, D. Ö., & Toprak, I. H. (2021). Android-Based Skin Cancer Recognition System Using Convolutional Neural Network. In D. Yadav, A. Bansal, M. Bhatia, M. Hooda, & J. Morato (Eds.), *Diagnostic Applications of Health Intelligence and Surveillance Systems* (pp. 59–85). IGI Global., doi:10.4018/978-1-7998-6527-8.ch003

Devare, M. H. (2019). Cloud Computing and Innovations. In G. Kecskemeti (Ed.), *Applying Integration Techniques and Methods in Distributed Systems and Technologies* (pp. 1–33). IGI Global., doi:10.4018/978-1-5225-8295-3.ch001

Dhamdhere, S. N., & Mane, D. (2021). Real-Time Recommendation Engine for Readers. In S. N. Dhamdhere (Ed.), *Big Data Applications for Improving Library Services* (pp. 165–177). IGI Global., doi:10.4018/978-1-7998-3049-8.ch011

Dhanapal, A. C., Sylvia Subapriya, M., Subramaniam, K., & Appukutty, M. (2022). Virtual Digital Twins. In D. Jeya Mala (Ed.), *Integrating AI in IoT Analytics on the Cloud for Healthcare Applications* (pp. 1–23). IGI Global., doi:10.4018/978-1-7998-9132-1.ch001

Dhawale, C. A., & Dhawale, K. (2020). Current Trends in Deep Learning Frameworks With Opportunities and Future Prospectus. In S., S., & M., J. (Ed.), Neural Networks for Natural Language Processing (pp. 63-77). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-1159-6.ch003

Dhawale, C. A., Dhawale, K., & Dubey, R. (2020). A Review on Deep Learning Applications. In J. J. Thomas, P. Karagoz, B. B. Ahamed, & P. Vasant (Eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics* (pp. 21–31). IGI Global., doi:10.4018/978-1-7998-1192-3.ch002

Dhaya, R., & Kanthavel, R. (2022). Futuristic Research Perspectives of IoT Platforms. In D. Jeya Mala (Ed.), *Integrating AI in IoT Analytics on the Cloud for Healthcare Applications* (pp. 258–275). IGI Global., doi:10.4018/978-1-7998-9132-1.ch015

Dhupia, B., & Rani, M. U. (2022). Assessment of Electric Consumption Forecast Using Machine Learning and Deep Learning Models for the Industrial Sector. In P. Krishna (Ed.), *Handbook of Research on Advances in Data Analytics and Complex Communication Networks* (pp. 206–218). IGI Global., doi:10.4018/978-1-7998-7685-4.ch016

Dif, N., & Elberrichi, Z. (2020). A New Deep Learning Model Selection Method for Colorectal Cancer Classification. [IJSIR]. *International Journal of Swarm Intelligence Research*, *11*(3), 72–88. doi:10.4018/IJSIR.2020070105

Dif, N., & Elberrichi, Z. (2020). A New Intra Fine-Tuning Method Between Histopathological Datasets in Deep Learning. [IJSSMET]. *International Journal of Service Science, Management, Engineering, and Technology*, *11*(2), 16–40. doi:10.4018/IJSSMET.2020040102

Dif, N., & Elberrichi, Z. (2020). Efficient Regularization Framework for Histopathological Image Classification Using Convolutional Neural Networks. [IJCINI]. *International Journal of Cognitive Informatics and Natural Intelligence*, *14*(4), 62–81. doi:10.4018/IJCINI.2020100104

Dikovic, L. (2021). Internet of Things in Healthcare as an Innovative Form of Personalized Medicine. In Khosrow-Pour D.B.A., M. (Eds.), Encyclopedia of Information Science and Technology, Fifth Edition (pp. 1933-1943). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3479-3.ch134

Dinc, E., Kuscu, M., Bilgin, B. A., & Akan, O. B. (2019). Internet of Everything. In P. J. Cardoso, J. Monteiro, J. Semião, & J. M. Rodrigues (Eds.), *Harnessing the Internet of Everything (IoE) for Accelerated Innovation Opportunities* (pp. 1–30). IGI Global., doi:10.4018/978-1-5225-7332-6.ch001

Durga, S., Mohan, S., Dinesh Peter, J., & Nittala, M. R. (2019). Towards Benefiting Both Cloud Users and Service Providers Through Resource Provisioning. [IJITSA]. *International Journal of Information Technologies and Systems Approach*, *12*(1), 37–51. doi:10.4018/IJITSA.2019010103

Dutta, P., Muppalaneni, N. B., & Patgiri, R. (2022). A Survey on Explainability in Artificial Intelligence. In P. Krishna (Ed.), *Handbook of Research on Advances in Data Analytics and Complex Communication Networks* (pp. 55–75). IGI Global., doi:10.4018/978-1-7998-7685-4.ch004

Edu, S. A., Agoyi, M., & Agozie, D. Q. (2020). Integrating Digital Innovation Capabilities Towards Value Creation. [IJIIT]. *International Journal of Intelligent Information Technologies*, *16*(4), 37–50. doi:10.4018/IJIIT.2020100103

El-Attar, N. E., El-Ela, N. A., & Awad, W. A. (2019). Integrated Learning Approaches Based on Cloud Computing for Personalizing e-Learning Environment. [IJWLTT]. *International Journal of Web-Based Learning and Teaching Technologies*, *14*(2), 67–87. doi:10.4018/IJWLTT.2019040105

Elangovan, R., & Prianga, M. (2019). Side Channel Attacks in Cloud Computing. In A. Haldorai & A. Ramu (Eds.), *Cognitive Social Mining Applications in Data Analytics and Forensics* (pp. 77–98). IGI Global., doi:10.4018/978-1-5225-7522-1.ch005

Elhosuieny, A., Salem, M., Thabet, A., & Ibrahim, A. (2019). ADOMC-NPR Automatic Decision-Making Offloading Framework for Mobile Computation Using Nonlinear Polynomial Regression Model. [IJWSR]. *International Journal of Web Services Research*, *16*(4), 53–73. doi:10.4018/IJWSR.2019100104

Elleuch, M., & Kherallah, M. (2019). Boosting of Deep Convolutional Architectures for Arabic Handwriting Recognition. [IJMDEM]. *International Journal of Multimedia Data Engineering and Management*, *10*(4), 26–45. doi:10.4018/IJMDEM.2019100102

Enshaei, N., Hammad, A., & Naderkhani, F. (2020). A Comprehensive Review on Advanced Maintenance Strategies for Smart Railways. In A. Awasthi & K. Grzybowska (Eds.), *Handbook of Research on Interdisciplinary Approaches to Decision Making for Sustainable Supply Chains* (pp. 433–457). IGI Global., doi:10.4018/978-1-5225-9570-0.ch020

Fantinato, M., Peres, S. M., Kafeza, E., Chiu, D. K., & Hung, P. C. (2021). A Review on the Integration of Deep Learning and Service-Oriented Architecture. [JDM]. *Journal of Database Management*, *32*(3), 95–119. doi:10.4018/JDM.2021070105

Fati, S. M., Jaradat, A. K., Abunadi, I., & Mohammed, A. S. (2020). Modelling Virtual Machine Workload in Heterogeneous Cloud Computing Platforms. [JITR]. *Journal of Information Technology Research*, *13*(4), 156–170. doi:10.4018/JITR.20201001.oa1

Fazzin, S. (2020). *Artificial Intelligence in Practice*. IGI Global., doi:10.4018/978-1-7998-2036-9.ch005

Fazzin, S. (2020). *The Fourth Industrial Revolution and the Internet of Things*. IGI Global., doi:10.4018/978-1-7998-2036-9.ch001

Feng, D., Wu, Z., Zuo, D., & Zhang, Z. (2020). Auto-Scaling Provision Basing on Workload Prediction in the Virtualized Data Center. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *12*(1), 53–69. doi:10.4018/IJGHPC.2020010104

França, R. P., Iano, Y., Monteiro, A. C., & Arthur, R. (2020). A Proposal of Improvement for Transmission Channels in Cloud Environments Using the CBEDE Methodology. In B. B. Gupta (Ed.), *Modern Principles, Practices, and Algorithms for Cloud Security* (pp. 184–202). IGI Global., doi:10.4018/978-1-7998-1082-7.ch009

França, R. P., Monteiro, A. C., Arthur, R., & Iano, Y. (2021). An Overview of Narrowband Internet of Things (NB-IoT) in the Modern Era. In S. K. Routray & S. Mohanty (Eds.), *Principles and Applications of Narrowband Internet of Things (NBIoT)* (pp. 26–45). IGI Global., doi:10.4018/978-1-7998-4775-5.ch002

G., S. S., L., M., & Patil, R. (2021). Real-Time Problems to Be Solved by the Combination of IoT, Big Data, and Cloud Technologies. In Velayutham, S. (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 265-276). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch015

Gaba, P., & Raw, R. S. (2020). Vehicular Cloud and Fog Computing Architecture, Applications, Services, and Challenges. In R. S. Rao, V. Jain, O. Kaiwartya, & N. Singh (Eds.), *IoT and Cloud Computing Advancements in Vehicular Ad-Hoc Networks* (pp. 268–296). IGI Global., doi:10.4018/978-1-7998-2570-8.ch014

Gao, T., & Wang, G. Y. (2020). Brain Signal Classification Based on Deep CNN. [IJSPPC]. *International Journal of Security and Privacy in Pervasive Computing*, *12*(2), 17–29. doi:10.4018/IJSPPC.2020040102

Garcia-Santa, N., San Miguel, B., & Ugai, T. (2019). Converging Semantic Knowledge and Deep Learning for Medical Coding. [IJPHIM]. *International Journal of Privacy and Health Information Management*, *7*(2), 33–52. doi:10.4018/IJPHIM.2019070103

Garozzo, R., Pino, C., Santagati, C., & Spampinato, C. (2020). Harnessing the Power of Artificial Intelligence for Modelling and Understanding Cultural Heritage Data. In C. M. Bolognesi & C. Santagati (Eds.), *Impact of Industry 4.0 on Architecture and Cultural Heritage* (pp. 357–376). IGI Global., doi:10.4018/978-1-7998-1234-0.ch015

Gavrovska, A., & Samčović, A. (2020). Intelligent Automation Using Machine and Deep Learning in Cybersecurity of Industrial IoT. In M. D. Stojanović & S. V. Boštjančič Rakas (Eds.), *Cyber Security of Industrial Control Systems in the Future Internet Environment* (pp. 156–174). IGI Global., doi:10.4018/978-1-7998-2910-2.ch008

Gayathri, T. N., & Rajasekharababu, M. (2022). Privacy Preserving in Smart Cities Using Various Computing Technologies. In P. Krishna (Ed.), *Handbook of Research on Advances in Data Analytics and Complex Communication Networks* (pp. 47–54). IGI Global., doi:10.4018/978-1-7998-7685-4.ch003

Ghavifekr, S., & Wong, S. Y. (2022). Role of Big Data in Education. In P. Ordóñez de Pablos, X. Zhang, M. N. Almunawar, & J. E. Gayo (Eds.), *Handbook of Research on Big Data, Green Growth, and Technology Disruption in Asian Companies and Societies* (pp. 22–37). IGI Global., doi:10.4018/978-1-7998-8524-5.ch002

Gill, H. K., Verma, A. K., & Sandhu, R. (2019). An Adaptive Security Framework for the Internet of Things Applications Based on the Contextual Information. In M. T. Banday (Ed.), *Cryptographic Security Solutions for the Internet of Things* (pp. 244–267). IGI Global., doi:10.4018/978-1-5225-5742-5.ch009

Gill, S. S., & Shaghaghi, A. (2020). Security-Aware Autonomic Allocation of Cloud Resources. [JOEUC]. *Journal of Organizational and End User Computing*, *32*(3), 15–22. doi:10.4018/JOEUC.2020070102

Gite, S., & Agrawal, H. (2019). Early Prediction of Driver's Action Using Deep Neural Networks. [IJIRR]. *International Journal of Information Retrieval Research*, *9*(2), 11–27. doi:10.4018/IJIRR.2019040102

Gonzalez-Usach, R., Yacchirema, D., Julian, M., & Palau, C. E. (2019). Interoperability in IoT. In Kaur, G., & Tomar, P. (Ed.), Handbook of Research on Big Data and the IoT (pp. 149-173). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7432-3.ch009

Goundar, S., Bhardwaj, A., Singh, S., Singh, M., & L., G. H. (2021). Big Data and Big Data Analytics. In Goundar, S., & Rayani, P. K. (Ed.), *Applications of Big Data in Large- and Small-Scale Systems* (pp. 1-19). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-6673-2.ch001

Goyal, S., & Saxena, A. (2021). Creditworthiness Assessment Using Natural Language Processing. In P. Tanwar, A. Saxena, & C. Priya (Eds.), *Deep Natural Language Processing and AI Applications for Industry 5.0* (pp. 120–141). IGI Global., doi:10.4018/978-1-7998-7728-8.ch007

Greene, M., Clarke-Hagan, D., & Curran, M. (2020). Achieving Smarter Buildings and More Efficient Facilities Management. [IJDIBE]. *International Journal of Digital Innovation in the Built Environment*, *9*(2), 1–16. doi:10.4018/IJDIBE.2020070101

Gulzar, Z., Raj, L. A., & Leema, A. A. (2019). Ontology Supported Hybrid Recommender System With Threshold Based Nearest Neighbourhood Approach. [IJICTE]. *International Journal of Information and Communication Technology Education*, *15*(2), 85–107. doi:10.4018/IJICTE.2019040106

Güngör, D. Ö. (2019). Industry 4.0 Technologies Used in Project Management. In H. Bolat & G. T. Temur (Eds.), *Agile Approaches for Successfully Managing and Executing Projects in the Fourth Industrial Revolution* (pp. 40–63). IGI Global., doi:10.4018/978-1-5225-7865-9.ch003

Gunjan, V. K., Pathak, R., & Singh, O. (2019). Understanding Image Classification Using TensorFlow Deep Learning - Convolution Neural Network. [IJHIoT]. *International Journal of Hyperconnectivity and the Internet of Things*, *3*(2), 19–37. doi:10.4018/IJHIoT.2019070103

Gupta, M., Choudhary, A., & Parmar, J. (2021). Analysis of Text Identification Techniques Using Scene Text and Optical Character Recognition. [IJCVIP]. *International Journal of Computer Vision and Image Processing*, *11*(4), 39–62. doi:10.4018/IJCVIP.2021100104

Hammouri, Q., & Abu-Shanab, E. A. (2020). Major Factors Influencing the Adoption of Cloud Computing in Jordan. [IJTHI]. *International Journal of Technology and Human Interaction*, *16*(4), 55–69. doi:10.4018/IJTHI.2020100104

Hanafizadeh, P., & Zareravasan, A. (2020). A Systematic Literature Review on IT Outsourcing Decision and Future Research Directions. [JGIM]. *Journal of Global Information Management*, *28*(2), 160–201. doi:10.4018/JGIM.2020040108

Haque, M. A., Haque, S., Kumar, K., & Singh, N. K. (2021). A Comprehensive Study of Cyber Security Attacks, Classification, and Countermeasures in the Internet of Things. In P. F. Anunciação, C. M. Pessoa, & G. L. Jamil (Eds.), *Handbook of Research on Digital Transformation and Challenges to Data Security and Privacy* (pp. 63–90). IGI Global., doi:10.4018/978-1-7998-4201-9.ch004

Harrath, Y., & Bahlool, R. (2019). Multi-Objective Genetic Algorithm for Tasks Allocation in Cloud Computing. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(3), 37–57. doi:10.4018/IJCAC.2019070103

Hasib, K. M., Towhid, N. A., & Islam, M. R. (2021). HSDLM. [IJCAC]. *International Journal of Cloud Applications and Computing*, *11*(4), 1–13. doi:10.4018/IJCAC.2021100101

Hawamdeh, M., & Adamu, I. (2021). Problem-Based Learning (PBL). In G. Kurubacak-Meric & S. Sisman-Ugur (Eds.), *Improving Scientific Communication for Lifelong Learners* (pp. 38–56). IGI Global., doi:10.4018/978-1-7998-4534-8.ch003

Herzfeldt, A., Floerecke, S., Ertl, C., & Krcmar, H. (2019). Examining the Antecedents of Cloud Service Profitability. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(4), 37–65. doi:10.4018/IJCAC.2019100103

Herzfeldt, A. B., Rauer, H. P., Weißbach, R., & Ertl, C. (2020). Cloud Computing as the Next Utility. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(4), 28–47. doi:10.4018/IJCAC.2020100103

Hossain, K., Rahman, M., & Roy, S. (2019). IoT Data Compression and Optimization Techniques in Cloud Storage. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(2), 43–59. doi:10.4018/IJCAC.2019040103

Huda, M., Habib, M. I., Khan, M. Z., & Aziz, A. A. (2021). Critical Comprehensive Study of Big Data Analytics Management. In B. A. Khan, M. H. Kuofie, & S. Suman (Eds.), *Handbook of Research on Future Opportunities for Technology Management Education* (pp. 235–247). IGI Global., doi:10.4018/978-1-7998-8327-2.ch014

Hudaib, A., & Albdour, L. (2019). Fog Computing to Serve the Internet of Things Applications. [IJFC]. *International Journal of Fog Computing*, *2*(2), 44–56. doi:10.4018/IJFC.2019070103

Huseynov, F. (2021). Big Data in Business. In K. Sandhu (Ed.), *Disruptive Technology and Digital Transformation for Business and Government* (pp. 235–249). IGI Global., doi:10.4018/978-1-7998-8583-2.ch012

Husna, A., Amin, S. H., & Shah, B. (2021). Demand Forecasting in Supply Chain Management Using Different Deep Learning Methods. In A. Taghipour (Ed.), *Demand Forecasting and Order Planning in Supply Chains and Humanitarian Logistics* (pp. 140–170). IGI Global., doi:10.4018/978-1-7998-3805-0.ch005

Hussain, M. M., & Beg, M. S. (2019). Using Vehicles as Fog Infrastructures for Transportation Cyber-Physical Systems (T-CPS). [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, *11*(1), 47–69. doi:10.4018/IJSSCI.2019010104

Ibrahim, N. M., & Zainal, A. (2019). An Adaptive Intrusion Detection Scheme for Cloud Computing. [IJSIR]. *International Journal of Swarm Intelligence Research*, *10*(4), 53–70. doi:10.4018/IJSIR.2019100104

Ibrahim, N. M., & Zainal, A. (2020). A Distributed Intrusion Detection Scheme for Cloud Computing. [IJDST]. *International Journal of Distributed Systems and Technologies*, *11*(1), 68–82. doi:10.4018/IJDST.2020010106

Ilmudeen, A. (2021). Big Data, Artificial Intelligence, and the Internet of Things in Cross-Border E-Commerce. In M. R. Hoque & R. E. Bashaw (Eds.), *Cross-Border E-Commerce Marketing and Management* (pp. 257–272). IGI Global., doi:10.4018/978-1-7998-5823-2.ch011

Imran, M., Iqbal, J., & Saleem, H. M. (2020). The Role of Internet of Things, Knowledge Management, and Open Innovation in SME Sustainability. In P. Ordoñez de Pablos, X. Zhang, & K. Chui (Eds.), *Innovative Management and Business Practices in Asia* (pp. 295–303). IGI Global., doi:10.4018/978-1-7998-1566-2.ch015

Iqbal, S., Ahmad, S., Bano, B., Akkour, K., Alghamdi, M. A., & Alothri, A. M. (2021). A Systematic Review. [IJIIT]. *International Journal of Intelligent Information Technologies*, *17*(1), 1–18. doi:10.4018/IJIIT.2021010101

Iqbal, S., Iqbal, U., & Hassan, S. A. (2019). Intelligent Tracking and Positioning of Targets Using Passive Sensing Systems. In I. Comşa & R. Trestian (Eds.), *Next-Generation Wireless Networks Meet Advanced Machine Learning Applications* (pp. 286–305). IGI Global., doi:10.4018/978-1-5225-7458-3.ch012

Issac, T., Silas, S., & Rajsingh, E. B. (2020). Modelling a Deep Learning-Based Wireless Sensor Network Task Assignment Algorithm. In Sagayam, K., Bhushan, B., Andrushia, A., & Albuquerque, V. C. (Ed.), Deep Learning Strategies for Security Enhancement in Wireless Sensor Networks (pp. 84-109). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-5068-7.ch005

Iyer, S. S., & Rajagopal, S. (2020). Applications of Machine Learning in Cyber Security Domain. In P. Ganapathi & D. Shanmugapriya (Eds.), *Handbook of Research on Machine and Deep Learning Applications for Cyber Security* (pp. 64–82). IGI Global., doi:10.4018/978-1-5225-9611-0.ch004

J., G., R. K., J., & N., S. (2021). On Swarm Intelligence and Its Integration With Internet of Things. In Bouarara, H. A. (Ed.), *Advanced Deep Learning Applications in Big Data Analytics* (pp. 156-181). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2791-7.ch009

J., M., & S., A. (2021). Data Security and Privacy-Preserving in Edge Computing. In Ambika, P., Donald, A., & Kumar, A. (Ed.), *Cases on Edge Computing and Analytics* (pp. 188-202). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4873-8.ch010

Jadad, H. A., Touzene, A., & Day, K. (2020). Offloading as a Service Middleware for Mobile Cloud Apps. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(2), 36–55. doi:10.4018/IJCAC.2020040103

Jadad, H. A., Touzene, A., Day, K., Alziedi, N., & Arafeh, B. (2019). Context-Aware Prediction Model for Offloading Mobile Application Tasks to Mobile Cloud Environments. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(3), 58–74. doi:10.4018/IJCAC.2019070104

Jain, S., Jain, S., & Jain, A. K. (2021). Deep Learning Approach Towards Unstructured Text Data Utilization. In A. Sharaff, G. R. Sinha, & S. Bhatia (Eds.), *New Opportunities for Sentiment Analysis and Information Processing* (pp. 29–49). IGI Global., doi:10.4018/978-1-7998-8061-5.ch002

Jakobczak, D. J., & Chatterjee, A. (2021). The Rise of "Big Data" in the Field of Cloud Analytics. In D. J. Jakóbczak (Ed.), *Analyzing Data Through Probabilistic Modeling in Statistics* (pp. 204–225). IGI Global., doi:10.4018/978-1-7998-4706-9.ch008

Jani, K. A., & Chaubey, N. (2020). IoT and Cyber Security. In N. K. Chaubey & B. B. Prajapati (Eds.), *Quantum Cryptography and the Future of Cyber Security* (pp. 203–235). IGI Global., doi:10.4018/978-1-7998-2253-0.ch010

Jasmine, K. S. (2019). A New Process Model for IoT-Based Software Engineering. In D. J. Mala (Ed.), *Integrating the Internet of Things Into Software Engineering Practices* (pp. 1–13). IGI Global., doi:10.4018/978-1-5225-7790-4.ch001

Javali, A. (2021). NB-IoT for Agriculture. In S. K. Routray & S. Mohanty (Eds.), *Principles and Applications of Narrowband Internet of Things (NBIoT)* (pp. 175–194). IGI Global., doi:10.4018/978-1-7998-4775-5.ch008

Jayswal, A. K. (2020). Hybrid Load-Balanced Scheduling in Scalable Cloud Environment. [IJISMD]. *International Journal of Information System Modeling and Design*, *11*(3), 62–78. doi:10.4018/IJISMD.2020070104

Ji, B., Zhang, W., Li, R., & Ji, H. (2019). Deep Learning Models for Biomedical Image Analysis. In C. Chen & S. S. Cheung (Eds.), *Computational Models for Biomedical Reasoning and Problem Solving* (pp. 128–148). IGI Global., doi:10.4018/978-1-5225-7467-5.ch005

Jindal, U., & Gupta, S. (2021). Deep Learning-Based Knowledge Extraction From Diseased and Healthy Edible Plant Leaves. [IJISMD]. *International Journal of Information System Modeling and Design*, *12*(2), 67–81. doi:10.4018/IJISMD.2021040105

Joshi, D., Anwarul, S., & Mishra, V. (2020). Deep Leaning Using Keras. In M. Mahrishi, K. K. Hiran, G. Meena, & P. Sharma (Eds.), *Machine Learning and Deep Learning in Real-Time Applications* (pp. 33–60). IGI Global., doi:10.4018/978-1-7998-3095-5.ch002

Jouini, M., & Ben Arfa Rabai, L. (2020). A Computational Approach for Secure Cloud Computing Environments. In B. B. Gupta (Ed.), *Modern Principles, Practices, and Algorithms for Cloud Security* (pp. 129–144). IGI Global., doi:10.4018/978-1-7998-1082-7.ch006

Junior, M. R., Batista, C. L., Marques, M. E., & Pessoa, C. R. (2019). Business Models Applicable to IoT. In Jamil, G. L., Jamil, L. C., Pessoa, C. R., & Silveira, W. (Ed.), Handbook of Research on Business Models in Modern Competitive Scenarios (pp. 21-42). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7265-7.ch002

K. DeviPriya, & Lingamgunta, S. (2020). Multi Factor Two-way Hash-Based Authentication in Cloud Computing. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(2), 56–76. doi:10.4018/IJCAC.2020040104

K., N., Menon, V. G., B., J., & Kumar, A. (2020). Deep Learning for Next-Generation Inventive Wireless Networks. In Solanki, A., Kumar, S., & Nayyar, A. (Ed.), *Handbook of Research on Emerging Trends and Applications of Machine Learning* (pp. 183-199). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-9643-1.ch010

Kameswara Rao, M., & Santhi, S. G. (2020). Internet of Things-Based Authentication Mechanism for E-Health Applications. In P. Pankajavalli & G. Karthick (Eds.), *Incorporating the Internet of Things in Healthcare Applications and Wearable Devices* (pp. 122–136). IGI Global., doi:10.4018/978-1-7998-1090-2.ch007

Kanani, P., & Padole, M. C. (2021). ECG Image Classification Using Deep Learning Approach. In G. Rani & P. K. Tiwari (Eds.), *Handbook of Research on Disease Prediction Through Data Analytics and Machine Learning* (pp. 343–357). IGI Global., doi:10.4018/978-1-7998-2742-9.ch016

Kante, M., & Ndayizigamiye, P. (2021). A Systematic Mapping of Studies on the Adoption of Internet of Things to Provide Healthcare Services in Developing Countries. In P. Ndayizigamiye, G. Barlow-Jones, R. Brink, S. Bvuma, R. Minty, & S. Mhlongo (Eds.), *Perspectives on ICT4D and Socio-Economic Growth Opportunities in Developing Countries* (pp. 99–126). IGI Global., doi:10.4018/978-1-7998-2983-6.ch004

Karegowda, A. G. G., D., & T. P., R. S. (2020). IOT Technology, Applications, and Challenges. In Krit, S., Bălaș, V. E., Elhoseny, M., Benlamri, R., & Bălaș, M. M. (Ed.), Sensor Network Methodologies for Smart Applications (pp. 65-113). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4381-8.ch004

Karegowda, A. G. G., D., & M., G. (2021). Deep Learning Solutions for Agricultural and Farming Activities. In Senthilnathan, K., Shanmugam, B., Goyal, D., Annapoorani, I., & Samikannu, R. (Ed.), Deep Learning Applications and Intelligent Decision Making in Engineering (pp. 256-287). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2108-3.ch011

Karim, S., & Soomro, T. R. (2020). What Is Cloud Computing? In S. Gochhait, D. T. Shou, & S. Fazalbhoy (Eds.), *Cloud Computing Applications and Techniques for E-Commerce* (pp. 1–27). IGI Global., doi:10.4018/978-1-7998-1294-4.ch001

Karthikeyan, P., & Sathiyamoorthy, E. (2020). An Adaptive Service Monitoring System in a Cloud Computing Environment. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *12*(2), 47–63. doi:10.4018/IJGHPC.2020040103

Kashyap, R. (2019). Miracles of Healthcare With Internet of Things. In J. J. Rodrigues, A. Gawanmeh, K. Saleem, & S. Parvin (Eds.), *Smart Devices, Applications, and Protocols for the IoT* (pp. 120–164). IGI Global., doi:10.4018/978-1-5225-7811-6.ch007

Kaur, A., Kaur, B., & Singh, D. (2018). Meta-Heuristics Based Load Balancing Optimization in Cloud Environment on Underflow and Overflow Conditions. [JITR]. *Journal of Information Technology Research*, *11*(4), 155–172. doi:10.4018/JITR.2018100110

Kaur, R., & Rani, R. (2020). Comparative Study on ASD Identification Using Machine and Deep Learning. In T. Wadhera & D. Kakkar (Eds.), *Interdisciplinary Approaches to Altering Neurodevelopmental Disorders* (pp. 250–270). IGI Global., doi:10.4018/978-1-7998-3069-6.ch015

Kaushik, S., & Gandhi, C. (2020). Capability Based Outsourced Data Access Control with Assured File Deletion and Efficient Revocation with Trust Factor in Cloud Computing. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(1), 64–84. doi:10.4018/IJCAC.2020010105

Kaushik, S., & Gandhi, C. (2021). Cloud Computing Technologies. In R. S. Rao, N. Singh, O. Kaiwartya, & S. Das (Eds.), *Cloud-Based Big Data Analytics in Vehicular Ad-Hoc Networks* (pp. 233–253). IGI Global., doi:10.4018/978-1-7998-2764-1.ch011

Kero, A., Khanna, A., Kumar, D., & Agarwal, A. (2019). An Adaptive Approach Towards Computation Offloading for Mobile Cloud Computing. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *14*(2), 52–73. doi:10.4018/IJITWE.2019040104

Kesavan, S. J., S., Y., S., & V., M. (2021). IoT Device Onboarding, Monitoring, and Management. In Velayutham, S. (Ed.), Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing (pp. 196-224). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch012

Khan, A., Jhanjhi, N. Z., Humayun, M., & Ahmad, M. (2020). The Role of IoT in Digital Governance. In V. Ponnusamy, K. Rafique, & N. Zaman (Eds.), *Employing Recent Technologies for Improved Digital Governance* (pp. 128–150). IGI Global., doi:10.4018/978-1-7998-1851-9.ch007

Khan, G., Sarkar, A., & Sengupta, S. (2019). HBSD. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *14*(3), 37–63. doi:10.4018/IJITWE.2019070103

Khan, U. A. (2020). Semantic Analysis of Videos for Tags Prediction and Segmentation. In P. Kumar, V. Ponnusamy, & V. Jain (Eds.), *Industrial Internet of Things and Cyber-Physical Systems* (pp. 296–307). IGI Global., doi:10.4018/978-1-7998-2803-7.ch014

Khasawneh, M. A. (2020). Impact of Cloud Computing on Green Supply Chain Management. In A. Awasthi & K. Grzybowska (Eds.), *Handbook of Research on Interdisciplinary Approaches to Decision Making for Sustainable Supply Chains* (pp. 476–490). IGI Global., doi:10.4018/978-1-5225-9570-0.ch022

Khekare, G., & Sheikh, S. (2021). Autonomous Navigation Using Deep Reinforcement Learning in ROS. [IJAIML]. *International Journal of Artificial Intelligence and Machine Learning*, *11*(2), 63–70. doi:10.4018/IJAIML.20210701.oa4

Khekare, G., Verma, P., Dhanre, U., Raut, S., & Yenurkar, G. (2021). Analysis of Internet of Things Based on Characteristics, Functionalities, and Challenges. [IJHIoT]. *International Journal of Hyperconnectivity and the Internet of Things*, *5*(1), 44–62. doi:10.4018/IJHIoT.2021010103

Kioussi, A., Doulamis, A., Karoglou, M., & Moropoulou, A. I. (2020). Cultural Intelligence-Investigation of Different Systems for Heritage Sustainable Preservation. [IJACDT]. *International Journal of Art, Culture, Design, and Technology*, *9*(2), 16–30. doi:10.4018/IJACDT.2020070102

Koç, E. (2020). How Can Industrial Internet of Things (IIoT) Improve Enterprise Productivity? In E. Koç (Ed.), *Internet of Things (IoT) Applications for Enterprise Productivity* (pp. 112–133). IGI Global., doi:10.4018/978-1-7998-3175-4.ch005

Kohli, A., Kohli, R., Singh, B., & Singh, J. (2020). Smart Plant Monitoring System Using IoT Technology. In R. Singh, A. Gehlot, V. Jain, & P. K. Malik (Eds.), *Handbook of Research on the Internet of Things Applications in Robotics and Automation* (pp. 318–366). IGI Global., doi:10.4018/978-1-5225-9574-8.ch016

Kömeçoğlu, Y., Muftuoglu, Z., Umay, C., Tasdelen, A., & Ozdemir, S. (2019). Importance of Data Management in Digital Media Literacy. In M. N. Yildiz, M. Fazal, M. Ahn, R. Feirsen, & S. Ozdemir (Eds.), *Handbook of Research on Media Literacy Research and Applications Across Disciplines* (pp. 235–247). IGI Global., doi:10.4018/978-1-5225-9261-7.ch015

Kondaveeti, H. K., Priyatham Brahma, G., & Vijaya Sahithi, D. (2021). Deep Learning Applications in Agriculture. In P. Tomar & G. Kaur (Eds.), *Artificial Intelligence and IoT-Based Technologies for Sustainable Farming and Smart Agriculture* (pp. 325–345). IGI Global., doi:10.4018/978-1-7998-1722-2.ch020

Krishnamoorthy, M., Ahamed, B. B., Suresh, S., & Alagappan, S. (2020). Deep Learning Techniques and Optimization Strategies in Big Data Analytics. In J. J. Thomas, P. Karagoz, B. B. Ahamed, & P. Vasant (Eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics* (pp. 142–153). IGI Global., doi:10.4018/978-1-7998-1192-3.ch009

Krishnamurthi, R., & Goyal, M. (2019). Enabling Technologies for IoT. In B. Gupta & D. P. Agrawal (Eds.), *Handbook of Research on Cloud Computing and Big Data Applications in IoT* (pp. 243–270). IGI Global., doi:10.4018/978-1-5225-8407-0.ch013

Kuchuk, H., Podorozhniak, A., Hlavcheva, D., & Yaloveha, V. (2020). Application of Deep Learning in the Processing of the Aerospace System's Multispectral Images. In T. Shmelova, Y. Sikirda, & A. Sterenharz (Eds.), *Handbook of Research on Artificial Intelligence Applications in the Aviation and Aerospace Industries* (pp. 134–147). IGI Global., doi:10.4018/978-1-7998-1415-3.ch005

Kumar, A., & Sharma, S. (2021). Demur and Routing Protocols With application in Underwater Wireless Sensor Networks for Smart City. In N. Goyal, L. Sapra, & J. K. Sandhu (Eds.), *Energy-Efficient Underwater Wireless Communications and Networking* (pp. 262–278). IGI Global., doi:10.4018/978-1-7998-3640-7.ch017

Kumar, A. B., & Rao, P. V. (2020). Energy Efficient, Resource-Aware, Prediction Based VM Provisioning Approach for Cloud Environment. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *11*(3), 22–41. doi:10.4018/IJACI.2020070102

Kumar, K. D., Venkata Rathnam, T., Venkata Ramana, R., Sudhakara, M., & Poluru, R. K. (2020). Towards the Integration of Blockchain and IoT for Security Challenges in IoT. In D. S. Rajput, R. S. Thakur, & S. M. Basha (Eds.), *Transforming Businesses With Bitcoin Mining and Blockchain Applications* (pp. 45–67). IGI Global., doi:10.4018/978-1-7998-0186-3.ch003

Kumar, P., Chouhan, L., & Songara, A. (2019). Cloud- and IoT-Powered Smart Connected Cities. In N. Dey & S. Tamane (Eds.), *Big Data Analytics for Smart and Connected Cities* (pp. 71–102). IGI Global., doi:10.4018/978-1-5225-6207-8.ch004

Kumar, R. (2021). IoT and Deep Learning for Livestock Management. In R. Raut & A. D. Mihovska (Eds.), *Examining the Impact of Deep Learning and IoT on Multi-Industry Applications* (pp. 80–96). IGI Global., doi:10.4018/978-1-7998-7511-6.ch006

Kumar, S., & Raza, Z. (2019). A Priority-Based Message Response Time Aware Job Scheduling Model for the Internet of Things (IoT). [IJCPS]. *International Journal of Cyber-Physical Systems*, *1*(1), 1–14. doi:10.4018/IJCPS.2019010101

Kumar, S., Yadav, A., & Sharma, D. K. (2020). Deep Learning and Computer Vision in Smart Agriculture. In N. Pradeep, S. Kautish, C. Nirmala, V. Goyal, & S. Abdellatif (Eds.), *Modern Techniques for Agricultural Disease Management and Crop Yield Prediction* (pp. 66–88). IGI Global., doi:10.4018/978-1-5225-9632-5.ch004

Kumar, U., Tripathi, E., Tripathi, S. P., & Gupta, K. K. (2019). Deep Learning for Healthcare Biometrics. In D. R. Kisku, P. Gupta, & J. K. Sing (Eds.), *Design and Implementation of Healthcare Biometric Systems* (pp. 73–108). IGI Global., doi:10.4018/978-1-5225-7525-2.ch004

Kumar, V., & Bhardwaj, A. (2020). Deploying Cloud-Based Healthcare Services. [IJSSMET]. *International Journal of Service Science, Management, Engineering, and Technology*, *11*(4), 87–100. doi:10.4018/IJSSMET.2020100106

Kumar, V., & Bhardwaj, A. (2020). Role of Cloud Computing in School Education. In L. A. Tomei & D. D. Carbonara (Eds.), *Handbook of Research on Diverse Teaching Strategies for the Technology-Rich Classroom* (pp. 98–108). IGI Global., doi:10.4018/978-1-7998-0238-9.ch008

Kumar, V., & Lalotra, G. S. (2021). Blockchain-Enabled Secure Internet of Things. In H. Patel & G. S. Thakur (Eds.), *Blockchain Applications in IoT Security* (pp. 45–55). IGI Global., doi:10.4018/978-1-7998-2414-5.ch003

Kuppusamy, P. (2019). Green Cloud Architecture to E-Learning Solutions. In Z. Ma & L. Yan (Eds.), *Emerging Technologies and Applications in Data Processing and Management* (pp. 358–384). IGI Global., doi:10.4018/978-1-5225-8446-9.ch016

L., A., & R., P. (2021). A Hybrid Computational Intelligence Algorithm to Transform Traditional IPC Into a Smart Camera. In Almunawar, M. N., Anshari Ali, M., & Ariff Lim, S. (Ed.), *Handbook of Research on Innovation and Development of E-Commerce and E-Business in ASEAN* (pp. 722-738). IGI Global. https://doi.org/doi:10.4018/978-1-7998-4984-1.ch035

Lacárcel, F. J., Polanco-Diges, L., & Debasa, F. (2021). A Better Understanding of Big Data and Marketing Analytics. In J. R. Saura (Ed.), *Advanced Digital Marketing Strategies in a Data-Driven Era* (pp. 1–15). IGI Global., doi:10.4018/978-1-7998-8003-5.ch001

Lal, A. M., Reddy, B. K., & D., A. (2020). Review on Various Machine Learning and Deep Learning Techniques for Prediction and Classification of Quotidian Datasets. In Voulodimos, A., & Doulamis, A. (Ed.), *Recent Advances in 3D Imaging, Modeling, and Reconstruction* (pp. 296-323). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-5294-9.ch014

Lal, P., & Bharadwaj, S. S. (2020). Understanding the Drivers of Cloud-Based Service Adoption and Their Impact on the Organizational Performance. [JGIM]. *Journal of Global Information Management*, *28*(1), 56–85. doi:10.4018/JGIM.2020010104

Lavanya, S., Susila, N., & Venkatachalam, K. (2019). Impact of Cloud of Clouds in Enterprises Applications. In P. Raj & S. Koteeswaran (Eds.), *Novel Practices and Trends in Grid and Cloud Computing* (pp. 21–33). IGI Global., doi:10.4018/978-1-5225-9023-1.ch002

Li, B., Sun, X., & Yu, S. (2019). Designing of Internet of Things Sensor Based Information Gateway Using SDN Concept. [IJDST]. *International Journal of Distributed Systems and Technologies*, *10*(1), 13–24. doi:10.4018/IJDST.2019010102

Li, L., Wang, Y., Jin, L., Zhang, X., & Qin, H. (2019). Two-Stage Adaptive Classification Cloud Workload Prediction Based on Neural Networks. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *11*(2), 1–23. doi:10.4018/IJGHPC.2019040101

Li, Y. Z., Zhang, S. P., Li, Y., & Wang, S. (2020). Research on Intrusion Detection Algorithm Based on Deep Learning and Semi-Supervised Clustering. [IJCRE]. *International Journal of Cyber Research and Education*, *2*(2), 38–60. doi:10.4018/IJCRE.2020070105

Liao, S., & Chiu, W. (2021). Investigating the Behaviors of Mobile Games and Online Streaming Users for Online Marketing Recommendations. [IJOM]. *International Journal of Online Marketing*, *11*(1), 39–61. doi:10.4018/IJOM.2021010103

Liu, Q., Yang, F., & Tang, X. (2021). A New Pedestrian Feature Description Method Named Neighborhood Descriptor of Oriented Gradients. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *16*(1), 23–55. doi:10.4018/IJITWE.2021010102

Loey, M., ElSawy, A., & Afify, M. (2020). Deep Learning in Plant Diseases Detection for Agricultural Crops. [IJSSMET]. *International Journal of Service Science, Management, Engineering, and Technology*, *11*(2), 41–58. doi:10.4018/IJSSMET.2020040103

Loey, M., Naman, M. R., & Zayed, H. H. (2020). A Survey on Blood Image Diseases Detection Using Deep Learning. [IJSSMET]. *International Journal of Service Science, Management, Engineering, and Technology*, *11*(3), 18–32. doi:10.4018/IJSSMET.2020070102

Lu, J., & Yan, W. Q. (2020). Comparative Evaluations of Human Behavior Recognition Using Deep Learning. In B. B. Gupta & D. Gupta (Eds.), *Handbook of Research on Multimedia Cyber Security* (pp. 176–189). IGI Global., doi:10.4018/978-1-7998-2701-6.ch009

Lu, Y. (2020). Deep Learning of Data Analytics in Healthcare. In J. Khuntia, X. Ning, & M. Tanniru (Eds.), *Theory and Practice of Business Intelligence in Healthcare* (pp. 151–165). IGI Global., doi:10.4018/978-1-7998-2310-0.ch007

Lu, Y., Chen, W., Wang, X., Ailsworth, Z., Tsui, M., Al-Ghaib, H., & Zimmerman, B. (2020). Deep Learning-Based Models for Porosity Measurement in Thermal Barrier Coating Images. [IJMDEM]. *International Journal of Multimedia Data Engineering and Management*, *11*(3), 20–35. doi:10.4018/IJMDEM.2020070102

Lukyamuzi, A., Ngubiri, J., & Okori, W. (2020). Towards Ensemble Learning for Tracking Food Insecurity From News Articles. [IJSDA]. *International Journal of System Dynamics Applications*, *9*(4), 129–142. doi:10.4018/IJSDA.2020100107

M., K., & S., S. V. (2021). Deep Learning Approaches to Overcome Challenges in Forensics. In Misra, S., Arumugam, C., Jaganathan, S., & S., S. (Ed.), *Confluence of AI, Machine, and Deep Learning in Cyber Forensics* (pp. 81-92). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4900-1.ch005

M., S. Y., & R., K. (2021). A Survey on Network Intrusion Detection Using Deep Generative Networks for Cyber-Physical Systems. In Luhach, A. K., & Elçi, A. (Ed.), *Artificial Intelligence Paradigms for Smart Cyber-Physical Systems* (pp. 137-159). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-5101-1.ch007

Maazouzi, F., & Zarzour, H. (2021). AI-Driven Big Healthcare Analytics. In Z. Sun (Ed.), *Intelligent Analytics With Advanced Multi-Industry Applications* (pp. 172–184). IGI Global., doi:10.4018/978-1-7998-4963-6.ch008

Mahmood, Z. (2021). Cloud Computing Technologies for Connected Digital Government. In Z. Mahmood (Ed.), *Web 2.0 and Cloud Technologies for Implementing Connected Government* (pp. 19–35). IGI Global., doi:10.4018/978-1-7998-4570-6.ch002

Mahmoud, A. S., & Hemdan, E. E. (2021). Digital Twins Concepts, Challenges, and Future Trends. In M. Del Giudice & A. Osello (Eds.), *Handbook of Research on Developing Smart Cities Based on Digital Twins* (pp. 48–60). IGI Global., doi:10.4018/978-1-7998-7091-3.ch003

Mahyoub, F. H., & Abdullah, R. (2020). Protein Secondary Structure Prediction Approaches. In J. J. Thomas, P. Karagoz, B. B. Ahamed, & P. Vasant (Eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics* (pp. 251–273). IGI Global., doi:10.4018/978-1-7998-1192-3.ch015

Majhi, V., Saikia, A., Datta, A., Sinha, A., & Paul, S. (2020). Comprehensive Review on Deep Learning for Neuronal Disorders. [IJNCR]. *International Journal of Natural Computing Research*, *9*(1), 27–44. doi:10.4018/IJNCR.2020010103

Makana, M., Nwulu, N., & Dogo, E. (2021). Automated Microcontroller-Based Irrigation System. In R. Raut & A. D. Mihovska (Eds.), *Examining the Impact of Deep Learning and IoT on Multi-Industry Applications* (pp. 45–60). IGI Global., doi:10.4018/978-1-7998-7511-6.ch004

Makori, E. O. (2020). Blockchain Applications and Trends That Promote Information Management. In B. Holland (Ed.), *Emerging Trends and Impacts of the Internet of Things in Libraries* (pp. 34–51). IGI Global., doi:10.4018/978-1-7998-4742-7.ch002

Malar, B. A., & Prabhu, J. (2020). Data Storage Security in the Cloud Environment Using the Honey Pot System. [IJeC]. *International Journal of e-Collaboration*, *16*(4), 1–14. doi:10.4018/IJeC.2020100101

Malarvizhi, N., Aswini, J., & Neeba, E. A. (2019). Resource-Aware Allocation and Load-Balancing Algorithms for Cloud Computing. In P. Raj & S. Koteeswaran (Eds.), *Novel Practices and Trends in Grid and Cloud Computing* (pp. 90–106). IGI Global., doi:10.4018/978-1-5225-9023-1.ch007

Maleh, Y. (2019). Malware Classification and Analysis Using Convolutional and Recurrent Neural Network. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 233–255). IGI Global., doi:10.4018/978-1-5225-7862-8.ch014

Maleh, Y., Sahid, A., & Belaissaoui, M. (2019). *Managing the Cloud for Information System Agility in Organizations*. IGI Global., doi:10.4018/978-1-5225-7826-0.ch006

Mallikarjuna, B. (2020). Feedback-Based Fuzzy Resource Management in IoT-Based-Cloud. [IJFC]. *International Journal of Fog Computing*, *3*(1), 1–21. doi:10.4018/IJFC.2020010101

Mani, N., Moh, M., & Moh, T. (2021). Defending Deep Learning Models Against Adversarial Attacks. [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, *13*(1), 72–89. doi:10.4018/IJSSCI.2021010105

Mani, V. R. (2020). Deep Learning Models for Semantic Multi-Modal Medical Image Segmentation. In A. Suresh, R. Udendhran, & S. Vimal (Eds.), *Deep Neural Networks for Multimodal Imaging and Biomedical Applications* (pp. 12–30). IGI Global., doi:10.4018/978-1-7998-3591-2.ch002

Manu, C., Vijaya Kumar, B. P., & Naresh, E. (2019). Anomaly Detection Using Deep Learning With Modular Networks. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 256–290). IGI Global., doi:10.4018/978-1-5225-7862-8.ch015

Mardi, V., Naresh, E., & Vijaya Kumar, B. P. (2019). A Survey on Deep Learning Techniques Used for Quality Process. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 131–152). IGI Global., doi:10.4018/978-1-5225-7862-8.ch008

Marques, G. (2019). Ambient Assisted Living and Internet of Things. In P. J. Cardoso, J. Monteiro, J. Semião, & J. M. Rodrigues (Eds.), *Harnessing the Internet of Everything (IoE) for Accelerated Innovation Opportunities* (pp. 100–115). IGI Global., doi:10.4018/978-1-5225-7332-6.ch005

Martin, M. S., Hugues, R. E., & Puliatte, A. (2019). The Use of Cloud-Computing to Promote Collaborative Learning in Higher Education. In H. L. Schnackenberg & C. Johnson (Eds.), *Preparing the Higher Education Space for Gen Z* (pp. 32–45). IGI Global., doi:10.4018/978-1-5225-7763-8.ch002

Martínez-Serrano, L. M. (2020). The Pedagogical Potential of Design Thinking for CLIL Teaching. In M. E. Gómez-Parra & C. Huertas Abril (Eds.), *Handbook of Research on Bilingual and Intercultural Education* (pp. 427–446). IGI Global., doi:10.4018/978-1-7998-2588-3.ch018

Masoodi, I. S., & Javid, B. (2019). A Review of Cryptographic Algorithms for the Internet of Things. In M. T. Banday (Ed.), *Cryptographic Security Solutions for the Internet of Things* (pp. 67–93). IGI Global., doi:10.4018/978-1-5225-5742-5.ch003

Mathur, S., & Arora, A. (2020). Internet of Things (IoT) and PKI-Based Security Architecture. In P. Kumar, V. Ponnusamy, & V. Jain (Eds.), *Industrial Internet of Things and Cyber-Physical Systems* (pp. 25–46). IGI Global., doi:10.4018/978-1-7998-2803-7.ch002

McKenna, H. (2019). *Sensing Cities and Getting Smarter*. IGI Global., doi:10.4018/978-1-5225-7882-6.ch002

Meedeniya, D., & Rubasinghe, I. (2020). A Review of Supportive Computational Approaches for Neurological Disorder Identification. In T. Wadhera & D. Kakkar (Eds.), *Interdisciplinary Approaches to Altering Neurodevelopmental Disorders* (pp. 271–302). IGI Global., doi:10.4018/978-1-7998-3069-6.ch016

Mekonnen, A. A., Seid, H. W., Mohapatra, S. K., & Prasad, S. (2021). Developing Brain Tumor Detection Model Using Deep Feature Extraction via Transfer Learning. In M. Panda & H. Misra (Eds.), *Handbook of Research on Automated Feature Engineering and Advanced Applications in Data Science* (pp. 119–137). IGI Global., doi:10.4018/978-1-7998-6659-6.ch007

Mellman, L., Williams, M. K., & Slykhuis, D. A. (2021). Using Generation Z's Learning Approaches to Create Meaningful Online Learning. In M. L. Niess & H. Gillow-Wiles (Eds.), *Handbook of Research on Transforming Teachers' Online Pedagogical Reasoning for Engaging K-12 Students in Virtual Learning* (pp. 149–169). IGI Global., doi:10.4018/978-1-7998-7222-1.ch008

Memon, M. S., Kumar, P., Mirani, A. A., Qabulio, M., & Sodhar, I. N. (2020). Deep Learning and IoT. In P. Kumar, V. Ponnusamy, & V. Jain (Eds.), *Industrial Internet of Things and Cyber-Physical Systems* (pp. 47–60). IGI Global., doi:10.4018/978-1-7998-2803-7.ch003

Menon, M. S., & Pothuraju, R. (2021). Fog Computing in Industrial Internet of Things. In Goundar, S., Avanija, J., Sunitha, G., Madhavi, K. R., & Bhushan, S. B. (Ed.), Innovations in the Industrial Internet of Things (IIoT) and Smart Factory (pp. 65-78). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3375-8.ch005

Menon, S. P. (2020). A Survey on Algorithms in Deep Learning. In P. Kumar, V. Ponnusamy, & V. Jain (Eds.), *Industrial Internet of Things and Cyber-Physical Systems* (pp. 339–350). IGI Global., doi:10.4018/978-1-7998-2803-7.ch017

Mezghani, K., & Almansour, M. A. (2019). Study of Intentions to Use Cloud CRM Within Saudi SMEs. In Mezghani, K., & Aloulou, W. (Ed.), Business Transformations in the Era of Digitalization (pp. 33-50). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7262-6.ch003

Mihailescu, M. I., & Nita, S. L. (2021). CSAP. In R. Raut & A. D. Mihovska (Eds.), *Examining the Impact of Deep Learning and IoT on Multi-Industry Applications* (pp. 249–269). IGI Global., doi:10.4018/978-1-7998-7511-6.ch014

Mikavica, B., & Kostić-Ljubisavljević, A. (2020). Security Issues of Cloud Migration and Optical Networking in Future Internet. In M. D. Stojanović & S. V. Boštjančič Rakas (Eds.), *Cyber Security of Industrial Control Systems in the Future Internet Environment* (pp. 91–106). IGI Global., doi:10.4018/978-1-7998-2910-2.ch005

Miller, B., & Thomas, M. (2021). Accessing Learning Management Systems With Smartphones. In C. M. Crawford & M. A. Simons (Eds.), *eLearning Engagement in a Transformative Social Learning Environment* (pp. 221–243). IGI Global., doi:10.4018/978-1-7998-6956-6.ch011

Milovanovic, D. A., & Pantovic, V. (2022). Interoperability in Internet of Media Things and Integration Big Media. In P. Y. Taser (Ed.), *Emerging Trends in IoT and Integration with Data Science, Cloud Computing, and Big Data Analytics* (pp. 59–79). IGI Global., doi:10.4018/978-1-7998-4186-9.ch004

Mishra, A., Gupta, N., & Gupta, B. B. (2020). Security Threats and Recent Countermeasures in Cloud Computing. In B. B. Gupta (Ed.), *Modern Principles, Practices, and Algorithms for Cloud Security* (pp. 145–161). IGI Global., doi:10.4018/978-1-7998-1082-7.ch007

Mishra, I., Bandyopadhyay, R., Ghosh, S., & Swetapadma, A. (2019). Analysis of Cutting-Edge Regression Algorithms Used for Data Analysis. In Z. Sun (Ed.), *Managerial Perspectives on Intelligent Big Data Analytics* (pp. 199–213). IGI Global., doi:10.4018/978-1-5225-7277-0.ch011

Mishra, P. S., & Nandi, D. (2021). Deep Learning for Feature Engineering-Based Improved Weather Prediction. In M. Panda & H. Misra (Eds.), *Handbook of Research on Automated Feature Engineering and Advanced Applications in Data Science* (pp. 195–217). IGI Global., doi:10.4018/978-1-7998-6659-6.ch011

Mishra, S., Sahoo, S., & Mishra, B. K. (2019). Addressing Security Issues and Standards in Internet of Things. In P. K. Mallick & S. Borah (Eds.), *Emerging Trends and Applications in Cognitive Computing* (pp. 224–257). IGI Global., doi:10.4018/978-1-5225-5793-7.ch010

Misra, P., Yadav, A. S., & Chaurasia, S. (2021). Social Media Induced Transformations in Healthcare. In A. Sharaff, G. R. Sinha, & S. Bhatia (Eds.), *New Opportunities for Sentiment Analysis and Information Processing* (pp. 72–89). IGI Global., doi:10.4018/978-1-7998-8061-5.ch004

Mitra, A. (2019). On Investigating Energy Stability for Cellular Automata Based PageRank Validation Model in Green Cloud. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(4), 66–85. doi:10.4018/IJCAC.2019100104

Mkrttchian, V. (2021). Artificial and Natural Intelligence Techniques as IoP- and IoT-Based Technologies for Sustainable Farming and Smart Agriculture. In P. Tomar & G. Kaur (Eds.), *Artificial Intelligence and IoT-Based Technologies for Sustainable Farming and Smart Agriculture* (pp. 40–53). IGI Global., doi:10.4018/978-1-7998-1722-2.ch003

Mkrttchian, V., Gamidullaeva, L., Finogeev, A., Chernyshenko, S., Chernyshenko, V., Amirov, D., & Potapova, I. (2021). Big Data and Internet of Things (IoT) Technologies' Influence on Higher Education. [IJWLTT]. *International Journal of Web-Based Learning and Teaching Technologies*, *16*(5), 137–157. doi:10.4018/IJWLTT.20210901.oa8

Mkrttchian, V., Gamidullaeva, L., Panasenko, S., & Sargsyan, A. (2019). Creating a Research Laboratory on Big Data and Internet of Things for the Study and Development of Digital Transformation. In Kaur, G., & Tomar, P. (Ed.), Handbook of Research on Big Data and the IoT (pp. 339-358). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7432-3.ch019

Mkrttchian, V., Gamidullaeva, L. A., Panasenko, S., & Sargsyan, A. (2019). Big Data Analytics and Internet of Things in Industrial Internet in Former Soviet Union Countries. In Kaur, G., & Tomar, P. (Ed.), Handbook of Research on Big Data and the IoT (pp. 359-378). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7432-3.ch020

Mohamed Jebran, P., & Banu, S. (2021). Deep Learning-Based Diabetic Retinopathy Detection. [IJOCI]. *International Journal of Organizational and Collective Intelligence*, *11*(3), 38–48. doi:10.4018/IJOCI.2021070103

Mohan, A. P., Mohamed Asfak, R., & Gladston, A. (2020). Merkle Tree and Blockchain-Based Cloud Data Auditing. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(3), 54–66. doi:10.4018/IJCAC.2020070103

Mohanty, S., Patra, P. K., Ray, M., & Mohapatra, S. (2019). An Approach for Load Balancing in Cloud Computing Using JAYA Algorithm. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *14*(1), 27–41. doi:10.4018/IJITWE.2019010102

Mondal, R. K., Ray, P., & Nandi, E., Biswajit.Biswas, Sanyal, M. K., & Sarddar, D. (2019). Load Balancing of Unbalanced Assignment Problem With Hungarian Method. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *10*(1), 46–60. doi:10.4018/IJACI.2019010103

More, N. S., & Ingle, R. B. (2020). Optimizing the Topology and Energy-Aware VM Migration in Cloud Computing. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *11*(3), 42–65. doi:10.4018/IJACI.2020070103

Mukherjee, S., Bhattacharjee, A. K., Bhattacharya, D., & Ghosal, M. (2019). Analysis of Industrial and Household IoT Data Using Computationally Intelligent Algorithm. In H. D. Purnomo (Ed.), *Computational Intelligence in the Internet of Things* (pp. 25–48). IGI Global., doi:10.4018/978-1-5225-7955-7.ch002

Mundhra, B., & Bose, R. (2021). *Cloud Computing and Its Implications for Construction IT*. IGI Global., doi:10.4018/978-1-7998-5291-9.ch007

Mundotiya, R. K., & Yadav, N. (2021). Forward Context-Aware Clickbait Tweet Identification System. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *12*(2), 21–32. doi:10.4018/IJACI.2021040102

Muramatsu, K., & Wangmo, S. (2019). Collaborative Design Education Using 3D Printing. In I. M. Santos, N. Ali, & S. Areepattamannil (Eds.), *Interdisciplinary and International Perspectives on 3D Printing in Education* (pp. 286–305). IGI Global., doi:10.4018/978-1-5225-7018-9.ch014

Murugan, R. (2020). Implementation of Deep Learning Neural Network for Retinal Images. In S. Velayutham (Ed.), *Handbook of Research on Applications and Implementations of Machine Learning Techniques* (pp. 77–95). IGI Global., doi:10.4018/978-1-5225-9902-9.ch005

Muthukumaran, V., & Ezhilmaran, D. (2020). A Cloud-Assisted Proxy Re-Encryption Scheme for Efficient Data Sharing Across IoT Systems. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *15*(4), 18–36. doi:10.4018/IJITWE.2020100102

Nadda, V., Chaudhary, H. S., & Arnott, I. (2020). Cloud Computing in Tourism. In J. D. Santos & Ó. L. Silva (Eds.), *Digital Marketing Strategies for Tourism, Hospitality, and Airline Industries* (pp. 141–155). IGI Global., doi:10.4018/978-1-5225-9783-4.ch007

Namasudra, S., & Roy, P. (2018). PpBAC. [JOEUC]. *Journal of Organizational and End User Computing*, *30*(4), 14–31. doi:10.4018/JOEUC.2018100102

Nambobi, M., Ruth, K., Alli, A. A., & Ssemwogerere, R. (2021). The Age of Autonomous Internet of Things Devices. In S. Velayutham (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 1–16). IGI Global., doi:10.4018/978-1-7998-3111-2.ch001

Nami, N., & Moh, M. (2019). Adversarial Attacks and Defense on Deep Learning Models for Big Data and IoT. In B. Gupta & D. P. Agrawal (Eds.), *Handbook of Research on Cloud Computing and Big Data Applications in IoT* (pp. 39–66). IGI Global., doi:10.4018/978-1-5225-8407-0.ch003

Namiot, D., & Sneps-Sneppe, M. (2020). Internet of Things and Cyber-Physical Systems at the University. In S. Balandin & E. Balandina (Eds.), *Tools and Technologies for the Development of Cyber-Physical Systems* (pp. 285–302). IGI Global., doi:10.4018/978-1-7998-1974-5.ch011

Nan, D. (2020). The Design of Power Security Defense System Based on Resource Pool Cloud Computing Technology. [IJISMD]. *International Journal of Information System Modeling and Design*, *11*(1), 1–11. doi:10.4018/IJISMD.2020010101

Narang, A., Gupta, D., & Kaur, A. (2020). Biometrics-Based Un-Locker to Enhance Cloud Security Systems. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(4), 1–12. doi:10.4018/IJCAC.2020100101

Narongou, D., & Sun, Z. (2021). Big Data Analytics for Smart Airport Management. In Z. Sun (Ed.), *Intelligent Analytics With Advanced Multi-Industry Applications* (pp. 209–231). IGI Global., doi:10.4018/978-1-7998-4963-6.ch010

Nayak, S. C., Parida, S., Tripathy, C., & Pattnaik, P. K. (2019). Dynamic Backfilling Algorithm to Increase Resource Utilization in Cloud Computing. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *14*(1), 1–26. doi:10.4018/IJITWE.2019010101

Nazir, S., Patel, S., & Patel, D. (2020). Assessing Hyper Parameter Optimization and Speedup for Convolutional Neural Networks. [IJAIML]. *International Journal of Artificial Intelligence and Machine Learning*, *10*(2), 1–17. doi:10.4018/IJAIML.2020070101

Neeba, E. A., Aswini, J., & Priyadarshini, R. (2019). Leveraging the Internet of Things (IoT) Paradigm Towards Smarter Applications. In P. Raj & S. Koteeswaran (Eds.), *Novel Practices and Trends in Grid and Cloud Computing* (pp. 306–324). IGI Global., doi:10.4018/978-1-5225-9023-1.ch017

Neves, P. C., & Bernardino, J. R. (2021). The Role of Big Data and Business Analytics in Decision Making. In H. Rahman (Ed.), *Human-Computer Interaction and Technology Integration in Modern Society* (pp. 226–257). IGI Global., doi:10.4018/978-1-7998-5849-2.ch010

Neware, R., & Shrawankar, U. (2020). Fog Computing Architecture, Applications and Security Issues. [IJFC]. *International Journal of Fog Computing*, *3*(1), 75–105. doi:10.4018/IJFC.2020010105

Nirmalan, R., Gokulakrishnan, K., & Jesu Vedha Nayahi, J. (2019). Load Balancing in Heterogeneous Cluster Cloud Computing. In P. Raj & S. Koteeswaran (Eds.), *Novel Practices and Trends in Grid and Cloud Computing* (pp. 146–169). IGI Global., doi:10.4018/978-1-5225-9023-1.ch010

Nisha Angeline, C. V., & Lavanya, R. (2019). Fog Computing and Its Role in the Internet of Things. In K. Munir (Ed.), *Advancing Consumer-Centric Fog Computing Architectures* (pp. 63–71). IGI Global., doi:10.4018/978-1-5225-7149-0.ch003

Odia, J. O., & Akpata, O. T. (2021). Role of Data Science and Data Analytics in Forensic Accounting and Fraud Detection. In B. Patil & M. Vohra (Eds.), *Handbook of Research on Engineering, Business, and Healthcare Applications of Data Science and Analytics* (pp. 203–227). IGI Global., doi:10.4018/978-1-7998-3053-5.ch011

Olawuyi, J. N., Akhigbe, B. I., Afolabi, B. S., & Okine, A. (2021). Semantic Medical Image Analysis. In O. Daramola & T. Moser (Eds.), *Advanced Concepts, Methods, and Applications in Semantic Computing* (pp. 128–146). IGI Global., doi:10.4018/978-1-7998-6697-8.ch007

Ouldkablia, M. E., Kechar, B., & Bouzefrane, S. (2020). IoT-Based Smart Home Process Management Using a Workflow Approach. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *15*(2), 50–76. doi:10.4018/IJITWE.2020040103

Owusu, A. (2020). Determinants of Cloud Business Intelligence Adoption Among Ghanaian SMEs. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(4), 48–69. doi:10.4018/IJCAC.2020100104

Özen, I. A. (2020). Internet of Things in Tourism. In E. Çeltek (Ed.), *Handbook of Research on Smart Technology Applications in the Tourism Industry* (pp. 131–154). IGI Global., doi:10.4018/978-1-7998-1989-9.ch007

P., C., D., R., V., P., S. V., A. J., & B., B. (2021). Data Security and Privacy Requirements in Edge Computing. In Ambika, P., Donald, A., & Kumar, A. (Ed.), *Cases on Edge Computing and Analytics* (pp. 171-187). IGI Global. https://doi.org/doi:10.4018/978-1-7998-4873-8.ch009

P., S., Mujumdar, A. M., & Biswas, A. (2021). Cognitive Computing and Its Applications. In Musiolik, T. H., & Cheok, A. D. (Ed.), *Analyzing Future Applications of AI, Sensors, and Robotics in Society* (pp. 47-68). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3499-1.ch004

P., S., P., K., V., S., K., L., R., M. D., K., S., M., S., & C., S. (2021). Cloud-Based Big Data Analysis Tools and Techniques Towards Sustainable Smart City Services. In Butun, I. (Ed.), *Decision Support Systems and Industrial IoT in Smart Grid, Factories, and Cities* (pp. 63-90). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-7468-3.ch004

Padayachee, N., & Civilcharran, S. (2021). Predicting Student Intention to Use Cloud Services for Educational Purposes Based on Perceived Security and Privacy. In P. Ndayizigamiye, G. Barlow-Jones, R. Brink, S. Bvuma, R. Minty, & S. Mhlongo (Eds.), *Perspectives on ICT4D and Socio-Economic Growth Opportunities in Developing Countries* (pp. 335–365). IGI Global., doi:10.4018/978-1-7998-2983-6.ch013

Panda, S., Sahu, P. C., & Chandran, K. (2021). Generic Issue in Edge Computing and Use Cases. In P. Ambika, A. Donald, & A. Kumar (Eds.), *Cases on Edge Computing and Analytics* (pp. 153–170). IGI Global., doi:10.4018/978-1-7998-4873-8.ch008

Panwar, A., & Bhatnagar, V. (2020). Data Lake Architecture. [IJOCI]. *International Journal of Organizational and Collective Intelligence*, *10*(1), 63–75. doi:10.4018/IJOCI.2020010104

Parvathi, R., & Pattabiraman, V. (2019). A Similarity-Based Object Classification Using Deep Neural Networks. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 197–219). IGI Global., doi:10.4018/978-1-5225-7862-8.ch012

Patra, S. S., Goswami, V., & Mund, G. B. (2020). Performance Evaluation of Cloud Systems by Switching the Virtual Machines Power Mode Between the Sleep Mode and Active Mode. In P. Sahoo (Ed.), *Handbook of Research on Developments and Trends in Industrial and Materials Engineering* (pp. 145–168). IGI Global., doi:10.4018/978-1-7998-1831-1.ch007

Peng, H., & Xu, Y. (2019). Research on Resource Allocation Strategy of PaaS Platform. [JITR]. *Journal of Information Technology Research*, *12*(1), 63–76. doi:10.4018/JITR.2019010105

Potluri, R. M., & Vajjhala, N. R. (2021). Risks in Adoption and Implementation of Big Data Analytics. [IJRCM]. *International Journal of Risk and Contingency Management*, *10*(3), 1–11. doi:10.4018/IJRCM.2021070101

Prabhudas, J., & Reddy, C. H. (2020). Leveraging Natural Language Processing Applications Using Machine Learning. In A. Solanki, S. Kumar, & A. Nayyar (Eds.), *Handbook of Research on Emerging Trends and Applications of Machine Learning* (pp. 338–360). IGI Global., doi:10.4018/978-1-5225-9643-1.ch016

Pramanik, P. K., Pal, S., & Mukhopadhyay, M. (2020). Big Data and Big Data Analytics for Improved Healthcare Service and Management. [IJPHIM]. *International Journal of Privacy and Health Information Management*, *8*(1), 13–51. doi:10.4018/IJPHIM.2020010102

Prasad, V. K., & Bhavsar, M. D. (2020). Monitoring IaaS Cloud for Healthcare Systems. [IJEHMC]. *International Journal of E-Health and Medical Communications*, *11*(3), 54–70. doi:10.4018/IJEHMC.2020070104

Preethi, D., & Khare, N. (2021). An Intelligent Network Intrusion Detection System Using Particle Swarm Optimization (PSO) and Deep Network Networks (DNN). [IJSIR]. *International Journal of Swarm Intelligence Research*, *12*(2), 57–73. doi:10.4018/IJSIR.2021040104

Priya, G., & Jaisankar, N. (2019). A Fuzzy Based Trust Evaluation Model for Service Selection in Cloud Environment. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *11*(4), 13–27. doi:10.4018/IJGHPC.2019100102

Priyadarshinee, P. (2020). Examining Critical Success Factors of Cloud Computing Adoption. [IJDSST]. *International Journal of Decision Support System Technology*, *12*(2), 80–96. doi:10.4018/IJDSST.2020040105

Priyadarshini, R., Barik, R. K., & Mishra, B. K. (2020). Meta-Heuristic and Non-Meta-Heuristic Energy-Efficient Load Balancing Algorithms in Cloud Computing. In B. B. Gupta (Ed.), *Modern Principles, Practices, and Algorithms for Cloud Security* (pp. 203–222). IGI Global., doi:10.4018/978-1-7998-1082-7.ch010

Qafzezi, E., Bylykbashi, K., Spaho, E., & Barolli, L. (2019). A New Fuzzy-Based Resource Management System for SDN-VANETs. [IJMCMC]. *International Journal of Mobile Computing and Multimedia Communications*, *10*(4), 1–12. doi:10.4018/IJMCMC.2019100101

Quiñonez, Y. (2021). An Overview of Applications of Artificial Intelligence Using Different Techniques, Algorithms, and Tools. In Negrón, A. P., & Muñoz, M. (Eds.), Latin American Women and Research Contributions to the IT Field (pp. 325-347). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-7552-9.ch015

R., D., & R., K. (2022). Insights on the Prospects of Multi-Access Edge Computing. In Saini, K., & Raj, P. (Ed.), *Advancing Smarter and More Secure Industrial Applications Using AI, IoT, and Blockchain Technology* (pp. 37-55). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-8367-8.ch002

R., H., & Devi, A. (2021). Chemical Named Entity Recognition Using Deep Learning Techniques. In Tanwar, P., Saxena, A., & Priya, C. (Ed.), *Deep Natural Language Processing and AI Applications for Industry 5.0* (pp. 59-73). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-7728-8.ch004

R., K., B., A., A., M., & R., V. (2021). Edge Computing and IoT Technologies for Medical Applications. In Ambika, P., Donald, A., & Kumar, A. (Ed.), *Cases on Edge Computing and Analytics* (pp. 270-287). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4873-8.ch013

R., T., & Sevugan, P. (2021). Hyperspectral Image Classification Through Machine Learning and Deep Learning Techniques. In Swarnalatha, P., & Prabu, S. (Ed.), *Applications of Artificial Intelligence for Smart Technology* (pp. 103-121). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3335-2.ch008

Rababah, K. A., Bilal Ali Al-nassar, & Al-Nsour, S. N. (2020). Factors Influencing the Adoption of Cloud Computing in Small and Medium Enterprises in Jordan. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(3), 96–110. doi:10.4018/IJCAC.2020070106

Radhakrishnan, S., & Vijayarajan, V. (2020). Optimized Deep Learning System for Crop Health Classification Strategically Using Spatial and Temporal Data. In J. J. Thomas, P. Karagoz, B. B. Ahamed, & P. Vasant (Eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics* (pp. 233–250). IGI Global., doi:10.4018/978-1-7998-1192-3.ch014

Rajangam, V. N., S., R., K., & Mallikarjuna, K. (2021). Performance Analysis of VGG19 Deep Learning Network Based Brain Image Fusion. In Raj, A. J., Mahesh, V. G., & Nersisson, R. (Ed.), Handbook of Research on Deep Learning-Based Image Analysis Under Constrained and Unconstrained Environments (pp. 145-166). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-6690-9.ch008

Rajasekaran, R., Govinda, K., Masih, J., & Sruthi, M. (2020). Health Monitoring System for Individuals Using Internet of Things. In P. Pankajavalli & G. Karthick (Eds.), *Incorporating the Internet of Things in Healthcare Applications and Wearable Devices* (pp. 150–164). IGI Global., doi:10.4018/978-1-7998-1090-2.ch009

Rajendran, R., Kalidasan, A., & B., C. R. (2021). Convergence of AI, ML, and DL for Enabling Smart Intelligence. In Velayutham, S. (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 180-195). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch011

Rajesh, S., & Nalini, N. J. (2021). Recognition of Musical Instrument Using Deep Learning Techniques. [IJIRR]. *International Journal of Information Retrieval Research*, *11*(4), 41–60. doi:10.4018/IJIRR.2021100103

Raju, P. S., Rajendran, R. A., & Mahalingam, M. (2021). Perspectives of Machine Learning and Deep Learning in Internet of Things and Cloud. In S. Velayutham (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 248–264). IGI Global., doi:10.4018/978-1-7998-3111-2.ch014

Raković, R. M. (2020). Application of Cloud Computing in Electric Power Utility Systems. In M. D. Stojanović & S. V. Boštjančič Rakas (Eds.), *Cyber Security of Industrial Control Systems in the Future Internet Environment* (pp. 229–247). IGI Global., doi:10.4018/978-1-7998-2910-2.ch011

Rangra, A., Sehgal, V. K., & Shukla, S. (2019). A Novel Approach of Cloud Based Scheduling Using Deep-Learning Approach in E-Commerce Domain. [IJISMD]. *International Journal of Information System Modeling and Design*, *10*(3), 59–75. doi:10.4018/IJISMD.2019070104

Rani, S. (2021). A Study on COVID-19 Prediction and Detection With Artificial Intelligence-Based Real-Time Healthcare Monitoring Systems. In M. Roy & L. R. Gupta (Eds.), *Machine Learning and Data Analytics for Predicting, Managing, and Monitoring Disease* (pp. 52–63). IGI Global., doi:10.4018/978-1-7998-7188-0.ch004

Rao, K. S. S., S., Kumar, P. S., V., R., & Raghu, K. (2021). Leveraging Big Data Analytics and Hadoop in Developing India's Healthcare Services. In Zhao, J., & Kumar, V. (Ed.), Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies (pp. 396-407). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-6870-5.ch027

Rao, N. R. (2019). Cloud Computing for E-Governance. In N. Rao (Ed.), *Global Virtual Enterprises in Cloud Computing Environments* (pp. 113–144). IGI Global., doi:10.4018/978-1-5225-3182-1.ch005

Rao, N. R. (2019). Providing Healthcare Services in the Virtual Environment. In N. Rao (Ed.), *Global Virtual Enterprises in Cloud Computing Environments* (pp. 145–170). IGI Global., doi:10.4018/978-1-5225-3182-1.ch006

Rath, M. (2019). Security Challenges and Resolution in Cloud Computing and Cloud of Things. In G. Kecskemeti (Ed.), *Applying Integration Techniques and Methods in Distributed Systems and Technologies* (pp. 79–102). IGI Global., doi:10.4018/978-1-5225-8295-3.ch004

Rawat, A., Sushil, R., & Agarwal, A. (2020). Review of Fault Tolerance Frameworks in the Cloud. [IJISMD]. *International Journal of Information System Modeling and Design*, *11*(3), 79–99. doi:10.4018/IJISMD.2020070105

Rebbah, M., Slimani, Y., Debakla, M., & Smail, O. (2018). Toward A Performing Resource Provisioning Model for Hybrid Cloud. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *10*(4), 15–42. doi:10.4018/IJGHPC.2018100102

Rege, P. P., & Akhter, S. (2020). Text Separation From Document Images. In M. Mahrishi, K. K. Hiran, G. Meena, & P. Sharma (Eds.), *Machine Learning and Deep Learning in Real-Time Applications* (pp. 283–313). IGI Global., doi:10.4018/978-1-7998-3095-5.ch013

Revathi, T., Muneeswaran, K., & Blessa Binolin Pepsi, M. (2019). *Big Data Architecture Components*. IGI Global., doi:10.4018/978-1-5225-3790-8.ch002

Robinson, M., & Jones, K. (2020). Securing the Cloud for Big Data. In R. C. Joshi & B. B. Gupta (Eds.), *Security, Privacy, and Forensics Issues in Big Data* (pp. 1–23). IGI Global., doi:10.4018/978-1-5225-9742-1.ch001

Rodge, J., & Jaiswal, S. (2019). Comprehensive Overview of Neural Networks and Its Applications in Autonomous Vehicles. In H. D. Purnomo (Ed.), *Computational Intelligence in the Internet of Things* (pp. 159–173). IGI Global., doi:10.4018/978-1-5225-7955-7.ch007

Rolo, A., & Alves, R. C. (2021). Digital Economy and Cloud. In Carvalho, L. C., Reis, L., Prata, A., & Pereira, R. (Ed.), Handbook of Research on Multidisciplinary Approaches to Entrepreneurship, Innovation, and ICTs (pp. 22-41). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4099-2.ch002

Routray, S. K. (2021). Narrowband Internet of Things. In Khosrow-Pour D.B.A., M. (Eds.), Encyclopedia of Information Science and Technology, Fifth Edition (pp. 913-923). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3479-3.ch063

Roy, D. (2021). Critical Success Factors of Analytics and Digital Technologies Adoption in Supply Chain. In Khosrow-Pour D.B.A., M. (Eds.), Encyclopedia of Organizational Knowledge, Administration, and Technology (pp. 2458-2471). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3473-1.ch170

S., A. R. (2021). Recommender System Techniques and Approaches to Improve the Modern Learning Challenges. In Gulzar, Z., & Leema, A. A. (Eds.), *Machine Learning Approaches for Improvising Modern Learning Systems* (pp. 114-143). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-5009-0.ch005

S., B. P., & Arumugam, V. (2019). Business Aspects, Models, and Opportunities of IoT. In Nagarajan, G., & Minu, R. (Ed.), *Edge Computing and Computational Intelligence Paradigms for the IoT* (pp. 69-99). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-8555-8.ch006

S. G. S., D. S., B., & L., M. (2021). Big Data Analytics and Its Applications in IoT. In Velayutham, S. (Ed.), Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing (pp. 146-158). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch009

S., J. F., S., K., & S. A. R., S. M. (2021). A New Approach of Deep Learning-Based Tamil Vowels Prediction Using Segmentation and U-Net Architecture. In Raj, A. J., Mahesh, V. G., & Nersisson, R. (Ed.), *Handbook of Research on Deep Learning-Based Image Analysis Under Constrained and Unconstrained Environments* (pp. 186-206). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-6690-9.ch010

S., K., & D., T. (2020). Deep Learning Approach for Extracting Catch Phrases from Legal Documents. In S., S., & M., J. (Ed.), *Neural Networks for Natural Language Processing* (pp. 143-158). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-1159-6.ch009

S., S., R., D. P., & K., K. (2021). Big Data Analytics in Cloud Platform. In Velayutham, S. (Ed.), *Challenges and Opportunities for the Convergence of IoT, Big Data, and Cloud Computing* (pp. 159-179). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3111-2.ch010

S., U., D., S., Mouliganth, C., & M., V. E. (2021). KidNet. In Senthilnathan, K., Shanmugam, B., Goyal, D., Annapoorani, I., & Samikannu, R. (Ed.), *Deep Learning Applications and Intelligent Decision Making in Engineering* (pp. 114-129). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2108-3.ch004

Sabuncu, I. (2021). Understanding Tourist Perceptions and Expectations During Pandemic Through Social Media Big Data. In M. Demir, A. Dalgıç, & F. D. Ergen (Eds.), *Handbook of Research on the Impacts and Implications of COVID-19 on the Tourism Industry* (pp. 330–350). IGI Global., doi:10.4018/978-1-7998-8231-2.ch016

Sadasivam, U. M., & Ganesan, N. (2021). Detecting Fake News Using Deep Learning and NLP. In Misra, S., Arumugam, C., Jaganathan, S., & S., S. (Ed.), Confluence of AI, Machine, and Deep Learning in Cyber Forensics (pp. 117-133). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4900-1.ch007

Sahana, S., Mukherjee, T., & Sarddar, D. (2020). A Conceptual Framework Towards Implementing a Cloud-Based Dynamic Load Balancer Using a Weighted Round-Robin Algorithm. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(2), 22–35. doi:10.4018/IJCAC.2020040102

Sahu, H., & Gaytri,. (2021). Security Issues in Fog Computing and ML-Based Solutions. In Dua, M., & Jain, A. K. (Ed.), *Handbook of Research on Machine Learning Techniques for Pattern Recognition and Information Security* (pp. 209-234). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3299-7.ch013

Sahu, N. K., Patnaik, M., & Snigdh, I. (2021). Data Analytics and Its Applications in Brief. In S. Goundar & P. K. Rayani (Eds.), *Applications of Big Data in Large- and Small-Scale Systems* (pp. 115–125). IGI Global., doi:10.4018/978-1-7998-6673-2.ch008

Sahu, P., Chug, A., Singh, A. P., Singh, D., & Singh, R. P. (2021). Challenges and Issues in Plant Disease Detection Using Deep Learning. In M. Dua & A. K. Jain (Eds.), *Handbook of Research on Machine Learning Techniques for Pattern Recognition and Information Security* (pp. 56–74). IGI Global., doi:10.4018/978-1-7998-3299-7.ch004

Saikia, A., & Paul, S. (2020). Application of Deep Learning for EEG. In D. S. Sisodia, R. Pachori, & L. Garg (Eds.), *Handbook of Research on Advancements of Artificial Intelligence in Healthcare Engineering* (pp. 106–123). IGI Global., doi:10.4018/978-1-7998-2120-5.ch007

Saiod, A. K., & van Greunen, D. (2021). The Impact of Deep Learning on the Semantic Machine Learning Representation. In O. Daramola & T. Moser (Eds.), *Advanced Concepts, Methods, and Applications in Semantic Computing* (pp. 22–49). IGI Global., doi:10.4018/978-1-7998-6697-8.ch002

Salgotra, K., Khullar, V., Singh, H. P., & Khan, S. A. (2021). Diagnosis of Attention Deficit Hyperactivity Disorder. In R. Raut & A. D. Mihovska (Eds.), *Examining the Impact of Deep Learning and IoT on Multi-Industry Applications* (pp. 31–44). IGI Global., doi:10.4018/978-1-7998-7511-6.ch003

Salhi, D. E., Tari, A., & Kechadi, M. T. (2021). Using Clustering for Forensics Analysis on Internet of Things. [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, *13*(1), 56–71. doi:10.4018/IJSSCI.2021010104

Sambrekar, K., & Rajpurohit, V. S. (2019). Fast and Efficient Multiview Access Control Mechanism for Cloud Based Agriculture Storage Management System. [IJCAC]. *International Journal of Cloud Applications and Computing*, *9*(1), 33–49. doi:10.4018/IJCAC.2019010103

Sani, R. M. (2019). Adopting Internet of Things for Higher Education. In A. Raman & M. Rathakrishnan (Eds.), *Redesigning Higher Education Initiatives for Industry 4.0* (pp. 23–40). IGI Global., doi:10.4018/978-1-5225-7832-1.ch002

Sarhan, A. (2019). Cloud-Based IoT Platform. In P. J. Cardoso, J. Monteiro, J. Semião, & J. M. Rodrigues (Eds.), *Harnessing the Internet of Everything (IoE) for Accelerated Innovation Opportunities* (pp. 116–147). IGI Global., doi:10.4018/978-1-5225-7332-6.ch006

Sarkar, K., & Li, B. (2021). Deep Learning for Medical Image Segmentation. In S. Saxena & S. Paul (Eds.), *Deep Learning Applications in Medical Imaging* (pp. 40–77). IGI Global., doi:10.4018/978-1-7998-5071-7.ch002

Sasidharan, S., Salmasi, M. Y., Pirola, S., & Jarral, O. A. (2021). Relevance of Machine Learning to Cardiovascular Imaging. In S. Saxena & S. Paul (Eds.), *Deep Learning Applications in Medical Imaging* (pp. 78–99). IGI Global., doi:10.4018/978-1-7998-5071-7.ch003

Sasikumar, K., & Vijayakumar, B. (2020). An Efficient Multi-Objective Model for Data Replication in Cloud Computing Environment. [IJEIS]. *International Journal of Enterprise Information Systems*, *16*(1), 69–91. doi:10.4018/IJEIS.2020010104

Sasirekha, K., & Thangavel, K. (2020). A Novel Biometric Image Enhancement Approach With the Hybridization of Undecimated Wavelet Transform and Deep Autoencoder. In P. Ganapathi & D. Shanmugapriya (Eds.), *Handbook of Research on Machine and Deep Learning Applications for Cyber Security* (pp. 245–269). IGI Global., doi:10.4018/978-1-5225-9611-0.ch012

SATTA, A., & Mostefai, S. (2020). Strategic Outsourcing to Cloud Computing. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(1), 11–27. doi:10.4018/IJCAC.2020010102

Saxena, A. B., & Dave, M. (2020). Certification Attainment - A Gizmo to Evaluate Provider's Trust. [IJNCR]. *International Journal of Natural Computing Research*, *9*(1), 1–12. doi:10.4018/IJNCR.2020010101

Saxena, D. (2021). Big Data for Digital Transformation of Public Services. In K. Sandhu (Ed.), *Disruptive Technology and Digital Transformation for Business and Government* (pp. 250–266). IGI Global., doi:10.4018/978-1-7998-8583-2.ch013

Saxena, S., Paul, S., Garg, A., Saikia, A., & Datta, A. (2020). Deep Learning in Computational Neuroscience. In R. Kashyap & A. Kumar (Eds.), *Challenges and Applications for Implementing Machine Learning in Computer Vision* (pp. 43–63). IGI Global., doi:10.4018/978-1-7998-0182-5.ch002

Sedkaoui, S. (2019). *Entrepreneurship and Big Data*. IGI Global., doi:10.4018/978-1-5225-7609-9.ch007

Sedkaoui, S. (2019). *First of All, Understand Data Analytics Context and Changes*. IGI Global., doi:10.4018/978-1-5225-7609-9.ch004

Sedkaoui, S., & Khelfaoui, M. (2019). Building an Analytics Culture to Boost a Data-Driven Entrepreneur's Business Model. In Z. Sun (Ed.), *Managerial Perspectives on Intelligent Big Data Analytics* (pp. 260–291). IGI Global., doi:10.4018/978-1-5225-7277-0.ch014

Sen Saxena, V., Johri, P., & Kumar, A. (2021). AI-Enabled Support System for Melanoma Detection and Classification. [IJRQEH]. *International Journal of Reliable and Quality E-Healthcare*, *10*(4), 58–75. doi:10.4018/IJRQEH.2021100104

Shah, J. L., Bhat, H. F., & Khan, A. I. (2019). CloudIoT. [IJDCF]. *International Journal of Digital Crime and Forensics*, *11*(3), 1–22. doi:10.4018/IJDCF.2019070101

Shaila, S. G., Rajkumari, S., & Ayyasamy, V. (2020). Introducing the Deep Learning for Digital Age. In S. Velayutham (Ed.), *Handbook of Research on Applications and Implementations of Machine Learning Techniques* (pp. 317–333). IGI Global., doi:10.4018/978-1-5225-9902-9.ch017

Shally, Sharma, S. K., & Kumar, S. (2020). A Dynamic Threshold Based Energy Efficient Method for Cloud Datacenters. [IJSI]. *International Journal of Software Innovation*, *8*(2), 54–67. doi:10.4018/IJSI.2020040104

Sharma, A. (2021). Smart Agriculture Services Using Deep Learning, Big Data, and IoT (Internet of Things). In A. K. Gupta, D. Goyal, V. Singh, & H. Sharma (Eds.), *Smart Agricultural Services Using Deep Learning, Big Data, and IoT* (pp. 166–202). IGI Global., doi:10.4018/978-1-7998-5003-8.ch010

Sharma, A., Chudey, A. S., & Singh, M. (2021). COVID-19 Detection Using Chest X-Ray and Transfer Learning. In M. Dua & A. K. Jain (Eds.), *Handbook of Research on Machine Learning Techniques for Pattern Recognition and Information Security* (pp. 171–186). IGI Global., doi:10.4018/978-1-7998-3299-7.ch011

Sharma, A., & Kaur, J. (2021). Artificial Intelligence Based System. [IRMJ]. *Information Resources Management Journal*, *34*(2), 80–90. doi:10.4018/IRMJ.2021040105

Sharma, L., & Lohan, N. (2019). Internet of Things With Object Detection. In Kaur, G., & Tomar, P. (Ed.), Handbook of Research on Big Data and the IoT (pp. 89-100). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7432-3.ch006

Sharma, M., Kumar, R., & Jain, A. (2020). A Proficient Approach for Load Balancing in Cloud Computing-Join Minimum Loaded Queue. [IJISMD]. *International Journal of Information System Modeling and Design*, *11*(1), 12–36. doi:10.4018/IJISMD.2020010102

Sharma, N., & Ahlawat, P. (2020). 5G for IoT. In S. M. El-Kader & H. Hussein (Eds.), *Fundamental and Supportive Technologies for 5G Mobile Networks* (pp. 1–23). IGI Global., doi:10.4018/978-1-7998-1152-7.ch001

Sharma, P., Sengupta, J., & Suri, P. K. (2019). WLI Fuzzy Clustering and Adaptive Lion Neural Network (ALNN) for Cloud Intrusion Detection. [IJDAI]. *International Journal of Distributed Artificial Intelligence*, *11*(1), 1–17. doi:10.4018/IJDAI.2019010101

Sharma, R., Morwal, S., & Agarwal, B. (2021). Entity-Extraction Using Hybrid Deep-Learning Approach for Hindi text. [IJCINI]. *International Journal of Cognitive Informatics and Natural Intelligence*, *15*(3), 1–11. doi:10.4018/IJCINI.20210701.oa1

Sharma, S., & Kumar, V. (2019). Transfer Learning in 2.5D Face Image for Occlusion Presence and Gender Classification. In A. E. Hassanien, A. Darwish, & C. L. Chowdhary (Eds.), *Handbook of Research on Deep Learning Innovations and Trends* (pp. 97–113). IGI Global., doi:10.4018/978-1-5225-7862-8.ch006

Sharma, S., & Saini, H. (2019). Efficient Solution for Load Balancing in Fog Computing Utilizing Artificial Bee Colony. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *10*(4), 60–77. doi:10.4018/IJACI.2019100104

Shelar, M., Sane, S., & Kharat, V. (2019). A Novel Energy Efficient and SLA-Aware Approach for Cloud Resource Management. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *11*(2), 63–84. doi:10.4018/IJGHPC.2019040104

Shin, J., Guo, Q., & Gierl, M. J. (2021). Automated Essay Scoring Using Deep Learning Algorithms. In Khosrow-Pour D.B.A., M. (Ed.), Handbook of Research on Modern Educational Technologies, Applications, and Management (pp. 37-47). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3476-2.ch003

Shirazi, Z. A., de Souza, C. P., Kashef, R., & Rodrigues, F. F. (2020). Deep Learning in the Healthcare Industry. In M. Gul, E. Celik, S. Mete, & F. Serin (Eds.), *Computational Intelligence and Soft Computing Applications in Healthcare Management Science* (pp. 220–245). IGI Global., doi:10.4018/978-1-7998-2581-4.ch010

Shrivastav, U., & Singh, S. K. (2019). Digital Image Classification Techniques. In B. Pandey & A. Khamparia (Eds.), *Hidden Link Prediction in Stochastic Social Networks* (pp. 162–187). IGI Global., doi:10.4018/978-1-5225-9096-5.ch009

Shu, J., Jain, H., & Liang, C. (2019). Business Process Driven Trust-Based Task Scheduling. [IJWSR]. *International Journal of Web Services Research*, *16*(3), 1–28. doi:10.4018/IJWSR.2019070101

Shukla, R., Yadav, A. K., & Singh, T. R. (2021). Application of Deep Learning in Biological Big Data Analysis. In H. Saini, G. Rathee, & D. K. Saini (Eds.), *Large-Scale Data Streaming, Processing, and Blockchain Security* (pp. 117–148). IGI Global., doi:10.4018/978-1-7998-3444-1.ch006

Siddiqui, S., Darbari, M., & Yagyasen, D. (2020). Enhancing the Capability of Load Management Techniques in Cloud Using H_FAC Algorithm Optimization. [IJeC]. *International Journal of e-Collaboration*, *16*(2), 65–81. doi:10.4018/IJeC.2020040105

Siham, K., & Belabbas, Y. (2019). Combination of Scheduling and Dynamic Data Replication for Cloud Computing Workflows. [IJIRR]. *International Journal of Information Retrieval Research*, *9*(4), 23–35. doi:10.4018/IJIRR.2019100103

Şimşek, M. A., & Orman, Z. (2021). A Study on Deep Learning Methods in the Concept of Digital Industry 4.0. In K. Sandhu (Ed.), *Handbook of Research on Management and Strategies for Digital Enterprise Transformation* (pp. 318–339). IGI Global., doi:10.4018/978-1-7998-5015-1.ch016

Sindhwani, N., Verma, S., Bajaj, T., & Anand, R. (2021). Comparative Analysis of Intelligent Driving and Safety Assistance Systems Using YOLO and SSD Model of Deep Learning. [IJISMD]. *International Journal of Information System Modeling and Design*, *12*(1), 131–146. doi:10.4018/IJISMD.2021010107

Singh, A., Sinha, U., & Sharma, D. K. (2020). Cloud-Based IoT Architecture in Green Buildings. In A. Solanki & A. Nayyar (Eds.), *Green Building Management and Smart Automation* (pp. 164–183). IGI Global., doi:10.4018/978-1-5225-9754-4.ch008

Singh, D., Mishra, M., & Sahana, S. (2019). Big-Data-Based Techniques for Predictive Intelligence. In P. Gupta, T. Ören, & M. Singh (Eds.), *Predictive Intelligence Using Big Data and the Internet of Things* (pp. 1–18). IGI Global., doi:10.4018/978-1-5225-6210-8.ch001

Singh, J. (2019). Exploring Expansion and Innovations in Cloud Computing. [IJRDIS]. *International Journal of R&D Innovation Strategy*, *1*(1), 46–59. doi:10.4018/IJRDIS.2019010104

Singh, J., & Raghuvanshi, K. K. (2020). Regulations and Standards in Public Cloud. [JITR]. *Journal of Information Technology Research*, *13*(3), 21–36. doi:10.4018/JITR.2020070102

Singhania, U., & Tripathy, B. K. (2021). Text-Based Image Retrieval Using Deep Learning. In Khosrow-Pour D.B.A., M. (Ed.), Encyclopedia of Information Science and Technology, Fifth Edition (pp. 87-97). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-3479-3.ch007

Sivashanmugam, G., Shantharajah, S., & Iyengar, N. N. (2019). Avian Based Intelligent Algorithm to Provide Zero Tolerance Load Balancer for Cloud Based Computing Platforms. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *11*(4), 42–67. doi:10.4018/IJGHPC.2019100104

Solanki, A., & Nayyar, A. (2019). Green Internet of Things (G-IoT). In Kaur, G., & Tomar, P. (Ed.), Handbook of Research on Big Data and the IoT (pp. 379-405). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-7432-3.ch021

Soltane, M., Okba, K., Makhlouf, D., & Eom, S. B. (2018). Smart Configuration and Auto Allocation of Resource in Cloud Data Centers. [IJBAN]. *International Journal of Business Analytics*, *5*(4), 1–23. doi:10.4018/IJBAN.2018100101

Soltani, B., Ghenai, A., & Zeghib, N. (2020). Execution of Long-Duration Multi-Cloud Serverless Functions Using Selective Migration-Based Approach. [IJCAC]. *International Journal of Cloud Applications and Computing*, *10*(4), 70–97. doi:10.4018/IJCAC.2020100105

Somula, R., & Sasikala, R. (2019). A Load and Distance Aware Cloudlet Selection Strategy in Multi-Cloudlet Environment. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *11*(2), 85–102. doi:10.4018/IJGHPC.2019040105

Sonti, V. J., & G., S. (2019). Need for Internet of Things. In Nagarajan, G., & Minu, R. (Ed.), *Edge Computing and Computational Intelligence Paradigms for the IoT* (pp. 46-59). IGI Global. https://doi.org/ doi:10.4018/978-1-5225-8555-8.ch004

Srivastav, M. K. (2021). Study of Risk Analysis in the Internet of Things. In N. Chowdhury & G. Chandra Deka (Eds.), *Multidisciplinary Functions of Blockchain Technology in AI and IoT Applications* (pp. 26–45). IGI Global., doi:10.4018/978-1-7998-5876-8.ch002

Srivastava, M. (2020). A Surrogate Data-Based Approach for Validating Deep Learning Model Used in Healthcare. In R. Wason, D. Goyal, V. Jain, S. Balamurugan, & A. Baliyan (Eds.), *Applications of Deep Learning and Big IoT on Personalized Healthcare Services* (pp. 132–146). IGI Global., doi:10.4018/978-1-7998-2101-4.ch009

Srivastava, P. R., Zuopeng (Justin) Zhang, & Eachempati, P. (2021). Deep Neural Network and Time Series Approach for Finance Systems. [JOEUC]. *Journal of Organizational and End User Computing*, *33*(5), 204–226. doi:10.4018/JOEUC.20210901.oa10

Stephens, R. T. (2019). A Requirements Approach for Building an Enterprise Cloud Service Catalog. In N. Rao (Ed.), *Global Virtual Enterprises in Cloud Computing Environments* (pp. 86–111). IGI Global., doi:10.4018/978-1-5225-3182-1.ch004

Strømmen-Bakhtiar, A. (2019). Digital Economy, Business Models, and Cloud Computing. In N. Rao (Ed.), *Global Virtual Enterprises in Cloud Computing Environments* (pp. 19–44). IGI Global., doi:10.4018/978-1-5225-3182-1.ch002

Stupar, S., Bičo Ćar, M., & Šahić, E. (2020). Application of Cloud Computing in Companies. In I. Karabegović, A. Kovačević, L. Banjanović-Mehmedović, & P. Dašić (Eds.), *Handbook of Research on Integrating Industry 4.0 in Business and Manufacturing* (pp. 75–100). IGI Global., doi:10.4018/978-1-7998-2725-2.ch004

Subashini, P., Krishnaveni, M., Dhivyaprabha, T. T., & Shanmugavalli, R. (2020). Review on Intelligent Algorithms for Cyber Security. In P. Ganapathi & D. Shanmugapriya (Eds.), *Handbook of Research on Machine and Deep Learning Applications for Cyber Security* (pp. 1–22). IGI Global., doi:10.4018/978-1-5225-9611-0.ch001

Sudarsanam, S. K., & Umasankar, V. (2019). Cloud-Based Design and Manufacturing. In N. Rao (Ed.), *Global Virtual Enterprises in Cloud Computing Environments* (pp. 202–230). IGI Global., doi:10.4018/978-1-5225-3182-1.ch008

Sudhakara, M., Kumar, K. D., Poluru, R. K., Kumar, R. L., & Bhushan, S. B. (2020). Towards Efficient Resource Management in Fog Computing. In S. Goundar, S. B. Bhushan, & P. K. Rayani (Eds.), *Architecture and Security Issues in Fog Computing Applications* (pp. 158–182). IGI Global., doi:10.4018/978-1-7998-0194-8.ch010

Suganya, R., Rajaram, S., & Kameswari, M. (2021). A Literature Review on Thyroid Hormonal Problems in Women Using Data Science and Analytics. In B. Patil & M. Vohra (Eds.), *Handbook of Research on Engineering, Business, and Healthcare Applications of Data Science and Analytics* (pp. 416–428). IGI Global., doi:10.4018/978-1-7998-3053-5.ch021

Sugiyama, S. (2019). *Artificial Intelligence in General*. IGI Global., doi:10.4018/978-1-5225-8217-5.ch001

Sultana, S., Akter, S., Kyriazis, E., & Wamba, S. F. (2021). Architecting and Developing Big Data-Driven Innovation (DDI) in the Digital Economy. [JGIM]. *Journal of Global Information Management*, *29*(3), 165–187. doi:10.4018/JGIM.2021050107

Sun, Z., & Stranieri, A. (2021). The Nature of Intelligent Analytics. In Z. Sun (Ed.), *Intelligent Analytics With Advanced Multi-Industry Applications* (pp. 1–21). IGI Global., doi:10.4018/978-1-7998-4963-6.ch001

Sundaram, S. M., & Murgod, T. R. (2022). Edge Computing for Secured IoT Analytics on the Cloud. In D. Jeya Mala (Ed.), *Integrating AI in IoT Analytics on the Cloud for Healthcare Applications* (pp. 162–175). IGI Global., doi:10.4018/978-1-7998-9132-1.ch010

Sundu, M., & Ozdemir, S. (2020). The Effect of Artificial Intelligence on Management Process. In Ahmad, N. H., Iqbal, Q., & Halim, H. A. (Ed.), Challenges and Opportunities for SMEs in Industry 4.0 (pp. 22-41). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-2577-7.ch003

T. S. S. M., & Sreeja, P. S. (2021). FNDNLSTM. In Sharaff, A., Sinha, G. R., & Bhatia, S. (Ed.), New Opportunities for Sentiment Analysis and Information Processing (pp. 218-232). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-8061-5.ch012

T. T. M., J., B., S., A. T., Jeswin Nallathambi, D., & Muthukumar, A. (2021). Deep Learning-Based Malware Detection and Classification. In Misra, S., Arumugam, C., Jaganathan, S., & S., S. (Ed.), Confluence of AI, Machine, and Deep Learning in Cyber Forensics (pp. 93-116). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-4900-1.ch006

Tali, R. V., Borra, S., & Mahmud, M. (2021). Detection and Classification of Leukocytes in Blood Smear Images. [IJACI]. *International Journal of Ambient Computing and Intelligence*, *12*(2), 111–139. doi:10.4018/IJACI.2021040107

Tanaka Kinoshita, R., & Datta, S. (2021). Possibilities for "Proactive, Interactive Deep Learning" in Japan's New National Curriculum Evident From an Analysis of IBDP's Theory of Knowledge. In D. G. Coulson, S. Datta, & M. J. Davies (Eds.), *Educational Reform and International Baccalaureate in the Asia-Pacific* (pp. 177–193). IGI Global., doi:10.4018/978-1-7998-5107-3.ch010

Tanque, M., & Foxwell, H. J. (2020). The Intersection of Data Analytics and Data-Driven Innovation. In M. Strydom & S. Buckley (Eds.), *AI and Big Data's Potential for Disruptive Innovation* (pp. 317–343). IGI Global., doi:10.4018/978-1-5225-9687-5.ch012

Tatnall, A., & Davey, B. (2019). Rise of the Non-Human Actors. In M. Spöhrer (Ed.), *Analytical Frameworks, Applications, and Impacts of ICT and Actor-Network Theory* (pp. 138–155). IGI Global., doi:10.4018/978-1-5225-7027-1.ch006

Tezcan, O., Akcay, C., & Gazioglu, B. (2019). A Review on BIM and Information Technologies Research in the Construction Industry. [IJDIBE]. *International Journal of Digital Innovation in the Built Environment*, *8*(2), 1–19. doi:10.4018/IJDIBE.2019070101

Thangaraj, R., Rajendar, S., & Kandasamy, V. (2020). Internet of Things in Healthcare. In P. Pankajavalli & G. Karthick (Eds.), *Incorporating the Internet of Things in Healthcare Applications and Wearable Devices* (pp. 23–39). IGI Global., doi:10.4018/978-1-7998-1090-2.ch002

Thilagavathi, N., Dharani, D. D., Sasilekha, R., Suruliandi, V., & Uthariaraj, V. R. (2019). Energy Efficient Load Balancing in Cloud Data Center Using Clustering Technique. [IJIIT]. *International Journal of Intelligent Information Technologies*, *15*(1), 84–100. doi:10.4018/IJIIT.2019010104

Thiyagarajan, P. (2020). A Review on Cyber Security Mechanisms Using Machine and Deep Learning Algorithms. In P. Ganapathi & D. Shanmugapriya (Eds.), *Handbook of Research on Machine and Deep Learning Applications for Cyber Security* (pp. 23–41). IGI Global., doi:10.4018/978-1-5225-9611-0.ch002

Thomas, J. J., Tran, T. H., Lechuga, G. P., & Belaton, B. (2020). Convolutional Graph Neural Networks. In J. J. Thomas, P. Karagoz, B. B. Ahamed, & P. Vasant (Eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics* (pp. 107–123). IGI Global., doi:10.4018/978-1-7998-1192-3.ch007

Thomas, J. J., Wei, L. T., Jinila, Y. B., & Subhashini, R. (2020). Smart Computerized Essay Scoring Using Deep Neural Networks for Universities and Institutions. In J. J. Thomas, U. Fiore, G. P. Lechuga, V. Kharchenko, & P. Vasant (Eds.), *Handbook of Research on Smart Technology Models for Business and Industry* (pp. 125–152). IGI Global., doi:10.4018/978-1-7998-3645-2.ch006

Tiwari, P. K., & Joshi, S. (2019). Resource Management Techniques to Manage the Load Balancing in Cloud Computing. In P. Raj & S. Koteeswaran (Eds.), *Novel Practices and Trends in Grid and Cloud Computing* (pp. 170–195). IGI Global., doi:10.4018/978-1-5225-9023-1.ch011

Tiwari, P. K., & Joshi, S. (2020). Dynamic Management of Resources in Cloud Computing. [IJSI]. *International Journal of Software Innovation*, *8*(1), 65–81. doi:10.4018/IJSI.2020010104

Tiwari, S. (2020). A Comparative Study of Deep Learning Models With Handcraft Features and Non-Handcraft Features for Automatic Plant Species Identification. [IJAEIS]. *International Journal of Agricultural and Environmental Information Systems*, *11*(2), 44–57. doi:10.4018/IJAEIS.2020040104

Tiwari, S. (2021). An Ensemble Deep Neural Network Model for Onion-Routed Traffic Detection to Boost Cloud Security. [IJGHPC]. *International Journal of Grid and High Performance Computing*, *13*(1), 1–17. doi:10.4018/IJGHPC.2021010101

Trovati, M., Asimakopoulou, E., & Bessis, N. (2021). Deep Learning Approaches in Pandemic and Disaster Management. In E. Asimakopoulou & N. Bessis (Eds.), *Data Science Advancements in Pandemic and Outbreak Management* (pp. 108–124). IGI Global., doi:10.4018/978-1-7998-6736-4.ch006

Turcu, C. O., & Turcu, C. E. (2019). New Perspectives on Industrial Engineering Education. In P. Sahoo (Ed.), *Optimizing Current Strategies and Applications in Industrial Engineering* (pp. 1–19). IGI Global., doi:10.4018/978-1-5225-8223-6.ch001

Tuzova, L. N., Tuzoff, D. V., Nikolenko, S. I., & Krasnov, A. S. (2019). Teeth and Landmarks Detection and Classification Based on Deep Neural Networks. In K. Kamalanand, B. Thayumanavan, & P. Jawahar (Eds.), *Computational Techniques for Dental Image Analysis* (pp. 129–150). IGI Global., doi:10.4018/978-1-5225-6243-6.ch006

Tyagi, A. (2021). Healthcare-Internet of Things and Its Components. In N. Wickramasinghe (Ed.), *Optimizing Health Monitoring Systems With Wireless Technology* (pp. 258–277). IGI Global., doi:10.4018/978-1-5225-6067-8.ch018

Tyagi, A. K., & Rekha, G. (2020). Challenges of Applying Deep Learning in Real-World Applications. In R. Kashyap & A. Kumar (Eds.), *Challenges and Applications for Implementing Machine Learning in Computer Vision* (pp. 92–118). IGI Global., doi:10.4018/978-1-7998-0182-5.ch004

Udendhran, R., & M., B. (2020). Demystification of Deep Learning-Driven Medical Image Processing and Its Impact on Future Biomedical Applications. In Suresh, A., Udendhran, R., & Vimal, S. (Ed.), *Deep Neural Networks for Multimodal Imaging and Biomedical Applications* (pp. 155-171). IGI Global. https://doi.org/doi:10.4018/978-1-7998-3591-2.ch010

Usman, U. M., Ahmad, M. N., & Zakaria, N. H. (2019). The Determinants of Adoption of Cloud-Based ERP of Nigerian's SMES Manufacturing Sector Using Toe Framework and Doi Theory. [IJEIS]. *International Journal of Enterprise Information Systems*, *15*(3), 27–43. doi:10.4018/IJEIS.2019070102

Uzaman, S. K., Khan, A. U., Shuja, J., Maqsood, T., Rehman, F., & Mustafa, S. (2019). A Systems Overview of Commercial Data Centers. [IJITWE]. *International Journal of Information Technology and Web Engineering*, *14*(1), 42–65. doi:10.4018/IJITWE.2019010103

Vaissnave, V., & Deepalakshmi, P. (2020). Data Transcription for India's Supreme Court Documents Using Deep Learning Algorithms. [IJEGR]. *International Journal of Electronic Government Research*, *16*(4), 21–41. doi:10.4018/IJEGR.2020100102

Valiveti, S. R., Manglani, A., & Desai, T. (2021). Anomaly-Based Intrusion Detection Systems for Mobile Ad Hoc Networks. [IJSSSP]. *International Journal of Systems and Software Security and Protection*, *12*(2), 11–32. doi:10.4018/IJSSSP.2021070102

Vambe, W. T., Chang, C., & Sibanda, K. (2020). A Review of Quality of Service in Fog Computing for the Internet of Things. [IJFC]. *International Journal of Fog Computing*, *3*(1), 22–40. doi:10.4018/IJFC.2020010102

Vanani, I. R., & Amirhosseini, M. (2019). Deep Learning for Opinion Mining. In R. Agrawal & N. Gupta (Eds.), *Extracting Knowledge From Opinion Mining* (pp. 40–65). IGI Global., doi:10.4018/978-1-5225-6117-0.ch003

Vanitha, N., & Ganapathi, P. (2020). Traffic Analysis of UAV Networks Using Enhanced Deep Feed Forward Neural Networks (EDFFNN). In P. Ganapathi & D. Shanmugapriya (Eds.), *Handbook of Research on Machine and Deep Learning Applications for Cyber Security* (pp. 219–244). IGI Global., doi:10.4018/978-1-5225-9611-0.ch011

Venkatraman, S. (2019). A Self-Learning Framework for the IoT Security. In J. J. Rodrigues, A. Gawanmeh, K. Saleem, & S. Parvin (Eds.), *Smart Devices, Applications, and Protocols for the IoT* (pp. 34–53). IGI Global., doi:10.4018/978-1-5225-7811-6.ch003

Verma, J. K., & Katti, C. P. (2018). Resource Request Based Energy Efficient Heuristic for Server Offloading in Cloud Computing Environment. [JGIM]. *Journal of Global Information Management*, *26*(4), 1–17. doi:10.4018/JGIM.2018100101

Verma, R. (2019). Smart Internet of Things (IoT) Applications. In K. Saravanan, G. Julie, & H. Robinson (Eds.), *Handbook of Research on Implementation and Deployment of IoT Projects in Smart Cities* (pp. 33–42). IGI Global., doi:10.4018/978-1-5225-9199-3.ch003

Verma, S., Chug, A., Singh, A. P., Sharma, S., & Rajvanshi, P. (2019). Deep Learning-Based Mobile Application for Plant Disease Diagnosis. In N. Razmjooy & V. V. Estrela (Eds.), *Applications of Image Processing and Soft Computing Systems in Agriculture* (pp. 242–271). IGI Global., doi:10.4018/978-1-5225-8027-0.ch010

Videc, T., Zoroja, J., & Pejic-Bach, M. (2020). What Business Leaders Should Know About Cloud Computing. In K. Sandhu (Ed.), *Leadership, Management, and Adoption Techniques for Digital Service Innovation* (pp. 264–286). IGI Global., doi:10.4018/978-1-7998-2799-3.ch014

Vidushi, A. M., Khamparia, A., & Khatoon, N. (2020). Wireless Environment Security. In Sagayam, K., Bhushan, B., Andrushia, A., & Albuquerque, V. C. (Ed.), Deep Learning Strategies for Security Enhancement in Wireless Sensor Networks (pp. 65-83). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-5068-7.ch004

Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2019). A Comparative Analysis of Deep Learning Approaches for Network Intrusion Detection Systems (N-IDSs). [IJDCF]. *International Journal of Digital Crime and Forensics*, *11*(3), 65–89. doi:10.4018/IJDCF.2019070104

Wang, T. (2019). Key Technology for Intelligent Interaction Based on Internet of Things. [IJDST]. *International Journal of Distributed Systems and Technologies*, *10*(1), 25–36. doi:10.4018/IJDST.2019010103

Wang, Z., Yang, J., Guo, B., & Cheng, X. (2019). Security Model of Internet of Things Based on Binary Wavelet and Sparse Neural Network. [IJMCMC]. *International Journal of Mobile Computing and Multimedia Communications*, *10*(1), 1–17. doi:10.4018/IJMCMC.2019010101

Weissbach, R., Herzfeldt, A. B., Floerecke, S., & Ertl, C. (2020). The Complex and Opaque Cloud Ecosystem. In S. Gochhait, D. T. Shou, & S. Fazalbhoy (Eds.), *Cloud Computing Applications and Techniques for E-Commerce* (pp. 134–158). IGI Global., doi:10.4018/978-1-7998-1294-4.ch007

Woolman, T. A., & Lee, P. (2021). Effects of Deep Learning Technologies on Employment in the Field of Digital Communication Systems. [IJIDE]. *International Journal of Innovation in the Digital Economy*, *12*(4), 35–42. doi:10.4018/IJIDE.2021100103

Wu, J., & Tien, C. (2020). *Deep Learning Theory and Software*. IGI Global., doi:10.4018/978-1-7998-1554-9.ch002

Wu, J., & Tien, C. (2020). *Handwriting 99 Multiplication on App Store*. IGI Global., doi:10.4018/978-1-7998-1554-9.ch004

J. Wu, & C. Tien (2020). *iOS App and Architecture of Convolutional Neural Networks*. IGI Global. https://doi.org/ doi:10.4018/978-1-7998-1554-9.ch001

Wu, J., & Tien, C. (2020). *Mobile-Aided Breast Cancer Diagnosis by Deep Convolutional Neural Networks*. IGI Global., doi:10.4018/978-1-7998-1554-9.ch006

Wu, J., & Tien, C. (2020). *Transformation Across Deep Learning Frameworks*. IGI Global., doi:10.4018/978-1-7998-1554-9.ch003

Yadav, N., & Badal, N. (2020). Introduction to IoT Technologies and Its Applications. In R. Singh, A. Gehlot, V. Jain, & P. K. Malik (Eds.), *Handbook of Research on the Internet of Things Applications in Robotics and Automation* (pp. 238–264). IGI Global., doi:10.4018/978-1-5225-9574-8.ch012

Yerpude, S., & Singhal, T. K. (2020). Value Enablement of Collaborative Supply Chain Environment Embedded With the Internet of Things. [IJIIT]. *International Journal of Intelligent Information Technologies*, *16*(3), 19–51. doi:10.4018/IJIIT.2020070102

Yfantis, V., & Ntalianis, K. (2020). The Exploration of Government as a Service Through Community Cloud Computing. [IJHIoT]. *International Journal of Hyperconnectivity and the Internet of Things*, *4*(2), 58–67. doi:10.4018/IJHIoT.2020070104

YILMAZ. B., Genc, H., Agriman, M., Demirdover, B. K., Erdemir, M., Simsek, G., & Karagoz, P. (2020). Recent Trends in the Use of Graph Neural Network Models for Natural Language Processing. In Thomas, J. J., Karagoz, P., Ahamed, B. B., & Vasant, P. (Ed.), Deep Learning Techniques and Optimization Strategies in Big Data Analytics (pp. 274-289). IGI Global. https://doi.org/ doi:10.4018/978-1-7998-1192-3.ch016

Youssef, F., El Habib, B. L., & Hamza, R., Labriji El Houssine, Ahmed, E., & Hanoune, M. (2018). A New Conception of Load Balancing in Cloud Computing Using Tasks Classification Levels. [IJCAC]. *International Journal of Cloud Applications and Computing*, *8*(4), 118–133. doi:10.4018/IJCAC.2018100107

Yu, P., & Xu, Z. (2021). The Impact of News on Public-Private Partnership Stock Price in China via Text Mining Method. In Z. Sun (Ed.), *Intelligent Analytics With Advanced Multi-Industry Applications* (pp. 264–286). IGI Global., doi:10.4018/978-1-7998-4963-6.ch013

Zairi, K. (2021). DeepLearning for Computer Vision Problems. In H. A. Bouarara (Ed.), *Advanced Deep Learning Applications in Big Data Analytics* (pp. 92–109). IGI Global., doi:10.4018/978-1-7998-2791-7.ch005

Zayed, N. M., & Elnemr, H. A. (2019). Deep Learning and Medical Imaging. In N. Bouchemal (Ed.), *Intelligent Systems for Healthcare Management and Delivery* (pp. 101–147). IGI Global., doi:10.4018/978-1-5225-7071-4.ch005

Zertal, S., Batouche, M., & Laboudi, Z. (2020). A Novel Hybrid Optimization-Based Approach for Efficient Development of Business-Applications in Cloud. [IJISSS]. *International Journal of Information Systems in the Service Sector*, *12*(4), 14–35. doi:10.4018/IJISSS.2020100102

Zhao, X., Jiang, Z., & Gray, J. (2020). Text Classification and Topic Modeling for Online Discussion Forums. In A. Fiori (Ed.), *Trends and Applications of Text Summarization Techniques* (pp. 151–186). IGI Global., doi:10.4018/978-1-5225-9373-7.ch006

Zheng, R. Z. (2020). Learning With Immersive Technology. In R. Z. Zheng (Ed.), *Cognitive and Affective Perspectives on Immersive Technology in Education* (pp. 1–21). IGI Global., doi:10.4018/978-1-7998-3250-8.ch001

Zhou, L., Yao, Y., Guo, R., & Xu, W. (2021). Core Technology of Smart Cities. In K. Lyu, M. Hu, J. Du, & V. Sugumaran (Eds.), *AI-Based Services for Smart Cities and Urban Infrastructure* (pp. 47–90). IGI Global., doi:10.4018/978-1-7998-5024-3.ch003

About the Contributors

P. Karthikeyan obtained the Bachelor of Engineering (B.E.) in Computer Science and Engineering from Anna University, Chennai, and Tamil Nadu, India in 2005 and received his Master of Engineering (M.E.) in Computer Science and Engineering from Anna University, Coimbatore India in 2009. He has completed Ph.D. degree in Anna University, Chennai in 2018. Skilled in developing projects and carrying out research in the area of Cloud computing and Data science with the programming skill in Java, Python, R and C. He published more than 20 International journals with good impact factor and presented more than 10 International conferences. He was the reviewer of Elsevier, Springer, Inderscience and reputed Scopus indexed journals. He is acting as editorial board members in EAI Endorsed Transactions on Energy Web, The International Arab Journal of Information Technology and Blue Eyes Intelligence Engineering and Sciences Publication journal.

Polinpapilinho F. Katina is an Assistant Professor in the Department of Informatics and Engineering Systems at the University of South Carolina Upstate (Spartanburg, South Carolina, USA). He has served in various capacities at the National Centers for System of Systems Engineering (Norfolk, Virginia, USA), Old Dominion University (Norfolk, Virginia, USA), Politecnico di Milano (Milan, Italy), and The University of Alabama in Huntsville (Huntsville, Alabama, USA). He holds B.S. in Engineering Technology, M.Eng. in Systems Engineering, and Ph.D. in Engineering Management and Systems Engineering (Old Dominion University, Norfolk, Virginia, USA). He received additional training at the Politecnico di Milano (Milan, Italy). Dr. Katina's profile includes nearly 200 scholarly outputs of peer-reviewed journal articles, conference papers, book chapters, and technical reports. He has also co-authored six (6) books. He focuses on teaching and research in the areas of Complex System Governance, Critical Infrastructure Systems, Decision Making and Analysis, Emerging Technologies (e.g., IoT), Energy Systems (Smart Grids), Engineering Management, Infranomics, Manufacturing Systems, System of Systems, Systems Engineering, Systems Pathology, Systems Theory, and Systems Thinking. He has demonstrable experience leading large-scale research projects and has achieved many established research outcomes.

S. P. Anandaraj received B.E, (CSE) degree from Madras University, Chennai in the year 2004, M.Tech (CSE) with Gold Medal from Dr. MGR Educational and Research Institute, University in the year 2007 (Distinction with Honor) and Ph.D. in August 2014.He is presently working as Associate Professor CSE in Presidency University, Bangalore. He has 15+ Years of Teaching Experience. His areas of interest include Information security, Data Science, Machine Learning and Networks. He wrote two book chapters in IGI Global Publications USA. He filled the 4 patents for the growth of research. He has published 50+ papers in various national and International Journals, national and International Conferences. He serving as an editorial member for reputed journals. He also attended many National Workshops/ FDP/Seminars etc. He is a member of ISTE, CSI, Member of IACSIT and Member of IAENG.

* * *

Anishka Bhatia is pursuing B.Sc. Hons. in Computer Science from Indraprastha College for Women, University of Delhi, Delhi, India. An open source contributor, she held 7th position overall India in MIT COVID Hackathon. Her research interests span the fields of Data Science and Machine learning.

Arulkumar Diana worked as Assistant professor, Department of computer Science and Engineering, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore, Tamilnadu, India. She received her M.Tech degree in Information Technology at Kalasalingam Academy of Research and Education, Anand Nagar, Tamilnadu, India in 2012 and Pursing Ph.D in the Department of Computer Science and Engineering at Kalasalingam Academy of Research and Education, Anand Nagar Tamilnadu, India. She has guided undergraduates and post graduates from her areas of interests such as Cyber Security and its threats, Network traffic and security, game theory and Vehicular adhoc network, Internet of Things and Biometrics and Deep learning. she has published papers and book chapter in SCI journals and Scopus indexed journals.

Menaka E. is Associate Professor of Department of Information Technology in Vivekanandha Engineering College for Women, Trichengode. She has obtained her B.E in Computer Science and Engineering from Manonmanium Sundranar University, Tirunelveli in 2002, M.E in Computer Science Engineering from Vinayaka Mission University in 2006 and obtained Ph.D in Information and Communication Engineering from Anna University, Chennai during 2017. She is a member of ISSE, ISTE and CSI. Her research areas include Image Processing, Internet of Things and Wireless Sensor Networks. She has 10 years of research experience in wireless sensor networks and IoT. She has published twenty papers in international conferences, twenty five papers in International journals. She has done many projects and hands on experience related to Internet of Things.

Arulkumar G. is working as assistant Professor in the department of computer science and engineering in Bule Hora University, Bule Hora, Ethiopia.

Pon Harshavardhanan is working as Associate Professor and Division Head, Data science division, School of Computing Science and Engineering VIT Bhopal University, Bhopal, Madhya Pradesh, India since June 2019. He was awarded with B.E. (CSE) from Bharathiar University, Coimbatore in the year 2001, M.E. (CSE) from Sathyabama Institute of Science and Technology, Chennai in the year 2006 and doctoral degree Ph.D. (CSE) from Anna University, Chennai. He has 20 years of teaching experience from engineering institutions and 2 years of industrial experience from IT industries. He has published 9 research articles in international and national journals. His research areas include semantic Web services, distributed systems, cloud computing, machine learning, artificial intelligence, and data science.

Kamatchi Kartheeban is working as Associate Professor, Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India. Also he was worked as a Deputy Director Academic. He received his M.E degree in computer science and Engineering from College of Engineering, Anna University, and Chennai in 2007 and Completed PhD in the title of "Development of Efficient Algorithms for Secure Communication in Distributed Computing Environment" in the Department of Computer Science and Engineering at Kalasalingam Academy of Research and Education, Anandnagar, Tamilnadu, India in 2014. He worked as a faculty from Adhiyamaan College of Engineering, Hosur, India between 1996-1998. Currently 1 scholar completed his PhD and 6 students are doing their PhD under him with the topics such as Internet of Things, Medical Image Processing, Video Analytics, Cyber Forensics, Sentimental Analysis and Scheduling in Cloud computing. He has published many papers in SCI journals and Scopus indexed conferences. Also he has submitted proposals to DeIT, SERB and DRDO. His areas of interests include IoT, Medical image processing, cryptography and bioinformatics and grid and cloud computing.

Sarabjeet Kaur Kochhar was awarded a Ph.D. Degree in Computer Science from Department of Computer Science, University of Delhi, Delhi, India. She is an Associate Professor in the Department of Computer Science, Indraprastha College for Women, University of Delhi, Delhi, India. She has published extensively in international journals and conferences. Her research interests are currently aligned along with the fields of Data Mining, Data Analytics, and Natural Language Processing.

Babu N. is working as an Assistant Professor at Siddharth Institute of Engineering and Technology located at Puttur, Andhra Pradesh. He has 15 years experience in the field of teaching. He graduated B.E. CSE from Anna University, Chennai then completed his M.E. CSE from St. Peter's Institute of Education and Research, Chennai and currently pursuing his Ph.D. at St. Peter's Institute of Education and Research, Chennai. He has 4 publications in National and International journals and conferences. His area of interest includes Internet of Things, Machine Learning and Big Data Analytics.

Mohamed Imtiaz N. is working as an Assistant Professor in the Department of Information Science and Engineering at HKBK College of Engineering, Bangalore. He obtained his Bachelor's degree (B.Tech) in Information Technology from Anna University, Chennai (2007), Master Degree (M.Tech) Information Technology from Sathyabama University (2011) and currently pursuing his doctoral degree in Anna University, Chennai. His research focuses on Cloud Computing, Big Data Analytics, Artificial Intelligence.

Rajagopal R. received the Bachelor of Engineering(B.E) in Computer Science and Engineering) from Anna university, Chennai, Tamilnadu, India, in 2005 and received his Master of Engineering (M.E,) in Computer Science and Engineering from Anna University, Coimbatore, India in 2009 and PhD degree in Anna University, Chennai, in 2018.He was the reviewer of Inderscience and reputed Scopus indexed journals. He acts as an editorial board member in international journal of mathematical science, Engineering, and Sciences Publication journal. He published more than 22 International journals and presented more than 20 International conferences.

Surya S. received the BE degree and M.E degree in Computer Science and Engineering from Anna University Chennai. She is currently working towards the Ph.D degree in the Department of Computer Science and Engineering at SRM Institute of Science and Technology, Chennai. She has worked as a Assistant Professor of CSE in various institutions since 2011.She is currently working as Assistant Professor of CSE in Saveetha Engineering College, Tamilnadu, India. She has published more than 5 papers in both International and National conferences and Journals. Her research interests include Machine Learning, Big Data Analytics, Deep Learning, and IOT.

Anu Sayal is currently working as Assistant Professor in Mathematics in JAIN (deemed to be University), Bangalore, Karnataka. She has done her schooling from Convent of Jesus and Mary, Dehradun. She has done Bachelors in Non-medical from Guru Nanak Dev University, Amritsar. Her Post Graduation was in Mathematics from Guru Nanak Dev University, Amritsar. She was the college topper in Mathematics in Graduation. She has completed her PhD in Mathematics from UTU Dehradun. She has won the "Young Scientist Award in Mathematics: Statistics and Computer Science from Uttarakhand Council for Science and Technology, Dehradun, Uttarakhand, during the 14th Uttarakhand State Science and Technology Congress 2019-20 held from 27th – 29th February 2020 at UCOST, Vigyan Dham, Dehradun. She has also received Teacher of the Year 2020 Award on 5th September 2020 from Divya Himgiri in association with Uttarakhand Council for Science & Technology, Government of Uttarakhand, Commission for Scientific & Technical Terminology, MHRD, Govt. of India, New Delhi. She received Guruwarya Samman 2021 award from Harvest educational transformational solutions (HETS). She also received Best woman scientist award from Novel research academy on 6th December 2021. She is also awarded as the Best teacher on 9th January 2022 by Harvest educational transformational solutions. She has authored numerous research papers in reputed and prestigious journal (including Scopus & SCI), presented papers in many national and international conferences. She is serving as a reviewer and editorial board member for various international journals (Scopus indexed). She has also attended various workshops and FDP's. Her research interest includes fuzzy mathematics, statistics, numerical techniques, inventory, and supply chain management.

Rajendran Thanikachalam has more than 13 years of teaching and research experience and currently working as an Associate Professor in the Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai. He received his Ph.D. from Anna University, Chennai and M.E. Computer Science and Engineering from St. Peter's University, Chennai. His research interests are Network Security, Ethical Hacking, Cloud Computing, Image Mining, and Semantic Ontology. He has published 15 articles in international journals and 16 publications in national and international conferences which are indexed in Scopus and SCI. He has published 2 patents in national level. He is a reviewer for reputed International Journals and has been a session chair for various International and national Conferences.

Nandini Tomar is pursuing B.Sc. Hons. in Computer Science from Indraprastha College for Women, University of Delhi, Delhi, India. She has been an active contributor in Machine learning based research projects. She is also working on Speech Recognition Project involving Data Mining and Machine learning.

K. Vyshnavi is pursuing B.Tech - Computer Science and Engineering from Kakatiya Institute of Technology and Science, Warangal.

Index