Journal of Advances in Science and Technology
Vol. 19, Issue No. 1, March-2022, ISSN 2230-9659

Big Data Analytics

Prasann Bhatlawande[1]* Onkar Waghmare[2] Harshad Gaikwad[3] Chirag Mulani[4] Punam Kapane[5] Mr. Kase S. S.[6]

[1,2,3,4,5] Student, Department of Computer Engineering, Sahakar Maharshi Shankarrao Mohite Patil Institute of Technology and Research, Akluj, Solapur, Maharashtra, India

[6] Lecturer, Department of Computer Engineering, Sahakar Maharshi Shankarrao Mohite Patil Institute of Technology and Research, Akluj, Solapur, Maharashtra, India

Abstract – Huge volumes of data have become available to decision makers in the digital age. Big data refers to datasets that are not only large, but also diverse and rapidly changing, making traditional tools and procedures ineffective. Due to the increasing rise of such data, solutions to handle and extract value and knowledge from these datasets must be investigated and given. Furthermore, decision-makers must be able to extract useful information from such a diverse and frequently changing set of data, which includes anything from everyday transactions to customer interactions and social media data. Big data analytics, which is the application of advanced analytics techniques to large amounts of data, can give such value. The purpose of this paper is to investigate as the opportunities provided by the application of big data analytics in various decision domains.

Keywords – Big Data, Data Mining, Analytics, Decision Making.

----------X----------

INTRODUCTION

Imagine a world where every detail about a person or organisation, every transaction completed, and every aspect that can be documented is lost immediately after use. Organizations would thus lose the ability to extract valuable information and expertise, undertake extensive analyses, as well as give new opportunities and advantages. Customer names and addresses, products available, purchases made, staff hired, and so on have all become critical for day-to-day operations. Data is the building block upon which any company thrives. Consider the breadth and volume of data and information available today as a result of technological advancements and the internet. As storage capacity and data collection methods have evolved, massive amounts of data have been readily available. New data is generated every second, and it must be kept and processed in order to extract value. Furthermore, because data storage has become less expensive, businesses must extract as much value as possible from massive amounts of data. Because of the volume, variety, and rapid change of such data, new types of big data analytics, as well as new storage and analysis methods, are required. Such massive amounts of large data require thorough analysis and extraction of relevant information.

LITERATURE REVIEW

Big data is a word for big data sets having large, more diversified and complicated structure with the difficulties of storing, analysing and visualising for further processes or outcomes. Big data analytics is the study of large volumes of data in order to uncover hidden patterns and hidden relationships. These helpful data for businesses or organisations can help them obtain a competitive advantage by providing them with richer and deeper insights. As a result, big data implementations must be thoroughly examined and carried out. This paper provides an overview of big data, including its content, scope, samples, methods, benefits, and challenges, as well as privacy concerns.

1. Big data:

One of the most important outcomes of the digital age is the accumulation of large amounts of raw data. Managing such important cash in various shapes and sizes based on The attention of the management is required by the organisation. Big data has the potential to impact every facet of life, from social to educational, and everything in between. As the amount of data increases especially in technology-based firms, the subject of managing raw data becomes considerably more crucial. When confronted with raw data characteristics such as variety, velocity, and volume,

specialised methods are required to overcome the complexity and hidden body of the data.

2. Deep learning:

Deep learning emerged as a subsection of machine learning as a result of factors such as increased chip processing, which results in the creation of massive volumes of data, lower computer hardware prices, and significant advancements in machine learning algorithms. Deep learning algorithms are divided into four types.

3. Machine learning:

Machine learning is defined as data-driven predictive algorithms that are followed by a learning algorithm in an unstructured programme. ML is divided into three categories: supervised, unsupervised, and reinforcement learning [47], and it takes place throughout the "data preprocessing," "learning," and "evaluation" phases. Preprocessing is the process of changing raw data into a format that can be used in the learning phase, and it includes steps like as cleaning the data, extracting, manipulating, and merging it. WORKING: In the evaluation phase, a data set will be chosen, and performance will be evaluated, statistical tests will be run, and errors or deviations will be estimated.

A data analyst cleans and analyses data sets to assist in issue solving. A data analyst collects, cleans, and evaluates data sets in order to a solve a problem. They can work in a variety of fields, such as business, finance, criminal justice, science, medical, and government, among others.

PROCESS:

1. Data Requirement Gathering

2. Data Collection

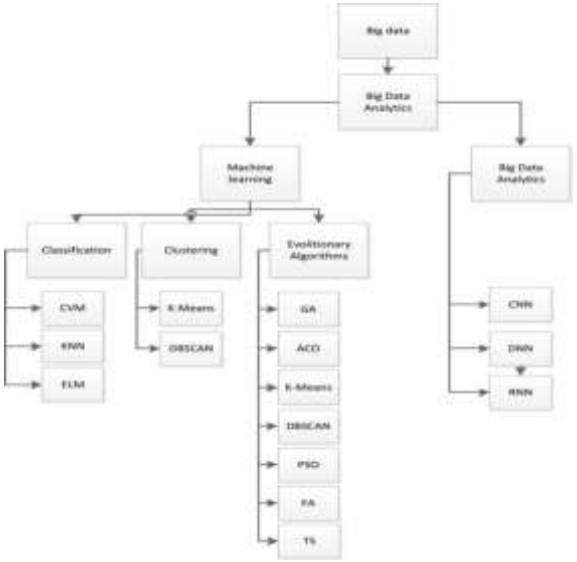

Fig. Big Data Analytics

3. Data Cleaning: It's time to clean up the data you've collected because not all of it will be valuable. White spaces, duplicate data, and simple errors are all removed during this step. Before sending the data to be analysed, it must first be cleaned.

4. Data visualization: "Graphically portray your knowledge in a way that people can read and understand it," is a sophisticated way of stating. Charts, graphs, maps, bullet points, and a variety of other ways can be used. By allowing you to compare datasets and observe relationships, visualisation assists you in gaining valuable insights

5. Data Interpretation

Fig 2: data analysis flow.

CONCLUSION

To begin, restate the study's overall goal. Then explain how the key finding relates to the study's ultimate goal. Then, from the results section, summaries any other noteworthy discoveries. Explain how the statistical results connect to the study's goal.

REFERENCES

[1] Chengang Zhu, Guang Cheng (senior member, IEEE) and Kun Wang (senior member) IEEE, “Big data analytics for program popularity prediction in broadcast TV industries”, IEEE Access, October’2017

[2] Han Hu, Yonggang Wen (Senior member, IEEE), Tat- Seng Chua and Xuelong Li (Fellow, IEEE), “Toward scalable systems for Big Data Analytics: A technology tutorial”, IEEE Access, June’2014/

[3] Kubick, W.R. (2012). Big Data, Information and Meaning. In: Clinical Trial Insights, pp. 26–28.

[4] Russom, P. (2011). Big Data Analytics. In: TDWI Best Practices Report, pp. 1–40.

[5] Asur, S., Huberman, B. A. (2010). Predicting the Future with Social Media. In: ACM International Conference on we Intelligence and Intelligent Agent Technology, vol. 1, pp. 492–499.

[6] Economist Intelligence Unit: The Deciding Factor: Big Data & Decision Making. In: Capgemini Reports, pp. 1–24 (2012).

Corresponding Author

Prasann Bhatlawande*

Student, Department of Computer Engineering, Sahakar Maharshi Shankarrao Mohite Patil Institute of Technology and Research, Akluj, Solapur, Maharashtra, India