The International Journal of Interdisciplinary Organizational Studies
ISSN: 2324-7649 (Print), ISSN: 2324-7657 (Online)
Volume 19, Issue 1, January-June, 2024

USE OF BIG DATA ANALYTICS TO POWER AI (ARTIFICIAL INTELLIGENCE) AND ML (MACHINE LEARNING) AUTOMATION

***Krovvidi S B Ambika**
Department of CSE, Aditya college of engineering and technology

ABSTRACT

Big data now forms a crucial component of professional and academic research as well as their areas of interest. The expansion of the Internet and of digital gadgets is the cause of data growth. The bigger data set include real-time structures, temporal structures, and structures that were collected from other sources. Predictive approaches propose strategies for extracting private information from enormous data sets. Numerous possibilities are available with this strategy, including the top optimization techniques or detection patterns. In the end, we presented the direction of future study in this field and concluded our studies. Artificial intelligence research and development is advancing quickly, and the combination of AI and automation has already begun to alter the corporate environment. To reach new heights of efficiency and quality, businesses are concentrating on merging current artificial intelligence with automation processes. The article illustrates artificial intelligence and automation, and it seeks to show the audience how the two are connected, how they can be more productive when used in tandem, and how they may provide a competitive edge.

Keywords: Big Data, Big Data and Analytics, Data Science, Predictive Analytics, Advanced Analytics, AI.

INTRODUCTION

Large data analysis has emerged as a major area of study for both teachers and students in recent years. Massive amounts of data are shown all day long on tablets, smartphones, and the expanding Internet. big data from numerous data sources in varied formats, as opposed to traditional data. For individuals who manage a data centre, the size, diversity, and speed of the data provide a special difficulty. The likelihood of computing, preserving, and analysing has, nonetheless, come to these concerns. Large data storage has become widespread and affordable. The corporation is aware of the importance of these social networking sites like Twitter or Facebook based on conventional business statistics. These technologies demonstrate the ability to gather the commercial information required to create a competitive strategy. In this post, we discussed how to construct a firm such that managers may use information to make choices, which can improve organisations' returns on investment. The notion of "Big Data," "Predictive Analysis," and "Big Data Analysis Program," as well as challenges and possibilities, are presented in the article. The industry has a great lot of opportunity to handle upcoming business challenges with the community and provide special advantages as required.

Artificial intelligence is now extensively employed throughout a range of corporate sectors and industries, from manufacturing to law to healthcare and finance. Doctors are now more quickly

able to identify illnesses thanks to machine learning. Applications like chatbots are assisting patients with billing and assisting consumers with appointment scheduling. AI in education may provide automatic grading, assist students in learning by meeting their needs, and keep them on course. In the legal sector, artificial intelligence (AI) has made it simpler for the attorneys to efficiently and correctly review thousands of massive legal papers, which is a rather tiresome task in general. Manufacturing is now more simpler and more effective than it was only a few years ago thanks to industrial robots.

CONCEPT OF AUTOMATION

The previous ten years have seen a surge in the usage of automation, which aims to save time and labour. Automation has replaced a system that was created by merging man and machine with a system of computers and machines. With the use of automation in numerous sectors, very difficult and repetitive operations have become effective, and the product quality has also improved.

DIFFERENCES BETWEEN ARTIFICIAL INTELLIGENCE (AI) AND AUTOMATION

It's crucial to recognise that both artificial intelligence and automation are phrases that are sometimes used interchangeably before continuing with this discussion. They are linked to machines that help us operate more productively and effectively, such as real or software robots. This may be done mechanically by putting something like an automobile together or by sending a follow-up email the day after you discover that your consumer hasn't finished placing their purchase. But what many miss is that there are also significant distinctions between the two. These variations reflect the degree of complexity in both systems. These differences are as follows:

- **Terms of Difference:** Automation [9] is the process of creating technology or software that can carry out tasks automatically and without the need for human involvement. On the other hand, artificial intelligence involves both science and engineering to create intelligent machines. Making robots resemble or even strive to outperform human intellect and behaviour is the goal of artificial intelligence (AI).
- **Data:** Artificial intelligence may or may not be used in automation. Between the first and third industrial revolutions, the whole practise of [4] automation underwent a transformation that led to its present state. Automatic testing, mechanical labour, control systems, computers, and operational equipment are all used in the manufacturing process. All of the many forms of automation that have appeared all around us are constrained by clear rules and programming. All that is required to make the same thing an AI is to feed it with data, according to [10]. The programme must include enormous amounts of data, including those from neural networks, graphs, and deep machine learning. How much you can make your system stimulate like a person will depend on your skill of coding. However, it's probable that you'll be imparting to the machine all you already know. If automated, sensor readings will allow you to quickly determine the output. However, much as with human brains, there is always some degree of ambiguity with AI.
- **Purpose:** Automated systems may carry out repeated activities. This allows individuals to engage in more significant tasks that call for logical judgement and thinking, freeing up

crucial time [5]. This increases the system's effectiveness and efficiency. Artificial intelligence is designed to not only look for patterns but also learn from past mistakes so that it can choose the best reaction for any given circumstance.

BIG DATA AND INTELLIGENT AUTOMATION

There are several methods for entering huge data. Large data is fantastic for conventional data usage, according to Waller and Fautchett, thus novel solutions are required to address these problems. [1]. Several tools and programmes, such Chen and Aldefin, are needed to manage and maintain the analysis and visibility of the[2] significant and sophisticated applications. Fan et al made the assumption that this was accessible for acquiring information [3]. Larger data may be chosen not only for the amount of the data but also for how quickly it can be processed throughout and how it relates to Kizi's nature [4]. As a result, huge data refers to five-byte data created in a short amount of time, such as a few hours, and setup data for minor things. Work is made simple by intelligent automation. Everything that you find useful, including new services, technologies, business models, partnerships, ecosystems, and more, may be integrated.

Automation may help you save money and deliver analytical information that is more precise, quick, and plentiful. One individual may do several swallowing-related duties. The current home automation (HA) system may make use of data collecting and processing. As fresh data is collected, the Sensor HA system has a lot of data. Therefore, it's critical to distinguish between crucial and superfluous data. These four steps—identifying business concerns, creating papers, evaluating, putting into practise, assessing, fixing, fixing, and measuring solutions—are crucial in establishing the ultimate course of action.

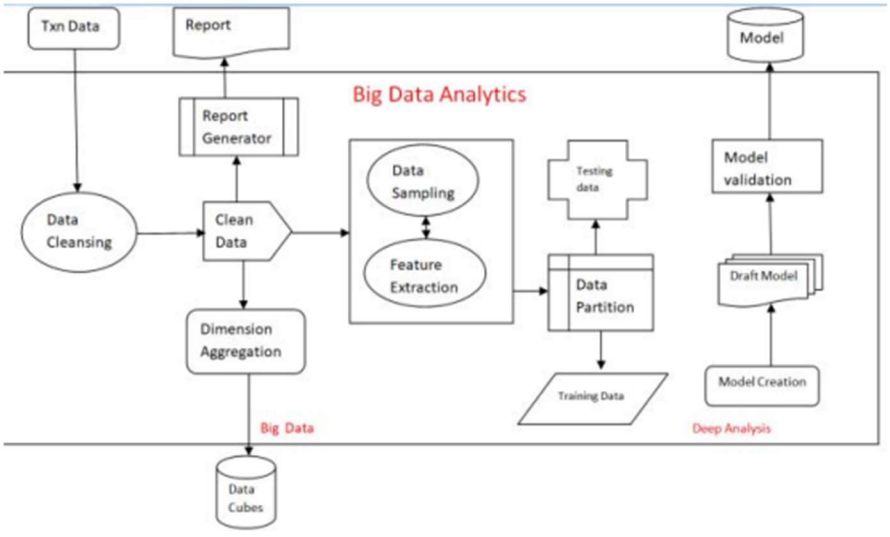

Figure 1: Overview of Big Data and Analytics

TYPES OF DATA TYPES

Informal and informal data are produced by many sources, and they may be categorised accordingly.

- **Structured data:** Structured data comes from organisational information systems, which include enterprise extensions like SCM systems and data from ERP processing. The database's table structure for this data is well established. The data has been analysed and arranged. Data about uncertainty is gathered from a variety of sources, including media, news, emails, files, videos, and audio. Over 80% of the primary automobiles are the quickest and fastest. Organizational data is the foundation for the study and development of this data.
- **Semi-database:** This is derived from several sources and includes both organised and unstructured data. To guarantee the quickness, effectiveness, and correctness of information interchange among stakeholders, information systems are deployed.

BIG DATA VS TRADITIONAL DATA

In general, business data are in decent shape. Over the last several decades, a lot of significant firms have used the ERP system to increase their efficacy and success [6]. These ERP apps or platforms gather data every 24 hours. They assist in making decisions on interim and short-term strategy. Data size and speed are influenced by the size of the firm. The nations participating are the only ones whose data are being collected. There are numerous circumstances when more data is gathered. each ten years. These instances are common, yet they are much too few. Despite the fact that the Internet serves as the foundation for communication and information, various data circumstances exist worldwide: In multinational corporations where worldwide communication is quick, simple, and transparent, access to the Internet has encouraged global collaboration and growth. Recently, the Internet has helped expand other businesses. To carry out all organisational and distribution tasks, the corporation is continually growing its IT system.

The updated information may be seen in its formal format by clicking on the blog blog on the blog. Social networking information on social media has also contributed to a steady growth. Data size and speed. Data originates from several sources without being permanently stored or coming from sources that have been updated. obtaining tens of thousands of songs per day from Facebook [4]. The likelihood of manual analysis is outweighed by the volume, variety, and speed of data acquisition. Keeping local data is sometimes very necessary. To handle the massive volumes of data needed for this large-scale research, industry giants like Microsoft, IBM, and Oracle employ specialised technologies like Hadoop [2][21].

ANALYTICS AND AUTOMATION

Analyzing a thorough data analysis of the prior data analysis approach, including utilisation methods such a transparency analysis with intelligent data flow and real-time analysis using sensor data [2]. It is a method for gathering and keeping track of business intelligence data to discover this phenomena by examining the model's development to forecast future outcomes based on variables, modelling, or optimization, and artificial intelligence based on changes in parts (eg travel analysis). use of statistical techniques like mathematical analysis of variables and events and statistical analysis of statistical analysis Levalle et al investigated value perception and data analysis. They said that due to advancements in technology, there is currently no better method to

examine this data. [9] The major challenge today is to discover answers to questions like what is occurring rather than worrying about what occurred and why it is known as mentioned in the inspection. Forward-looking analysis is the process of determining what actions should be made based on current and future occurrences. They are classified as analytical and are best suited for outcomes. So, as illustrated in Figure 2 below, business analysis may be divided into three categories: description, forecasting, and analysis. The need for coding and programming will decline in the age of genius and with the advancement of automation. Information and data will be assessed, and the need for analytical documentation will continue to be the first step in the processing of data for scientific purposes and unquestionably in the creation of any solutions for scientific data. With additional possibilities, company development potential will grow. a combination of data science and mathematics that may be used to tackle certain business issues. In order to enhance the overall performance of the company, this will need the broad use of the AI algorithm, which should be regarded as the hybridization of data.

A. Predictive Analytics

Data exploration for the default approach, statistical image recognition, machine learning, artificial intelligence, and data mining are all included in this procedure. [11] This entails using analytical data to respond to inquiries or resolve issues, much as a high-level analytical method [10]. Statistical techniques are integrated with the further growth of intelligent business (BI) and data extraction. Business intelligence aids in data analysis so that decision-makers in the business world may make wise choices. Professionals in the field of education have created questions, and data analysts have forecasted a variety of patterns and linkages. [11] The computer develops algorithms for models and related connections based on independence throughout this methodical analytical procedure. In order to reduce sample mistakes, it is made to find a better regression coefficient. To discover the best solutions for this issue, this technique makes use of sophisticated information systems.

The foundation for practise analysis based on event logs now comes from the production process. This allows for comparisons and introduces new regional initiatives like banking or retail health [12]. The Hadoop framework offers a response to these scaling requirements. There are several analytics strategies that support and evaluate data mining, depending on the source and kind of data. Many developing-world industries have used the text analysis approach, including those that analyse, index, sense, analyse, predict, and download data from article sources [2]. Since this product's problems began to surface in recent years, social media data and emotional index analysis have provided customers with immediate feedback on this product.

B. Sentiment Analysis

Many businesses record and use consumer sentiment research to examine client opinions posted on social media about their goods or services. The usage of English natural process algorithms in various companies and sectors may be assessed in a number of different ways. Some websites also

provide consumers with concepts for goods or services. A variety of services are available via Google Analytics, including Intelligent Advertising (based on effective and efficient advertising), Deeper Customer Understanding (CRM contact centre customer call centre - phone), etc.

C. Competitive Intelligence

Today's brand and company executives seek consumer input on their goods as well as those of their rivals in a market that is so fiercely competitive. Kim alludes to the process of gathering data via a system of social analysis about the items and services offered by their rivals in the same market. [14] The link between the diversity of market demand for the product and the emotions that surface in social networks is shown by an analysis of cellphones like the iPhone 2 and Samsung Galaxy S5. They looked at three factors: I the extent of social networks, (ii) the client's mood for more psychoanalysis, and (iii) the desire to purchase. There are various methods for a firm to acquire competitive intelligence via social networks to assist with learning about suppliers, the competitive environment, and general business trends.

D. Marketing and Brand Promotion Strategy

More and more individuals are producing very large data sets and geographic regions as a result of the adoption of smartphones and tablets [15]. The internet provides several chances for business relationships, including news, search queries, e-commerce websites, and social networking sites. Before making a purchase, consumers constantly search for goods and services on the Internet and on social media. The expense and danger of using a certain product determine the research level. These companies are able to understand the nature of the activity, assess pertinent data, and design their marketing efforts appropriately.

OPPORTUNITIES

Larger sizes and sizes determine larger data [3]. This kind of large-scale, erratic data analysis shows some receptivity to the difficulty of managing these enormous amounts of data. We had a detailed debate as a result. The previous art had an issue with a lot of information, and that problem is now the scope [19]. Numerous potential and competitive benefits are provided by big data. The first version of Amazon gathered data on users' interests, purchases, and book reviews. On the basis of this information, suggestions for services that assist clients in extending their purchases of the same product or comparable items have been made. Retailers of the future will be able to monitor each customer's activity and identify patterns of influence. Bigger data has numerous advantages over conventional data gathering techniques, such as templates.

A. Data Mining

Finding information from a dataset that aids in future result prediction is a difficult procedure. Finding models in the population and spotting looming differences with little data is helpful (Fan and al 2014). This is essential for acquiring accurate but ambiguous information.

B. Large Sample Size for Analytics

Large amounts of data produce a massive dome, which improves test findings [19]. The alterations in the team are explained by the illustrative method in which, apart from an example, another citizen may result in one model being erroneous owing to a lack of input parameters. We examine the complete data set, including everything, inside the enormous data. This gives us greater insights into the complex relationships between the model's dependent and independent variables as well as into the data sheet's many different components. The present analyst can also analyse vast amounts of data affordably. The data model delivers true information resulting in better business choices when the data is reflective of the complete population.

C. Technological Process

To search the same vast database of data for contact information, significant data from analytical data is required. [17]. The primary data problems have prompted improvements in computer system infrastructure and new storage techniques. An algorithm for optimising efficiency produces the highest point in statistical, scientific, and mathematical exercises and data, and it may be further extended to greater data for higher volumes [3].

D. In Real Time

Whether people who may acquire this data source can get particular information from the data from sources. Important connections to the power supply may be destroyed by natural disasters. Effects, particularly for nations like the influence of a more advanced Indian transport system in India, where the devastated infrastructure is insufficiently young players from the top [20]. Information gathered on climate, natural catastrophe development, and geographic location (floods, earthquakes, storms, etc.).

CHALLENGES WITH BIG DATA ANALYTICS

The large amount of data presents a unique issue. Most organisations control the data that is being gathered at such a rapid rate. Important issues are those involving size, appetite quality, waste, imbalance, misbehaviour, unintentional occurrences, and incorrect measurement [3]. One important point to remember is that enormous data is governed by massive amounts of data that are created everyday and are now more affordable than ever. Major data restrictions including inconsistency, relevance, and inefficiency must be addressed for statistical processes to be successful [3][22].

E. Size and Quality of Data

Size is a major barrier to large data resolution. Using various ETLs in the information system, local databases may be easily handled. For specialist systems, like Hadoop for handling data, basic data is required. Large data must be sent and stored with extreme accuracy since they are rapidly created. The quality of the information, which may be completely resolved via accuracy and timeliness, is the other issue.

F. High Size

As a consequence of the high intensity and excellent correlation, several factors may skew the sample's findings. Any variable that doesn't account for the outcomes' fluctuation may be included into the template. The choice of variables and the design are problematic because of the likelihood of erroneous references resulting from the magnitude of the data. This is because a sample of a bigger size yields an incorrect result.

G. Reliability

Another issue is the reliance on data, which is often false, susceptible, and lost. Such information may be found in a number of places, including social media, mobile devices, email, and messaging [5]. Data from many sources, including data systems and other sensors, are produced in the area of data production. This results in the extraction process requiring vast amounts of information rather than the emergence of discrete, pertinent information.

H. Completeness of Data

Despite being large, unstructured, and multilateral, the database does not represent a population. If we consider social media as a dataset without a significant media structure, we may share our opinions or have a discussion about the problem. A subset of data called data sets is accessible on certain websites. When analysing data, researchers need to be ignored.

PERFORMANCE ANALYSIS

The enhancement of the organization's broader data analysis project has a unique problem, just like any information systems project. There are some issues with conducting the organization's analysis due to the existence of skilled scientists and the inclusion of the essential infrastructure to modify the business environment. [10] Scientists will require scientific, data-based, and predictable analyses to improve the quality of crucial industrial data. This area requires significant subject knowledge and data, including statistics, predictions, optimization, and template construction [1]. Clear techniques and procedures are necessary to comprehend the underlying problems from several sources (legacy systems, ERP systems, social networks). Based on knowledge, expertise, and recording, a scientist is required. The Superior Team should evaluate the first project. After a Contextual Strategy for Business Context has been implemented successfully, its application may only be made to the other parts of the company (Davenport, 2014).

CONCLUSION

Digital recordings of social and business events started to become a significant source of information from 2000. Big Data Analysis is an anticipated area that integrates expertise in the fields of computer science and quantitative statistics. The media and social media provide data. Understanding and analysing potential outcomes is crucial. Predicting this capacity increases one's ability to strategize in advance of formidable rivals. By assessing emotional indices via participation in contests or by utilising social media companies to promote their intelligence, social media analysis has the capability to get first-hand knowledge about their goods from the market. as a venue. The kind of massive data has particular opportunities and problems.

REFERENCES

[1]. Waller, M. A., & Fawcett, S. E. (2013). Waller, "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management." Journal of Business Logistics 34.2 (2013)

[2]. Abbasi, Ahmed, Suprateek Sarker, "Big data research in information systems: Towards an inclusive research agenda." Journal of the Association for Information Systems 17.2 (2016).

[3]. Fan, Jianqing, Fang Han "Challenges of big data analysis." National science review 1.2 (2014): 293-314.

[4]. Zerillo, Anthony J., and Gregory K. Embry. "Deluxe emergency notification." U.S. Patent No. 8,976,938. 10 Mar. 2015.

[5]. Chen, Min, Shiwen Mao. "Big data: A survey." Mobile networks and applications 19.2 (2014): 171-209.

[6]. Bharathi, S. Vijayakumar, and Tanuja Mandal. "Prioritising & ranking critical factors for sustainabale cloud ERP adoption in SMEs." Internationale Journal of Automation and Logistics 1.3 (2015): 294-316.

[7]. Hazarn, Benjamine T., et al. "Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications." International Journal of Production Economics 154 (2014): 72-80.

[8]. Dubey, Rameshwar, and Angappa Gunasekarra. "Education and training for successful career in Big Data and Business Analytics." Industrial and Commercial Training 47.4 (2015): 174-181.

[9]. Fox, Stephen. "Getting real about Big Data: applying critical realism to analyse Big Data hype." International Journal of Managing Projects in Business 6.4 (2013): 739-760.

[10]. Jeble, Shirish. "Role of big data and predictive analytics." Internationale Journal of Automation and Logistics 2.4 (2016): 307-331.

[11]. Abbot, Deen. Applied predictive analytics: Principles and techniques for the professional data analyst. John Wiley & Sons, 2014.

[12]. Luis, Zhenjhua, et al. "Renewable and cooling aware workload management for sustainable data centers." ACM SIGMETRICS Performance Evaluation Review. Vol. 40. No. 1. ACM, 2012.
[13]. Jeble, Shirhis, Sneha Kumar, and Yogesh. "Role of big data and predictive analytics." International Journal of Automation and Logistics 2.4 (2016): 307-331.
[14]. Rai, Siddharth Shankar, Vishal Sharma, and Kunal Ganguly. "Logistics complexity in Indian garment supply chain." International Journal of Automation and Logistics 1.4 (2015): 419-430.
[15]. Stephens, Zachary D., et al. "Big data: astronomical or genomical?." PLoS biology 13.7 (2015): e1002195.
[16]. Kanamoria, Yuko, and Yuzuru Matsuokab. "Development of a Model for Estimation of Household Consumption and Environmental Load Generation." Refereed Sessions I-II Monday 10 March (2008): 347.
[17]. Marshale, Antony, Stefane Meck, and Rebeccaa Shockley. "How leading organizations use big data and analytics to innovate." Strategy & Leadership 43.5 (2015): 32-39.
[18]. Vilajosana, Ignasi. "Bootstraping smart cities through a selfsustainable model based on big data flows." IEEE Communications magazine 51.6 (2013): 128-134.
[19]. Ananth V, Sridar Seshadrii, and Roy Vaser. Toyota supply chain management: A strategic approach to Toyota's renowned system. Vol. 240. New York, NY: McGraw-Hill, 2009.
[20]. Tan, Kim Hua, et al. "Harvesting data to enhance suply chain innovation capabillities: An analytic infrastructure based on deduction graph." International Journal of Production Economics 165 (2015): 223- 233.
[21]. S Purri, T Choudhury et. al. ,Specialization of IoT applications in health care industries,ICBDAC, 2017
[22]. Tanupriya Choudhury et. al. Privacy and Security of Cloud-Based Internet of Things (IoT), CINE,2017