International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 11 | November - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Big Data Analytics Applications and challenges in Real Time

Dr. CK Gomathy.Assistant Professor in CSE, SCSVMV
Ms. Sai Nandini Peesapati- Final Year UG Scholar, CSE Dept, SCSVMV.
Mr.SG Sri Goutam- Final Year UG Scholar, CSE Dept, SCSVMV.
Ms.Rakshita K - Final Year UG Scholar, CSE Dept, SCSVMV.
Ms.Swarna Dharshini S- Final Year UG Scholar, CSE Dept, SCSVMV.
Mr. Nanda Pavan Kumar T-Final Year UG Scholar, CSE Dept, SCSVMV.
,

Abstract:
The phrase "Big Data" was used to describe a large amount of data that cannot be managed using typical data handling methods or procedures. In numerous industries, including agro, banking, data mining, education, chemistry, finance, cloud computing, marketing, and health care stocks, the field of big data is indispensable. Big data analytics is a technique for analysing large amounts of data to find hidden patterns, mysterious relationships, and other crucial information that can be used to make better decisions. Because of its rapid development and wide range of applications, there has been a steadily growing interest in big data.

Key words: *Big data, Analytics, Healthcare, weather-forecast, Education, Government, Hdfs*

1. INTRODUCTION

Big Data opens significant opportunities in every corner of the planet in almost every company and industry, viz. stock exchange , agriculture, telecommunications, healthcare, and education. This ample opportunity comes with big challenges and issues. likelihood is increasing as the volume of Big Data is also increasing and is predicted to grow enormously because of the technological revolution, which incorporates but is not limited to various mobile devices. The character of big data using use cases, real-time analysis, and data integration, eventually turns big data into an enormous value. The pressing issues identified during this paper are privacy, processing and analysis, and storage. During this paper, we explored various usages of massive Data and methodologies in Big Data.

Here are some samples of real-time uses for big data:

- Big Data Analytics in HealthCare,
- Big Data Analytics in meteorology ,
- Big Data Analytics in Education,
- Big Data Analytics in Government Sector, and
- Big Data in Media and Entertainment.

2. BIG DATA ANALYTICS IN HEALTHCARE

Healthcare organisations are being transformed by big data and will be in the near future. Big Data Analytics (BDA) is regarded as a vital instrument by scientific literature in the management context applied to healthcare organisations, to the point that it has drawn the interest of the scientific community and stakeholders. A foundational assumption should be made, though: data by themselves don't explain anything. As a result, in order for data to be relevant in the administration of healthcare organisations, their quality must first be verified, followed by the identification of the proper correlations. To put it another way, the data should be handled, examined, and analysed using the proper software.

A new era has begun, one in which machines will contribute to the development and management of human knowledge as well as no longer being the sole domain of humans. This tremendous data product potential, along with IoT applications and AI capabilities, will play a big role in the near future for pharmaceutical businesses and healthcare organisations. Today's IoT-based medical applications enable clinical data monitoring through the production of data collected by specialised equipment (such wearable devices) that are remotely accessible by a doctor rather than by caretakers.A new era has begun, one in which machines will contribute to the development and management of human knowledge as well as no longer being the sole domain of humans. This tremendous data product potential, along with IoT applications and AI capabilities, will play a big role in the near future for pharmaceutical businesses and healthcare organisations. Today's IoT-based medical applications enable clinical data monitoring through the production of data collected by specialised equipment (such wearable devices) that are remotely accessible by a doctor rather than by caretakers.

After an in-depth reading of abstracts and full papers, studies selected through a content analysis were clustered into 4 research areas (RAs):

- Potentialities of BDA (RA1)
- Resource management (RA2)
- BDA and management of health surveillance system (RA3)
- BDA technology for healthcare organisation (RA4)

The proposed clustering has been thought to give an easy-to-go research map and to support the healthcare managers.

*2.1 **RA1: Potentialities of BDA***

When applied to the healthcare industry, Wang and Hajli [1] describe BDA potentialities as "the ability to acquire, store, process, and analyse large amounts of health data in diverse forms, and offer meaningful information to users, allowing them to quickly uncover business values and insights." The "route to value chain" notion of the relationship between BDA and the advantages for healthcare companies has been clearly described [1]. This route represents a significant contribution to the study of business value because it empirically demonstrates how capabilities can be built and what benefits can be realised in healthcare organisations, in addition to drawing the general and well-established connection between big data capabilities and the benefits.The critical importance of BDA skills in creating healthcare supply chain integrations and its effects on hospital flexibility are the subject of another study in this area [2]. With regard to enhancing supply chain and operational flexibility and healthcare integration, the BDA is particularly important. Given the economic and health issues brought on by the Covid-19, managers have embraced this BDA factor as a particularly effective tool to increase the operational flexibility of healthcare companies. A key perspective of the BDA for assisting managers and healthcare professionals in their decision-making is its capacity to deliver real-time insights and prediction models.Regarding this, the literature offers a number of big data applications in healthcare that aid in the gathering, management, and integration of data in healthcare organisations [3]. Additionally, BDA makes it possible to integrate sizable datasets, supporting managers' decisions and keeping an eye on the managerial facets of healthcare companies. Finding the big data keys that may be used to apply ad-hoc strategies to increase efficiency along the healthcare value chain is the first step in developing a decision-making process based on BDA. In order to do this, Sousa et alresearch .'s [3] highlights the advantages that BDA may provide to the decision-making process through predictive models and real-time analytics, aiding in the gathering, administration, and integration of data in healthcare organisations.

*2.2 **RA2: Resource Management***

The good effect of the BDA on resource management emerged from the literature review as another significant study direction. The management of medical waste, energy consumption, and environmental load is inadequate, which limits resource conservation. The BDA is very helpful in this regard; it might make a significant contribution in the near future to the implementation of the circular economy processes and to the support of efforts for sustainable development in the healthcare organisations [4]. To this purpose, the research conducted by Kazançolu et al. [4] emphasises the significance of circularity and sustainability ideas in order to reduce the damaging environmental effects of the sector. The report also outlines the challenges the healthcare organisation faces in implementing the circular economy and offers solutions by putting BDA-based management systems in place.

In order to support healthcare management, many academics are focused on BDA-driven decision support systems nowadays [5]. These BDA-based analytical tools will offer administrators of healthcare companies helpful quantitative support. Using case studies, the authors have described the technical specifications and design elements of the system implementations. They have created a toolset that functions as a framework for resource management, enabling them to design strategic models and acquire analytical outcomes for deliberations that are informed by facts and managerial evaluations.

As previously stated, a successful implementation of BDA in the healthcare organisation will be essential to enhancing clinical outcomes management, providing useful information for managers and decision-makers to help prevent diseases, lower healthcare costs, and enhance the efficiency of the healthcare organisation [6]. To accomplish these lofty goals, the research will need to solve a complex problem: how to make heterogeneous data from various sources easily usable and economical. The study conducted by Kundella and Gobinath [6] makes a significant contribution to our understanding of the major difficulties, solutions, and emerging issues with BDA use in the healthcare industry, including privacy concerns, security algorithms, and future directions.

*2.3 **RA3: BDA and Management of Health Surveillance System***

Numerous healthcare issues in underdeveloped nations may be resolved as BDA gains popularity. The BDA was used in healthcare organisations to aid managers in resource allocation and the health system in providing patients with better care [7]. In this sense, the Zambian government is considering deploying BDA technologies to deliver healthcare services that are more effective and efficient. In developing nations, where severe resource shortages impede economic growth, a well-managed health surveillance system is a key factor in raising living standards and reducing medical waste. In order to provide new knowledge, enhance clinical treatment, and manage the public health surveillance system, Europe is investing in BDA efforts in the public health and oncology sectors [8]. More information has also been provided about the effective deployment of an improved public health surveillance system based on early detection, tracking contacts, and patient follow-up and assistance [9]. The administration and comprehension of COVID-19 were improved by Serrana's upgraded surveillance system. It was able to track SARS-CoV-2 positive cases and variants, follow the pandemic trajectory, provide patient care, and create pertinent research initiatives by merging community and academic institutions [9].

2.4 **RA 4: BDA Technology for Healthcare Organisation**

In addition to becoming the foundation of customised medicine, wearable technology and various types of sensors that can gather clinical data will also be essential instruments for enhancing the efficiency of healthcare organisations [10]. Another research area looked into by the articles included in this RA was the relationship between BDA and IoT technologies as a tool to incorporate the accessibility, ability to personalise, and practical conveyance of clinical data.These tools enable healthcare organisations to reduce costs, people to self-regulate their own treatments, and practitioners to make judgments as rapidly as possible while working remotely and being in continual communication with their patients [11]. These findings allow us to conclude that the Internet of Things (IoT), big data, and artificial intelligence as machine-learning algorithms are three of the most important breakthroughs in the healthcare sector. These firms are putting in place clever BDA systems based on machine learning and home-centric data collection networks.The authors emphasise the significance of combining IoT, big data, and artificial intelligence as methods to improve community health outcomes and the efficiency of healthcare systems. The effectiveness of public health interventions can be increased using the standardised data sets produced by these sources by the new generation of machine learning algorithms [12]. To this purpose, it is imperative that the next research focus R&D efforts toward fully standardised dataset processes, as noted by various studies in the field of BDA applied on healthcare companies.

3. BIG DATA ANALYTICS IN WEATHER

Weather forecasting, a crucial and vital process in people's everyday lives, assesses the change taking place in the atmosphere's current state. Massive data analytics is the practise of studying big data to uncover hidden patterns and useful information that might produce more beneficial outcomes. This essay offers a way for doing a systematic review of the literature for big data analytical techniques in weather forecasting (published between 2019 and August 2022). The currently evaluated studies are categorised as using technique-based, technology-based, and hybrid techniques in a workable way.

Weather predictions were first made in the nineteenth century [13, 14]. Weather forecasting is defined as the process of analysing atmospheric data, such as temperature, radiation, air pressure, wind speed, wind direction, humidity, and rainfall. There must be a large amount of data generated or collected in order to anticipate the weather. These data are also not well arranged. As a result, using meteorological data to predict the weather is a difficult undertaking since it has too many variables. These factors alter based on the rapidly changing weather conditions. We should take into account the specific properties of weather forecasting, such as continuity, data intensity, and multidimensional and chaotic behaviours [14, 15], before we suggest an algorithm.Originally a human-intensive endeavour, weather forecasting has evolved into a computational process [6], necessitating high-tech machinery. The accuracy of projections can be impacted by a number of things. Some of these important variables are the time of year, geographic location, data quality, classifications of the weather, lead time, and validity period [16, 17]. Big data is a term used to describe unstructured, diverse, and vast digital data [17]. Big data processing cannot be accomplished easily or effectively using traditional data management techniques [13, 14]. In order to efficiently analyse this type of data, we should look for a high-performance platform and a practical big data mining technique [17].Big data analytics is the process of searching through large amounts of data for hidden patterns, undiscovered relationships, and other pertinent information to aid in decision-making [12]. Big data can improve the decision-making process when it comes to weather forecasting [13]. We may employ big data analytics in weather forecasting to do this because it is essential to make exact predictions. In the past, forecasting relied heavily on human forecasters, but in the information era, it is now carried out using technology and data [14].Rainfall, humidity, air pressure, radiation, sunlight intensity, data collecting, and other variables make up atmospheric data sets. We also require a massive amount of datasets acquired from diverse sources (big data). High-tech technology and software are needed to process this amount of data [15,16].

3.1 Weather Big Data Analytic Factors

Each weather big data analysis approach has some evaluation factors containing QoS and weather factors for comparing the prevailing methods and solutions among recommended solutions. The important QoS and weather factors are defined as follows: • Accuracy, precision, and recall: so on evaluate an enormous data method, to match it with earlier methods to figure out a more efficient algorithm, and also to urge its pros and cons, we'd like QoS factors. Precision, accuracy, and recall are sort of these factors used in various articles. Within the next , we represent the foremost important QoS factors for large data analytics in weather forecasting: Accuracy, precision, and recall are defined as (1), (2), and (3). Accuracy is the share of correct prediction. The right positive prediction percentage is called precision. the share of occurrences predicted as positive is known as recall [14]. Additional factors utilised in articles for evaluation and comparison are execution time (second), the reaction time (second), scalability, reliability (a number between 0 and 1 that indicates the general consistency of a measure), Related Works and Motivation Some research about big data analytics are done. This section presents some survey papers within the massive data analytics feld and the descriptions of their important factors and limitations. Sahasrabuddhe and Jamsandekar [15] first introduced meteorology , including basic processes and different approaches, then big data. They then explained the kind of data structures used for big data and weather forecasting with an overview of a number of papers. This study didn't have a classification, a search methodology, and a discussion, and this paper wasn't a scientific review. [16] described meteorology , big data, and rain forecasts. The author reviewed several papers with the matter of predicting rainfall using data mining techniques. The author summarised them during a table then began to expound the papers. The limitation was explained. This paper didn't have a research methodology, the reviewed works, taxonomy, and discussion sections, so it had not been a scientific review. Hassani and Silva [18] administered one of the big data analytics studies on forecasting. They introduced various sorts of challenges related to forecasting with big data, including noisy data detection, hardware and software, statistical significance, the architecture of existing algorithms, and massive data itself. Also, the extant applications of statistical and processing techniques for prediction were discussed. To supply the researchers with a more satisfactory experience, the authors summarised and categorised the reviewed studies supporting the related fields, including economics, finance, population dynamics, crime,energy environment, bioscience , and media.

3.2 FORECASTING FRAMEWORK

The framework of weather prediction includes data acquisition, data preprocessing, model selection, and training, evaluation and visualisation of results. paper size. This data consists of both useful and useless information, which they're in the form of unstructured data. Once the info is gathered, the primary step is to preprocess the data to remove missing values and to clean the data. Data preprocessing tries to strengthen the quality of the input data. At this stage data is ready for training. It is often caused because of values that are noted and due to data corruption. As each value is validated, missing values can cause errors to the model, so these missing values are replaced with mean values in most cases. The info integration process merges data from different sources into a data store. Data reduction is employed to reduce the size of the data by eliminating unwanted features or clustering. Within the data transformation step, the info is normalised using standardisation techniques to convert suitable for processing. Data transformation techniques also are useful in reducing the training time of the models. Preprocessing isn't mutually exclusive, hence they'll function together. Once the pre-processing is completed an appropriate model is chosen , trained and tested using the datasets. Model selection and training may be a crucial step in any forecasting system. The knowledge about sorts of forecasting models will help researchers to select a suitable model matching the application domain is selected and trained using the datasets. After completion of coaching , the model performance is evaluated using statistical quantitative error indicators like MAE, MAPE, RMSE and R^2. Atlast the results are visualised using suitable plots. Plots like scatter plot, line plots are often used to visualise the results. These plots are wont to analyse the difference between actual and predicted values. The difference between the anticipated values and actual values are often more precisely viewed in semilog plots.

3.3 Technique-Based Approaches

A technique-based approach in big weather data analytics often focuses on the quality of the forecasts, which depends on the data quality in several levels of its life cycle. Most of these approaches try to discover the hidden knowledge in the observed data.Most of the studied papers used Non-heuristic algorithm and machine learning as their evaluation types and different tools for modelling the approaches. The main advantage of the technique-based approach is gaining high accuracy, while high execution time is its disadvantage. The approach is used in weather forecasting methods.

3.4 Technology-Based Approaches

A technology-based approach, which is predicated on big data ecosystems, tools, frameworks, etc., mostly tries to proportion and speed up weather data analysis in order to improve weather forecasting. The important testbed is the mainly used evaluation type; big data tools are used for modelling the approaches, and most of them didn't use any algorithms. The most advantageous technology based approaches are gaining low time, including execution time and reaction time , and their main disadvantage is low reliability. The technology-based solutions for giant data analytics in weather forecasting are utilised to improve scalability and time consumption via applicable approaches. the prevailing evaluation factors in technology-based approaches.

3.5 Hybrid Approaches

We review the chosen hybrid approaches. Commonly, complex predictive analysis can't be done by a basic approach, so two of them are merged to realize a value added approach. A hybrid approach may be a combination of big data techniques and technologies.A global climate change detection algorithm based on spatial cumulative sum (CUSUM) was proposed by Manogaran and Lopez. The CUSUMs of every sample value deviation from the target value (average) were determined by a CUSUM control chart. The massive day-wise climate data was stored in a distributed manner on HDFS.The hybrid approaches for giant weather data analysis were utilized, and both mentioned approaches were combined. So hybrid approaches proportion and speed up weather forecasting mechanisms and improve the quality of model outputs.The real testbed is the mainly used evaluation type, big data tools are used for modelling the approaches, and most of them used non-heuristic algorithms.

4. BIG DATA ANALYTICS IN EDUCATION

The subject of education is another one in which big data analytics is extremely important. It serves a variety of functions or is used in a variety of ways.

4.1 Interactions with Dyslexia students in the UK

It can be utilised, for instance, by those who struggle with dyslexia. Diverse interpretations[19] and definitions of the dyslexia concept are reflected in the use of background criteria, which is a contentious diagnostic problem. According to research and expert observation, not all higher education students who request a dyslexia diagnosis have a documented or even self-reported history of literacy or study skills issues. In their undergraduate and even graduate courses, many brilliant people experience their first academic challenges.

4.1.1 Guidelines for HE dyslexia diagnostic assessment

1. Background Information
2. Attainments in literacy
3. Underlying ability
4. Cognitive processing

4.1.2 Methods

1. Participants
2. Materials and procedures
   2.1 Data Collection
   2.2 Data Analysis

4.1.3 Results

The vague descriptive language provided by the Equality Act of 2010 does not make defining what is meant by disability in the HE sector any easier. Tasks that scream out for more precise criteria include determining where a "mental impairment" starts and stops, as well as how "substantial" its "adverse effect" has to be.

Assessors would do well to distinguish between assessments for statutory disability entitlement and assessments for intervention if there is one advice for policy and practise to come out of this study (Arnold, 2017).

Despite the fact that the former could point out a variety of functional challenges and helpfully use the term "dyslexia" in a brief descriptive manner (Bishop, 2012; Bishop Blog, 2014; Hulme & Snowling, 2009; Nicolson, 2016; Norwich, 2009; Ramus, 2014; the latter would necessitate a mutually agreed-upon, precise, and operational definition. Future research would benefit greatly from developing such an eligibility criterion for the HE sector as well as the standards to support it.

4.1.4 Conceptualizing the big data analytics among staff in Malaysia using TOE framework

Understanding an organization's readiness to adopt new technologies may help you better grasp how well you can keep up with market needs in your environment. Because of this, collective preparation at the organizational and individual level is necessary for any change management process inside an organization to be successful.

The TOE framework was created by [20], and it outlines three factors—technological, organisational, and environmental—that have an impact on how a technological innovation is adopted and implemented. All three factors were used for this study since they were considered appropriate and were heavily weighted in the literature when determining an organization's readiness. The technological part examines whether the organization's current skill infrastructure and technological capabilities align with the BDA. The organisational aspect examines top management's commitment to the implementation of BDA as well as a few other factors, such as the availability of a policy and legal framework and finance. The environmental aspect examines how the organisation interacts with outside parties regarding BDA.

4.1.5 Methodology

Study is predicated on a few underlying philosophical presumptions regarding what qualifies as credible research and the best research methodologies. The term "research paradigm" refers to these presumptions. Generally, there are three types[21] of study paradigms: positivist, interpretive, and critical. The way a research project was designed, how data was gathered and evaluated, and how the research findings were presented all reflect the researcher's beliefs during the research process. The interpretive paradigm was used in this investigation.

4.1.6 Research Setting

The National University of Malaysia is the only higher education institution included as a case study for this study. The case study approach is thought to be suitable for investigating modern phenomena using empirical research often connected to comprehending a particular topic (i.e., the state of BDA readiness). Based on the researcher's affiliation with the organisation, the research environment was chosen.

The documents served as a means of triangulating the claims that respondents would subsequently make in the interviews. Ten employees from different divisions within the university were interviewed in-depth for the data. To schedule the interviews, invitation emails were sent, followed by phone calls. Before the interview started, the respondents were asked if they would voluntarily participate in the study and allow their voices to be recorded.

Three components made up the interview questions: Section 1 on demographic data, Section 2 on staff awareness and knowledge of BDA, and Section 3 on BDA readiness with regard to the eight ODRA-adapted dimensions. The interviews, which ranged in length from 60 to 90 minutes on average and were audio recorded and transcribed, were done in a mixed-language setting (English and Malay). Thematic analysis was then done. The relevant ODRA dimensions and the TOE framework to which they related were used to organize the data chronologically and categorically. The data were then regularly evaluated and continuously coded.

4.1.6.1 Results

The general conclusion of BDA in operations is that it is relatively constrained to the usage of already-existing relational (structured) data, particularly data pertaining to students and academic programmes. The practical application of big data, namely enormous unstructured data from non-traditional sources like video or social media, is still under development. Additionally, because some of the academicians from the ICT and Sciences faculties are directly involved in the BDA teaching and research, they are more familiar with and exposed to BDA.

4.1.7 Saudi Arabian Higher Education

Developing countries are aware of the benefits of newly developed technology for increasing economic value. According to the Communication and Information Technology Commission, Saudi Arabia is one of these developing nations that is spending money on innovative technologies like the internet of things, agile innovation, and big data analytics (CITC 2016). In Saudi Arabia, particularly in the educational sectors, big data analytics (BDA) is regarded as one of the developing technologies. Consequently, this technology might boost enterprises' ability to realize value.

4.1.7.1 Theoretical Background

1. Socio-Technical Theory
2. Delone and Mclean IS Success Model

4.1.7.2 Technological Factors(Big Data Analytics Tasks- BDATs)

1. ***Storing Big Data***

   Big data has major implications for the company's capacity to develop adequate storage that can support quick access to many types of data due to its high volume, high velocity, and high variety characteristics (Watson 2014). Having the ability to store huge data from diverse sources could help senior management teams make decisions (Alghamdi 2016).

2. ***Analyzing Big Data***

   Big data analysis and storage frequently go hand in hand. Businesses increasingly double down on their big data management tools to assist data integration and analysis for improved information outputs in addition to storing large data (Magdanz 2014). For instance, colleges in Saudi Arabia are exploring the potential of data analysis for decision assistance.

3. ***Visualizing Big Data***

   As businesses increasingly use diverse business intelligence and data visualisation tools to improve the management and use of big data, data visualisation is the next significant step in big data analytics jobs[22]. It has been demonstrated that visualisation is useful for both conveying crucial information in large volumes of data and for guiding difficult analyses (Keim et al. 2013, p. 20).

In contrast to BD storage and analysis, BD visualisation represents the pinnacle of improved decision making because it offers a captivating and engaging means of doing so. BDATs (i.e., storage, analysis, and visualisation) will be positively correlated with system quality, according to propositions 1a–c.

4.1.7.3 Human Factors (Big Data Analytics Performers-BDAPs)

Academic and Non-Academic Information Technology Staff *:*

Other studies contend that IT personnel have an impact on the security of an information system (Betz 2016). Additionally, BD systems are vulnerable to the affects of the IT professionals that use and manage these systems, much like other IT systems (Russom 2011). In this study, academic and nonacademic staff members are important human factors for safeguarding big data in Saudi Arabian universities as well as making sure that BD privacy is being maintained. The system's performance in terms of BD security and privacy.

5. BIG DATA ANALYTICS IN GOVERNMENT SECTOR:

The public sector has a lot of promise with big data. Every day, massive volumes of data are generated and collected by a government's operations, including managing social benefits, collecting taxes, keeping an eye on the national health and education systems, capturing traffic statistics, and producing official papers. Governmental departments and agencies may identify areas that require attention, make more informed decisions more quickly, and implement essential changes thanks to information that is easily accessible in real time.

Big data and analytics can be used by governments to access important information and increase efficiency and transparency in public administration. By enhancing their use of IT, learning about new capabilities and services, identifying correlations and patterns in underlying data, and improving decision-making, government agencies can gain from using big data platforms.Government organisations are starting to employ big data technologies to analyse enormous data sets for science and research as well as mining data to stop terrorist actions from being committed by bad actors as well as to stop waste, fraud, and abuse.

Studies chosen through a content analysis were grouped into 4 research areas (RAs) after a thorough reading of their full papers :

1. Public sector - RA1
2. Taxation - RA2
3. Railways - RA3
4. National security - RA4

5.1 RA1 - PUBLIC SECTOR :

[23]Chief information officers are laying the groundwork to correlate and track dependencies across people, processes, and information to ensure that the right information is available to the right person at the right time as government leaders work to create data-driven organizations to successfully accomplish missions. Agencies must adopt a deliberate approach to where they store all of this data and on the platform that enables the easiest access when necessary. More departments, functions, use cases, goals, and most importantly, a stronger emphasis on making wise decisions and adding value. Big data analytics is ultimately what it is.

5.1.1 Big Data in Government and Society :

Recognize the escalating significance of big data in modern society and investigate big data's theory and application from an arts and humanities standpoint. Beyond the enormous data sets that may be examined for patterns, trends, and relationships, it is becoming more and more relevant to talk about our daily lives and how the data we produce is changing social, cultural, political, and economic processes as well as how knowledge is produced.

The first case study uses Madhya Pradesh, the largest state in India, a low-income state with a high proportion of previously excluded groups, to highlight two data-focused projects that have increased transparency throughout society and enhanced service delivery. The Chief Minister's Dashboard, which combines data from many sources to encourage evidence-based decision making, and the Chief Ministers Helpline, a centralized portal for registering complaints about all elements of government operation, are these.

5.1.2 Big Data for Policy Making :

Data's ability to lower ambiguity about the optimal course of action in policy design, guide a better policy making process, and result in more adequate, efficient, and successful public policies has come to more people's attention recently. Consequently, officials and policy advocates frequently use data to support their claims for certain policy solutions. Big data technologies have thus given decision-makers the chance to gain more in-depth, data-driven insights into the relevant problems and have made it possible for quantitative analysis to more thoroughly permeate the decision-making process than ever before. And because of this technical reality, there is now a chance and a need for a more sophisticated, complicated, and technologically driven method to turning data into policy action that goes beyond conventional empirical research.

5.1.3 Big Data for Sustainable Development :

Society is already being transformed by the data revolution, which includes the open data movement, the growth of crowdsourcing, new information communication technologies for data collection, and the explosion in the availability of big data. It is also accompanied by the emergence of artificial intelligence and the Internet of Things. Big data processing and analysis can now be done in real time thanks to developments in computing and data science. Official statistics and survey data can be supplemented with fresh insights obtained by data mining, giving information on human actions and experiences more depth and richness. This new information should be combined with existing data to provide high-quality, timely, and relevant information.Consumer profiling, customized services, and predictive analysis are now prevalent in the private sector and utilized for marketing, advertising, and management. Similar methods could be used to monitor people's wellbeing in real time and concentrate aid on the most vulnerable populations. If used responsibly, new data sources including satellite data, new technology, and new analytical methods can facilitate more swift, effective, and evidence-based decision making and improve how the Sustainable Development Goals (SDGs) are being achieved.

5.2 RA2 - TAXATION :

Authorities on taxes will be able to check information's accuracy or spot potential fraud with the aid of data from social networks and other sources. This list might not be exhaustive, and some people or businesses might have evaded IT's notice. Predictive analytics would enable the government to identify fraud and increase collection. Big data technology allows us to effectively gather data and Internet information, accomplish effective information management, and construct a full set of tax databases.[24]

5.2.1 Impact of Big Data on Corporate Tax Management:

Numerous businesses are doing well in the setting of the rapid economic expansion, and numerous businesses have adopted various tax management ideas. However, most businesses do not give tax administration in the financial sector adequate consideration, and businesses lack expert talent. Paying close attention to financial and tax management will not only have an undesirable impact on the business's daily operations but will also have some negative long-term implications. For the growth of the business, the management of the organization must be very conscious of financial risk.The business can, to a certain extent, avert many financial disasters. The system also organizes the gathered data. Of course, they must also possess strong professional abilities and the ability to employ intelligence.

5.2.2 Recommendations for Tax Management in the Context of Big Data :

The primary goal of an intelligent data analysis platform should be to gather a lot of tax-related data. Following the gathering of pertinent data, the platform should compare the data and establish a benchmark value to forecast tax crises. Those who exceed the standard value will receive stern reminders from the e data system, helping to prevent certain problems with the impending tax crisis. The demands on managers in tax management are very high. They must have good professional abilities in addition to understanding some legal concepts linked to taxation. Companies should pay more attention to professional abilities and cultivate and train more to increase the skills of employees since they will use numerous clever apps. The main priority in enterprise management is enterprise tax management. It affects not just how the business will develop in the future, but also how profitable the business will be. Therefore, the management of the company's finances and taxes should be given high importance. To give a good assurance for the financial administration of the organization, businesses should draw up pertinent rules and regulations with a clear division of labour.An organization's tax management is a labor-intensive operation. The procedure of each phase should be clearly explained by managers, who should also complete the preliminary tax accounting work, compute the correct amount, report it, and perform quality levelling.In compliance with legal requirements, businesses should pay taxes and fees on time, and employees should contact tax authorities promptly to avoid issues.

5.3. RA3 - RAILWAYS :

There are three main Railway transport system areas-maintenance, operations and safety-which have been benefited by BDA .

5.3.1 BIG DATA IN RAILWAY NETWORK MANAGEMENT:

[25]Big data analytics in the railway industry are still in their infancy. Decision-making for infrastructure repair based on available condition monitoring data as well as operational performance and train timetabling data are key applications of big data analytics within railway infrastructure. Integration of data from disparate sources is one of the data processing requirements for the collection of decision-making data, where low latency requirements for decision-making need to be fulfilled through parallel architectures.Among these is track circuit status monitoring, where data flows may be triggered on train movements through track sections in brief intervals of a few seconds or longer. Other strategies include gathering data on the state of the rolling stock for condition analyses to produce a plan for maintenance once the rolling stock arrives at a destination.[26]A summary of some big data characteristics relevant to railway network management is shown in the figure below.

Fig 1: Big Data in Railway Network Management

Through big data analytics, which continuously feed real-time information into a system, railroads are better able to respond to unforeseen concerns including unscheduled delays, critical maintenance, and accidents. Some of these advantages include freight cost savings of up to 25%, an increase in asset utilization of 5–15%, a decrease in inventory, and a better flexibility to add capacity.

5.4 RA4 - NATIONAL SECURITY:

Homeland Security, Defense, Public Safety, and Intelligence agencies are increasingly using big data and data analytics, mostly due to the world becoming more digital and linked. This is opening up new possibilities for intelligence processing, exploitation, dissemination, and analysis in addition to data gathering and storage. In many pertinent areas, such as the fight against crime and terrorism, defence against cyberattacks, public safety analytics, disaster and mass incident management, and the creation of predictive capabilities, big data and data analytics technologies can improve the investigative capabilities of intelligence organisations.[27]

5.4.1 Military big data:

Military big data is the application of big data in the military area, and its scope includes all data that can be used in the military field both directly and indirectly. Big data technology can effectively increase the utilization efficiency of equipment quality information, use the accumulated quality information, and pre-measure the equipment management and security requirements in peacetime and wartime. This is especially true of the equipment information management work. This is based on improving the effectiveness of equipment support and the accurate guarantee of organizational equipment. In addition, big data technology has shown the benefits of effective data fusion, processing, and security management, which may support military operations and combat equipment support with precise and effective prediction and decision-making references.

5.4.2 Equipment Quality Information Management and necessary of using military big data

The gathering of actual tactical technology status data that is embodied in the process of development, manufacture, and usage is known as equipment quality information. It serves as the framework for planning the application of equipment quality management and serves as a crucial foundation for the usage, upkeep, support, and advancement of equipment design. The management of equipment quality information has evolved into the foundation and a requirement for the growth of equipment quality work activities as a result of the constant advancement of information technology in the field of equipment construction in our military. Massive amounts of equipment quality information have led to "information fog" among equipment management workers, which complicates the management of equipment quality in various ways. It may successfully address the issue of vast amounts of information by processing and analyzing military big data in an effective, scientific, and rigorous manner and its challenging to handle.

6. CONCLUSION

Nowadays the challenge for healthcare organizations is the development of useful applications based on BDA. According to the circular economy view, the future research directions should be addressed considering the relationship between digitalization and management resources consumption. The data centralization combined with a BDA approach can effectively support circular economy processes in the healthcare supply chain by reducing waste and resource consumptions.

Future studies will need to focus on developing more efficient models for sharing data in order to improve the performance of healthcare organizations around the world.

The vacuity of Big Data, low- cost commodity tackle, and new information operation and logical software have produced a unique moment in the history of data analysis. The confluence of these trends means that we've the capabilities needed to dissect astonishing data sets snappily and bring- effectively for the first time in history. These capabilities are neither theoretical nor trivial. They represent a genuine vault forward and a clear occasion to realize enormous earnings in terms of effectiveness, productivity, profit, and profitability.

The Age of Big Data is then, and these are truly revolutionary times if both business and technology professionals continue to work together and deliver on the pledge.

7. REFERENCES

1. Wang Y Hajli, N.,. Exploring the path to big data analytics success in healthcare. J Bus Res. 2017;70:287–99.
2. Yu W, Zhao G, Liu Q, Song Y. Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility: An organizational information processing theory perspective. Technol Forecast Soc Change. 2021;163:120417.
3. Sousa MJ, Pesqueira AM, Lemos C, Sousa M, Rocha Á. Decision-making based on big data analytics for people management in healthcare organizations. J Med Syst. 2019;43:290.
4. Kazançoğlu Y, Sağnak M, Lafcı Ç, Luthra S, Kumar A, Taçoğlu C. Big Data-enabled solutions framework to overcoming the barriers to circular economy initiatives in healthcare sector. Int J Environ Res Public Health. 2021;18:7513.
5. Moutselos K, Maglogiannis I. Evidence-based public health policy models development and evaluation using big data analytics and web technologies. Med Arch (Sarajevo, Bosnia and Herzegovina). 2020;74:47–53.
6. Kundella S, Gobinath R. A survey on big data analytics in medical and healthcare using cloud computing. Int J Sci Technol Res. 2019;8:1061–5.
7. Chellah RC, Kunda D. An assessment of factors that affect the implementation of big data analytics in the Zambian health sector for strategic planning and predictive analysis: a case of Copperbelt province. Int J Electron Healthc. 2020;11:101–22.
8. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, Boccia S. Benefits and challenges of big data in healthcare: an overview of the European initiatives. Eur J Public Health. 2019;29:23–7.
9. Natasha N Ferreira, Pedro M M Garibaldi, Glenda R Moraes, José C Moura, Taline M Klein, Larissa E Machado, Lilian F B Scofoni, Simone K Haddad, Rodrigo T Calado, Dimas T Covas, Benedito A L Fonseca, Ricardo Palacios, Monica T R P Conde, Marcos C Borges, The impact of an enhanced health surveillance system for COVID-19 management in Serrana, Brazil. 2022.
10. Clim A, Zota RD, Tinica G. Big data in home healthcare: A new frontier in personalized medicine. Medical emergency services and prediction of hypertension risks. Int J Healthc Manage. 2019;12:241–9.

11. Bharathi MJ, Rajavarman VN. A survey on big data management in health care using IOT. Int J Recent Technol Eng. 2019;7:196–8.

12. Martinez F.E.L, Núñez-Valdez E.R, et al. Big data and machine learning: a way to improve outcomes in population health management. In: González García C, et al., editors. Protocols and applications for the industrial internet of things. Hershey: IGI Global; 2018. p. 225–39.

13.N. Anusha, M. Sai Chaitanya, Guru Jithendranath Reddy (2019). Weather Prediction using Multi Linear Regression algorithm. IOP Conf. Series: Materials Science and Engineering 590 012034.

14. Cadenas, E. Rivera, W. Campos-Amezccua, R. Heard, "Wind speed prediction using a univariate ARIMA model and multivariate NARX model", Wnergies 9(2), 109.

15. Kashiwao, K. Sarigul, M. Avci, M. Kandirmaz, T. Nakayama, K. Ando, S. Ikeda, K. Lee, M. Bahadori, A neural network based local rainfall prediction system using meteorological data on the internet: a case study using data from Japan meteorological agency.Appl. Soft Comput. 56, 317-330

16. Sahil Ratra (2020). Weather Prediction using Machine Learning Techniques.

17. Sindhu P. Menon, Ramith Bharadwaj, Pooja Shetty, Prajwal Sanu, Sai Nagendra (2017). Prediction of Temperature using Linear Regression. International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT).

18. Liu, X., Zhang, C., Liu, P., Yan, M., Wang, B., Zhang, J. and Higgs, R., 2018. Application of temperature prediction based on neural network in intrusion detection of IoT. Security and Communication Networks, (2018).

19. Denise Ryder, Brahm Norwich. What's in a name? Perspectives of dyslexia assessors working with students in the UK higher education sector. (2018)

20. L.G Tornatzky, and M.fleischer, The Process of Technological Innovation. Massachusetts:Lexington Books, 1990.

21. Mohamad Taha Ijab; Seri Manja Abdul Wahab; Mohd Azull Mohamad Salleh;Azlina Abu Bakar. Investigating Big Data Analytics Readiness in Higher Education Using the Technology-Organisation-Environment (TOE) Framework. (2019)

22. Maher Mohammed Aseeri, Kyeong Kang. Technological and Human Factors for Supporting Big Data Analytics in Saudi Arabian Higher Education. (2020)

23.Jung Wan LEE1. Big Data Strategies for Government, Society and Policy-Making.(2020).

24. Accounting School/Finance and Accounting Research Center, Fujian Jiangxia University, Fuzhou 350108. Application of Big Data Model in Financial Taxation Management . (2021) .

25. P. MCMAHON , T. ZHANG , AND R. DWIGHT. Requirements for Big Data Adoption for Railway Asset Management. (2020)

26. Faeze Ghofrania, Qing Hea,b,∗, Rob M.P. Goverdec, Xiang Liud. Recent applications of big data analytics in railway transportation systems: A survey (2018) .

27. Zheng Xiang, Liu Xiaofang, Gao Weigang. Analysis of the Application of Military Big Data in Equipment Quality Information Management. (2019)

28.Dr.C K Gomathy & et.al, A Developing Industrial Symbiosis In Large Cities Using Big Data, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

29.Dr.C K Gomathy & et.al, Big Data Analysis Of Airline Dataset Using Hive, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

30.Dr.C K Gomathy & et al, THE CUSTOMER DATA ANALYSIS USING SEGMENTATION WITH SPECIAL Reference Mall, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 9 | Sep 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

31.Dr.C K Gomathy & et al, THE INTEGRATING E-GOVERNANCE WITH BIG DATA ANALYTICS USING APACHE SPARK, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

32. Dr.C K Gomathy & et al, THE TWITTER BEHAVIOURAL ANALYTICS, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

33.Dr.C K Gomathy & et al, THE EFFICIENT MANAGEMENT FOR MALICIOUS USER DETECTION IN BIG DATA,COLLECTION., International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 05 Issue: 10 | Oct- 2021, ISSN: 2582-3930.

34.Dr.C K Gomathy & et al, ODD EVEN REDUCES TRANSPORT POLLUTION, Journal Of Engineering, Computing & Architecture Volume: 12 Issue: 03 March - 2022 , ISSN:1934-7197

35.Dr.C K Gomathy & et al, INSURANCE CLAIM ANALYSIS USING VEHICLE GPS DATA, Journal Of Engineering, Computing & Architecture, Volume: 12 Issue: 03 March – 2022

36.Dr.C K Gomathy & et al, INSTAGRAM AUTOMATION TOOL , Journal Of Engineering, Computing & Architecture, Volume: 12 Issue: 03 March - 2022 ISSN:1934-7197

37.Dr.C K Gomathy & et.al, Big Data-Driven Abnormal Behavior Detection In Healthcare Based On Association Rules, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072

Author's Profile

1. Sai Nandini Peesapati, Final Year Student, B.E. Computer Science and Engineering, Sri Chandrasekharendra SaraswathiViswa Mahavidyalaya deemed to be university, Enathur, Kanchipuram.

2. Raksitha K, Final Year Student, B.E. Computer Science and Engineering, Sri Chandrasekharendra SaraswathiViswa Mahavidyalaya deemed to be university, Enathur, Kanchipuram.

3. Sri Goutam SG, Final Year Student, B.E. Computer Science and Engineering, Sri Chandrasekharendra SaraswathiViswa Mahavidyalaya deemed to be university, Enathur, Kanchipuram.

4. Swarna Dharshini Final Year Student, B.E. Computer Science and Engineering, Sri Chandrasekharendra SaraswathiViswa Mahavidyalaya deemed to be university, Enathur, Kanchipuram.

5.  Nanda Pavan Kumar T, Final Year Student, B.E. Computer Science and Engineering, Sri Chandrasekharendra SaraswathiViswa Mahavidyalaya deemed to be university, Enathur, Kanchipuram.

6.  Dr. C.K. Gomathy, Assistant Professor in Computer Science and Engineering at Sri Chandrasekharendra Saraswathi Viswa Mahavidyala, Enathur, Kanchipuram, India. Her area of interest: Software Engineering, Web Services, Knowledge Management and IoT