ACADEMICIA
An International Multidisciplinary Research Journal
(Double Blind Refereed & Peer Reviewed Journal)

AN OVERVIEW OF BIG DATA IN EDUCATION

Dr. Ajay Rana*; Dr Tarun Kr. Sharma**

*Shobhit Institute of Engineering and Technology,
(Deemed to be University), Meerut, INDIA
Email id: ajay.rana@shobhituniversity.ac.in,

**School of Computer Science and Engineering,
Faculty of Engineering and Technology,
Shobhit Institute of Engineering and Technology,
(Deemed to be University), Meerut, INDIA
Email id: tarun.sharma@shobhituniversity.ac.in

ABSTRACT

The trendiest term in the twenty-first century is big data, which is collected by flooding data from billions of devices. Big data is data that is produced in large quantities, at a fast pace, and in a variety of formats that are complicated (veracity), necessitating analytical processes and active data management in order to extract useful insights. Similarly, big data is seen as a game-changer capable of altering the way companies function in various organizations in order to gain a lasting competitive edge. This article provides an overview of big data technologies in higher education and explores how big data analytics may provide value to Higher Education Institutions and help them maintain a competitive edge. Given the rising importance of educational institutions as a source of income that boosts the economy, the education industry is working in an increasingly complicated and competitive environment.

KEYWORDS: *Big Data, Higher Education, Innovation, Learning, Teaching.*

1. INTRODUCTION

We are living in the information era, and big data has become an integral component of most of our daily lives. As a result of the everyday explosion of data, new integrated innovative technology demands have reached an all-time high. Many organizations, including science, healthcare, engineering, business, and eventually society at large, are attempting to change their organizations using big data in order to gain a competitive edge over their counterparts in the

same setting. The term "big data" isn't defined in any manner (an efficient way of applying big data). As a result, big data may be described as a collection of data that is so complicated and vast that processing and analyzing it using on-hand database management tools is difficult. The primary aim of big data is to manage, process, and analyze its properties. Volume, Velocity, Variety, Veracity, and Value (i.e., the unknown insights of data) are the 5 Vs of data-related aspects. As a consequence, by evaluating performance and establishing competitive advantages, big data capabilities may provide long-term value.

The idea of using big data analytics in higher education is to get a more evidence-based competitive edge in terms of new teaching-learning approaches. Despite the fact that big data analytics is still a relatively new use in higher education, the technology has the potential to improve and contribute to institutions. Education institutions may use big data analytics to address some of their problems, as they face growing pressure from both national and international economic, global competitive advantage, and social and political change. Universities contain a lot of information. Although these databases are plentiful and increasing, they are often underused. As a result, the higher education industry has access to a significant quantity of data that may be utilized to make effective decisions for better educational quality. Higher and professional education is a domain that must be evaluated and transformed on a regular basis to keep up with the fast pace of changing trends in various market sectors, which in turn creates a variety of workforce needs. Technology has had a significant impact on the way education is delivered. Mobile devices and apparatuses, teleconference and remote access systems, educational platforms and services, and other technologies that students, teachers, academic faculty, evaluation specialists, researchers, and decision-makers in education interact with and use in an effort to impact and improve teaching and learning but also to realistically impact and improve teaching and learning.

The connection with these technologies produces a significant quantity of data, which may vary from a single access log file to institutional activities. The educational systems are still not fully prepared to deal with and exploit them for the purposes of continuous quality improvement. Health professions education, also known as health education, is one of the most common settings in which these technologies are used, resulting in a wide range of educational data. Furthermore, health education must always reflect the increasing corpus of medical knowledge and evidence in order to effectively integrate it in education and educate future health professionals to address the problems that healthcare systems will face in the future. The need to govern these challenges in health education is now more important than ever, and as a result, different approaches such as big data and analytics that could be useful in investigating and exploiting educational data have received attention.

1.1 Big Data Analytics in Higher Education Institutions:

As a result of the everyday explosion of data, new integrated innovative technology demands have reached an all-time high. Big data has developed as a result of this. Many studies have been conducted on the application of big data in a variety of areas, including data availability, cost, applicability, importance, and security. Despite this, just a few articles discuss the integrated use of Big Data in higher education. Higher education institutions have access to a massive quantity of data from a variety of sources, all of which are regulated by various procedures. As a result, big data has the ability to harness institutional data, which may help the education sector's future,

especially in decision-making. Big data analytics may be evaluated as a complete technological IT source in educational institutions by evaluating a set of processes that will contribute to long-term competitive advantage. The analytics of big data architectural component may be utilized to convert data from many sources and formats into useful insights in higher education using big data analytics technologies.

Big data component architecture is made up of five levels. Data, data analytics, data aggregation, information exploration, and data governance are examples of these levels. The data layer covers the data sources required to offer insights that will enable repeated operational procedures. As a result, the data layer can assist any difficulties that institutions have in making better decisions. analytics: The analytics layer assists with the analysis and visualization of data in HEIs for the purpose of improving learning and the environment in which it occurs data aggregation: The data aggregation layer assists with the analysis and visualization of data in HEIs for the purpose of improving learning and the environment in which it occurs. The processes of data management arising from different sources of data, such as data from communication channels, sizes, and formats, are referred to as the data aggregation layer (unstructured and structured). As a result, the data aggregation layer attempts to analyze data from a wide range of sources.

However, one of the major challenges in implementing Big data technology is the variability of incoming data; the information exploration layer produces outputs such as real-time or near real-time data by examining reports and visualizing them to derive meaningful insights; the data governance layer entails understanding data management, its lifecycle, and security; and the data governance layer entails understanding data management, its lifecycle, and security. The governance, standards, rules, tools, and procedures for data management are considered dominant data management. The process of managing data throughout its lifetime is known as data life-cycle management. Big data analytics offers innovativeness for data monitoring, auditing, and protection. Data security and privacy management is a concept in which big data analytics provides innovativeness for data monitoring, auditing, and protection.

As a result, each of the big data architectural components may be predicted to have the capacity of big data application inside educational institutions. As a result, the study aims to utilize the UTAUT model to grasp the anticipated capabilities of big data analytics to generate prolonged competitive advantage and business value, based on the complexity of big data analytics and sustained competitive advantage as stated by and. For example, big data may help maximize student learning success when some students are undecided about which main subject to take, leaving teachers wondering how to tailor learning pathways so that no student falls behind. Higher education is one of the areas where volume, diversity, and velocity coexist in data. In the higher educational environment, large quantities of educational data are collected and produced on a regular basis from many sources and in various forms. Educational data ranges from those generated by students' use of and interaction with learning management systems (LMSs) and platforms, to learning activities and course information, which includes learning objectives, syllabuses, learning material and activities, examination results, and course evaluation, as well as other types of data related to administrative, educational, and quality assurance.

The limited exploitation of large educational data, as well as the quantity and nature of these data in higher education, necessitates the use of specific methods to uncover new useful information

that is presently buried inside data. Such methods may be developed and modified from other areas that are characterized by big data, and they can be effectively used to large educational data. These methods may be utilized to generate insights "about student performance and learning approaches" and illustrate areas within large educational data that can be favorably affected, such as students' real performance according to taught curriculum. Big data and analytics, when used together, have recently showed potential in encouraging various activities in higher education. These actions pertain to "administrative decision-making and organizational resource allocation," "early identification of students at risk of failing," "development of effective instructional techniques," and "transforming the traditional view of the curriculum to reconsider it as a network of relations and connections between the various entities of data gathered and regularly produced from LMSs." The curriculum and its contents, as a significant component of large educational data, are one of the recognized areas in which big data and Analytics are properly relevant for inquiry and development in higher education.

1.2 Big Data Analysis Innovative Teaching-Learning:

Higher education institutions are now in the process of rethinking how they achieve their purpose in order to maintain a competitive edge. However, political and economic pressures have increased scrutiny of the quality of higher education institutions, particularly in terms of labor costs such as government funding and policy changes, access and continued growth of students, leadership changes, economic sustainable development, and updates of institutional policy are mostly the challenges the high education institutions face. Educational data analytics and mining may make previously ignored, invisible, and therefore inaccessible data apparent. The higher education industry, as may be observed, is rapidly expanding. The projection depicts how sustainability patterns are anticipated to change depending on different nation perspectives.

Most CEOs now identify data as a new digital innovation that is regarded as a need in the higher education system owing to the development and evolution of sustainable competitive advantage. The degree to which an organization integrates a concept while also producing a strategy that no existing or prospective competitors follow is referred to as competitive advantage[1]. Similarly, a company is considered to have a competitive advantage if no present rivals are adopting an innovation strategy at the same time and it has not yet been able to duplicate the advantages of the approach. As a result, when a company outperforms its rivals, it has a competitive advantage. Despite the fact that big data is still relatively new and little understood, it has the potential to provide organizations with much more development potential than conventional technology. Big data analytics integrative functions may utilize various statistical analyses and machine language to identify dangers, problems, and opportunities in the system, which can offer economic value for higher education. As a result, education may become more dynamic and less expensive, learning methods can be enhanced, and operations can be provided [2]. As a result, educational achievement may be improved. The compatibility of big data technology with sustainability, especially in the phase of multidimensional environment, intelligent distribution of sources, integrated practices, and expertise priorities, may be disputed.

1.3 Advantages Of Big Data In Education :

- *It assists you in finding solutions to difficult problems*:The greatest approach to brainstorm answers to the difficult problems confronting the education sector is to evaluate your current

data. The more you understand about your past, the more you can learn from it.If you work in higher education, for example, you may see a drop in enrolment. You can use big data to get the context cues you need to figure out where, when, and how your enrollments are changing[3].

- *It is easily accessible*: It's inconvenient and time-consuming to search through a wall-to-wall arrangement of file cabinets. It's much simpler to locate what you're searching for since big data depends on a technical infrastructure to collect, store, and manage information. Institutional silos, in addition to infrastructure, may make it difficult to exchange information. Leadership may have access to data that instructors do not, which may create obstacles to learning and development. With the proper tools and data analytics, you can create a more collaborative workplace. Because all of the data is in one place, all you need is internet connection to locate what you're looking for. You don't even need to install a program or plug-in since many software products are accessible via browsers like Google Chrome and Safari[4].

- *It may help you save money*: In higher education, proper resource allocation is critical, and your data is the key to efficiency. For one thing, your data may provide information about various class sections' enrollment figures. If just two of the five sections of ENG 102 are full, the remaining three may be merged to save resources such as classroom space, instructor time, and energy usage[5].From an infrastructure standpoint, cloud-based solutions have the ability to reduce data storage costs while also relieving strain on your IT staff. Manual sorting and transcribing of data has historically been needed, which is time intensive and may take weeks, if not months. It may also take a long time to generate customized reports on a regular basis. Your workers' time might be better spent on more productive activities[6]. An analytics software will automate most of this time-consuming labor, and the convenience of digital data makes data access fast and simple, saving you money in the long term. Efforts to attract students for higher education are another cost-cutting advantage. Your data may reveal you which potential students are more likely to succeed at your institution, as well as who are more likely to drop out or fail, by looking at previous school performance. This may assist you in developing more effective acceptance procedures that optimize the return on investment for each student[7].

- *It is Fast*: This was briefly stated under the cost-cutting advantage, but it needs repeating. You'll save a lot of time sifting through data to locate one specific report or piece of information on a certain kid now that all of your school's information is in one place. Big data is also accessible in real time, allowing you to make faster choices than ever before. This is especially useful during enrollment times, when instructors are keeping track of statistics in order to plan for the next semester. Year-Over-Year Enrollment reports may be automated to assist you see how enrollments are doing in comparison to the same time last year[6]. Then you may make choices on the spot to enhance or maintain the consistency of your enrollments.

- *It assists you in adapting*: You may create new courses, instructional techniques, and other ways to offer students what they need and desire by detecting patterns. Take, for example, community colleges. Because these institutions are mostly inhabited by adults who are juggling job, families, and education, a flexible timetable is one of their top priorities when

choosing a school. Many of these adult students now find that online courses are even more convenient than late-night or weekend sessions, thanks to the growth of online learning. Big data will reveal the exact statistics behind this since online learning is better suited for some kinds of courses such as those that do not need a lab or hands-on learning. Perhaps on-campus biology class enrollments have remained stable while on-campus English course enrollments have decreased[8]. You may modify your course offers appropriately to guarantee that your students always have the greatest choices.

1.4 Analysis Of Data Mining In Education:

Data mining methods are gaining traction in the education industry, and the results of these approaches may offer essential decision-making assistance. making. Data mining in education is referred to as as an example, educational data mining (EDM). EDM is a new genre. Data mining is a field that focuses on using data mining tools and techniques to solve problems. Methods for analyzing data in the field of education This section contains a list of for EDM, a literature review or survey articles are required. It emphasizes their most important achievements A review of current literature contains a review or survey article. This is a review ofshows twenty years of e-learning data mining research[9]. From a pedagogical standpoint, surroundings are important. The creators discovered and categorized problems that need to be researched in order to be solved Performances by student learners.

Another review of the literature focuses on EDM and learning analytics in higher education and was released in 2019. This study's findings are based on a review of the literature. four major areas were covered: Computer-assisted instruction CSLA) and the use of DM methods to obtain information that may be used depending on student engagement in the LMS (2) Predictive analytics with the help of computers (CSPA), as well as the usage of EDM and LA to predict student performance[10]. Assessment-based performance and retention in courses. In a learning activity, engagement and subject knowledge are important; (3) The use of computer-assisted behavioral analytics (CSBA) to detect student behavioral trends and use DM methods When it comes to engaging in online learning activities, I have a few preferences. and (4) visualization analytics with the aid of a computer (CSVA) as well as a mix of data visualization methods Data mining and knowledge representation have progressed as a result of these advancements. to provide a visual representation of student behavior in relation to the educational activity

2. DISCUSSION

The application of Big Data and Open Data in Education will be examined in this research. Also, how much data can be utilized and extracted to make something valuable, assisting the business in increasing revenues. As a result, defining Big Data and Open Data clarifies how these two technologies are categorized.Following that, it is critical to discuss the objectives and purposes of Big Data in education, as well as the value and impacts of Big Data in education. We investigate how Big Data's value potential has changed in recent years and what the future holds. Finally, we will examine the educational advantages of Big Data and Open Data, as well as a short explanation of how these technologies may contribute to a world-class educational system. In the field of education, assisting instructors and students in making more focused decisions. Big Data has the potential to significantly enhance education. May afford to create a contemporary, dynamic educational system from which every single student can benefit to the fullest extent possible. Furthermore, instructors now have important tools that they did not have

previously, allowing them to make more precise choices and pick from a wide range of innovative learning techniques. As a result, Big Data is actively engaged in changing the way sectors, including education, operate. Traditional problems will no longer exist in the new age of data, but excellent techniques will be maintained. New learning methods will be added to the school system, making it more efficient and focused. However, the journey into this new age has only just begun, and there are many challenges ahead, including a scarcity of skilled people in the fields of Big Data and Data analytics. Furthermore, instructors and academics must be trained and engaged in these new technologies, and students must embrace and utilize them.

3. CONCLUSION

The explosion of data has given rise to big data, which is used to analyze large amounts of data from a number of sources. Its capabilities may be very beneficial to higher education institutions. Higher education institutions must investigate the long-term implications of using big data analytics, which may help institutions expand their teaching-learning focus. Executives could suggest that administrators and educators assist the system and gain value from practical applications, such as creating a culture of data use for educational decision making, being keen with data users by asking critical questions about market deals and suggesting the most beneficial uses and features, involving IT departments in data collection and application planning, and starting with the most basic uses and features. As a consequence, institutions may save money by integrating educational abilities and making better decisions. This framework aims to improve the development of innovative teaching-learning orientations for educational institution performance as an insightful contribution to curriculum design for instructors and learners in education in particular, as well as the education sector's overall competitive advantage.

REFERENCES

1. F. Kalota, "Applications of Big Data in Education," *Int. J. Soc. Behav. Educ. Econ. Bus. Ind. Eng.*, 2015.

2. Y. Wang, "Big Opportunities and Big Concerns of Big Data in Education," *TechTrends*, 2016, doi: 10.1007/s11528-016-0072-1.

3. A. S. Drigas and P. Leliopoulos, "The Use of Big Data in Education," *Int. J. Comput. Sci. Issues*, 2014.

4. P. Schwerdtle and J. Bonnamy, "'Big Data in Nurse Education,'" *Nurse Education Today*. 2017, doi: 10.1016/j.nedt.2016.08.003.

5. C. F. Lynch, "Who prophets from big data in education? New insights and new challenges," *Theory Res. Educ.*, 2017, doi: 10.1177/1477878517738448.

6. A. Klašnja-Milićević, M. Ivanović, and Z. Budimac, "Data science in education: Big data and learning analytics," *Computer Applications in Engineering Education*. 2017, doi: 10.1002/cae.21844.

7. K. Sin and L. Muthu, ""APPLICATION OF BIG DATA IN EDUCATION DATA MINING AND LEARNING ANALYTICS – A LITERATURE REVIEW "," *ICTACT J. Soft Comput.*, 2015, doi: 10.21917/ijsc.2015.0145.

8. E. Zeide, "The Structural Consequences of Big Data-Driven Education," *Big Data*, 2017, doi: 10.1089/big.2016.0061.

9. J. A. Johnson, "The Ethics of Big Data in Higher Education," *Int. Rev. Inf. Ethics*, 2014, doi: 10.29173/irie365.

10. F. Kalota, "Applications of big data in industries," *Int. J. Soc. Behav. Eudcational, Econ. Bus. Ind. Eng.*, 2015.