**Explore – Journal of Research**
**Peer Reviewed Journal**
**ISSN 2278–0297 (Print)**
**ISSN 2278–6414 (Online)**
**Vol. XIV No. 2, 2022**

https://patnawomenscollege.in/explore-journal-of-research/

# Big Data Analytics Approach in Education

• **Satya** • **Prasanna Jha**

Received : April 2022
Accepted : May 2022
**Corresponding Author : Satya**

**Abstract:** This chapter introduces our readers to big educational data and how that data may be analyzed to give insights into many stakeholders and empower data-driven decisions regarding educational quality improvement. We present different techniques and popular applied scientific methods for data analysis and manipulation, such as analytics and other analytical approaches such as learning, academic, and visual analytics for the research and exploitation of big educational data and examples of how these techniques and methods could be used. The notion of educational quality enhancement is offered in connection to two factors: (a) to improve science and its influence on various processes in education such as learning, education, and academic processes, and (b) as a result, we introduce our readers to the idea of extensive educational data and how such data may be analyzed to offer insights into the practical application and execution of the analytical principles provided. The setting of health professions education is utilized to illustrate the various principles.

***Keywords:*** *Big data, educational data, data-driven decisions, data analytics, learning analytics*

**Satya**
MCA Student
Amity Institute of Information Technology
Amity University, Patna
Email-id: satyadubey6869@gmail.com

**Prasanna Jha**
Asst. Professor
Amity Institute of Information Technology
Amity University, Patna
Email-id: pkumar3@ptn.amity.edu

**Introduction:**

Higher and professional education is an area that is constantly being examined and altered to keep up with the quick speed of changing trends in many market sectors, which in turn produces a range of labour demands. Technology is a crucial aspect that has drastically impacted how education is delivered. Mobile devices and apparatuses, teleconference and remote access systems, educational platforms and services, and others are examples of technologies that students, teachers, academic faculty, evaluation specialists, researchers, and decision-makers engage with and utilize in education to affect and improve teaching and learning, but also to truly depict the employment of current technology in the learning stage create massive volumes of data, ranging from a single access log file to institutional-level activity. Still, educational systems are not entirely prepared to deal with and capitalize on them for the objectives of continual quality improvement. These technologies are most commonly employed in health professions or health education, generating a wide range of instructional data. Furthermore, to incorporate medical knowledge and evidence in education and train future health professionals to tackle the future problems of healthcare systems, health education is always in need of reflecting the rising corpus of medical knowledge and evidence—the necessity to govern these difficulties effective in studying and leveraging educational data.

**Big data and education:**

**Big data:** Today, "big data" is widely used to characterize and define large data volumes' recent rise and existence. It may be found in a wide range of industries. The public, commercial, and social sectors constantly receive and generate massive data from many sources and in various formats. In certain circumstances, data reach huge volumes, such as petabytes, beyond the hardware or human ability to warehouse, manipulate, and analyze it, hence referred to as big data. Nonetheless, this phrase refers to vast amounts of data, even if the scale might vary from sector to sector or, more particularly, across services within an industry [1]. Because of its vast scale, big data is referred to as such. Nonetheless, big data is defined by additional characteristics such as the disparate types and formats of data and the various sources from which they are collected, as well as the speed at which they are produced and, most importantly, the frequency with which they are processed, whether in real-time, frequently, or infrequently. All of these qualities are characterized as volume (size), diversity (sources, formats, and types), and velocity (speed and frequency), and they add complexity to the data, which is another worry. [2]. Data in a system or a specific domain is termed extensive data when the volume, diversity, and velocity are all high simultaneously, regardless of whether these three qualities are deemed "small" in another area. In this example, this is sufficient to test constraints in changing and analyzing data so that it may be utilized for various reasons. Data sizes can range from megabytes to petabytes depending on the domain. Thus, big data is context-specific and may relate to different volumes and data types depending on the domain. Still, the fundamental difficulty that all of these domains must face is making sense of the data by processing it at a high analytical level to allow a data-driven change in processes and procedures [3]. Big data and analytics have given value to data held in many settings, proving to be an exceptionally useful way of examining its potential influence in the industry in business intelligence and analytics [4]. Alternatively, educational data mining approaches and learning analytics are used in academia [5]. Given the lack of research on the use of big data and analytics in the context of health education, we will introduce the reader to the new subject of big educational data, which places big data in education and how educational data may be treated in many dimensions and from diverse viewpoints to bring to light insights for various stakeholders such as decision-makers, academic faculty, evaluation experts, researchers, and computer science students.

**Big educational data:** Higher education is one of the data domains where volume, diversity, and velocity coexist. Large volumes of educational data are acquired and created regularly from many sources and in various formats in the higher education ecosystem. Educational data range from those generated by students' use and interaction with learning management systems (LMSs) and platforms to learning activities and course information consisting of a curriculum such as learning objectives, syllabus, learning material and activities, examination results, and course evaluation, to data related to administrative, educational, and quality improvement processes and procedures. The restricted use of extensive educational data and the scale and variety of data in the context of higher education indicates the need for particular strategies that must be used to uncover new useful knowledge that is currently concealed inside data. Such a process may be derived and modified from other extensive data areas and effectively used to massive educational data. These strategies might generate insights "about student performance and learning approaches" and demonstrate areas within huge educational data that can be favorably influenced, such as students' actual performance according to the taught curriculum. Big data and analytics have recently shown potential in supporting various activities in higher education. These efforts include "administration decision-making and organizational resource allocation," the identification of pupils at risk of failure, the creation of effective teaching strategies, and the transformation of the traditional vision of the curriculum must rethink it as a network of relationships and linkages between the many entities of data acquired and created daily from LMSs, social networks, learning activities, and the curriculum More precisely, as a critical component of big educational data, one of the highlighted areas in which big data and analytics are suitably utilized for inquiry and development in higher education is the curriculum and its contents.

**Analytical approaches**

Analytics activities require several components to be effective. These components are the data (kind and source) and the context of interest. If these analytics components are in place, we can develop various analytics models that can live and evolve into an analytics engine capable of harnessing extensive data.

**Learning analytics**

LA is described as "the collecting, analysis, and reporting of data about learners and their settings to understand and optimize learning and the environments in which it occurs" and impacts actions and processes. We can find behavioral commonalities (e.g., user happiness) or aberrant patterns using LA (e.g., cheating). It can link past and future operations by feeding data about past occurrences into a LA engine and analyzing it to predict likely future results. Therefore, it may synthesize large amounts of educational data and provide a set of predictions to offer several decision possibilities, showing the ramifications of each decision option each time.

Teachers typically rely on their "gut instinct" based on their experience. To interpret students' behavior and predict if a student would drop out of a course or leave their studies. This can be demonstrated to be true or incorrect, but without proof, conclusions based only on experience have low confidence. An example displays the LA's ability to utilize evidence to strengthen this sort of conclusion. Data mining techniques were used in massive educational data as part of an analytics engine to recognize students who perform at high, moderate, and poor levels and inform them with various sorts of feedback. As a result, students in danger were discovered early on, while the school still had time to respond and take preventive measures.

**Academic analytics**

Academic analytics is defined as "the convergence of technology, information, management culture, and the use of information to govern the academic business." It focuses on reporting, modeling, analysis, and decision support for university and campus services. Admission, advising, funding, academic counseling, enrolling, and administration are examples of this type of service. The following is an example of a practical application of academic analytics, in which librarians used analytics on library usage data as part of the extensive educational data ecosystem to predict students' grades, demonstrating the value that data produced and processed in the library can provide the hosting institution.

**Visual analytics:**

In recent years, methods and strategies for manipulating complex data in various areas have been created. As an evolution of the sciences of information visualization and scientific visualization, visual analytics (VA) is the science of analytical reasoning aided by interactive graphical interfaces. VA integrates a variety of methodologies, including data visualization, data processing, and the power of human visual perception.

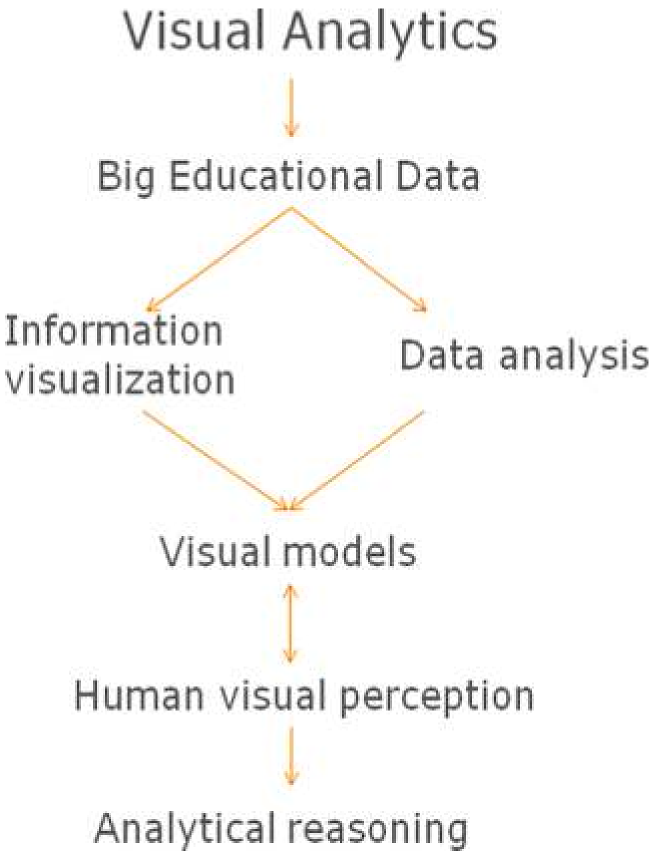


**Fig. 1 Visual Analytics**

Big educational data are modelled using information visualization and data analysis methodologies and portrayed in visual interfaces with which human visual perception interacts to influence the analytical reasoning process.

It can aid in manipulating and using big data by providing a comprehensive perspective of the data while disclosing complicated underlying information to the greatest extent feasible to affect analytical reasoning and decision-making positively. A survey of the literature yielded the following factors capable of supporting analytical thinking and decision making via VA and the interplay between human visual perception and visual interfaces:

Increased cognitive resources (V1)

Reduced requirement for information search (V2)

Pattern recognition is being improved (V3)

More accessible awareness of connection inference (V4)

Increased capacity to study and change data (V5)

**Quality improvement (QI):**

**Quality enhancement as a result of improvement science in education:** Quality improvement is described as "everyone's united and unrelenting efforts to effect improvements that will result in improved results, better system performance, and better professional growth." This concept encompasses all areas of health care that are integrally linked to change attempts. Improvement science instruments are all of the many parts and components required to actualize the efforts necessary for quality improvement to be a successful process. Improvement science has been used in numerous fields, including vehicle manufacture and health care, as an alternate method of putting new information into effect. Even inside education, projects anchored on improvement science began to show results. The comprehensive perspective of the studied setting is a distinguishing feature of improvement science. The crucial first step is to define the context (e.g., the organization, the actors and stakeholders, the routines and workflow) and treat it as a system; a thorough understanding of how tiny changes in one system instance might influence other parts of the system is critical.

Traditionally, improvement science was founded on the "plan-do-study-act" cycle, which attempted to answer fundamental issues:

What do we hope to achieve with the intended change?

What improvements can we make to make things better?

How will we know if a change is also a good thing?

Today, the application of analytics in massive educational data may be a "game-changer," playing an unquestionably necessary part in coordinating the components of improvement science actions to make improvements that effectively lead to improved outcomesin the enhancement of educational quality The method below uses massive educational data and integrates the essential components, as well as for analytics, within the framework of education to achieve the desired change that results in improvement effectively.

Each of the five elements is motivated by a different body of knowledge and has its traits and situations.

**Quality improvement of the learning process:** LA refers to operations at the micro and nano levels, such as teaching and learning activities in a course. Teachers, course designers, and study and programme directors are examples of people who undertake these operations. The scenario below exhibits the practical application of LA in a course's quality improvement circle.

Teachers might utilize curriculum mapping tools to pinpoint specific gaps during the course planning process. As a result, they can identify which learning objectives are not adequately covered by teaching or learning activities. They require suggestions for new, more appropriate, and motivating instructional activities to incorporate into their routine. They may use the provided Analytics tools to analyze the class further and forecast its needs, such as student demographics, performance, various learning methodologies, technologies utilized, and group dynamics. Multiple algorithms and prediction models handle this sort of data that can build the class's properties. For the next round, visualization tools may be utilized to provide different recommendations for constructing acceptable activities for this specific class and explain the impacts of each decision. The course director has control over the actions and may monitor the pupils' progress during the course. They may zoom in and out of the entire class to a specific working group or individual student. They can also track the flow of the previous social networks. They can assess overall dedication and identify pupils who are in danger. They are on a widely used platform,department, or even data from relevant programmes at other institutions. The findings and experiences gained can be utilized to create a knowledge database for various educational initiatives. This can aid in the establishment of new policies throughout the company and is an essential component of quality development and academic research.

**Quality improvement of the educational process:** We demonstrated how VA might be used to help the analytical thinking and decision-making of stakeholders involved in educational process quality improvement. This is accomplished when visual and analytical components serve as complementary and supportive instruments of a synchronized engine. The analytics component used on massive educational data reduces its complexity while retaining essential information and critical qualities, which are included at the top level of the displayed graphics. The other component is visualization, which reveals paths and relationships by leveraging the human tendency to digest and quickly grasp visual information. These two elements cannot exist apart and are applied to data with incoherent structure.

**Conclusion:**

The purpose of this chapter was to expose the reader to the notion of extensive educational data and the many kinds of analytics as applied scientific domains, as well as to dig further into them, considering popular data manipulation techniques and how they might be translated inside the health education system and employed as ways to the exploitation of massive educational data produced by such methods. Apart from the techniques themselves, the benefits and potential to use them for quality improvement purposes in health education are provided and discussed in detail. As a result, analytics is critical in establishing a solid foundation for a meaningful VA outcome. Before visualizations, the data analysis helps shape the inchoate massive educational data that visuals are then accountable for representing. The effort required to apply each of the elements is a significant consideration. The work necessary for the aesthetic and analytical components is incomparable, and their responsibilities are entirely distinct. Analytics necessitates tremendous creation to shape the data in question and assemble all of the individual parts required to represent the data accurately. On the other hand, visuals are a needless effort because the network of connections and linkages have already been built. However, it takes knowledge to pick and gradually construct suitable images to emphasize broadly. Consider the vital information in the network provided by data analysis and add scientific value by going beyond simple statistical-based visualizations. Of course, human visual perception is indispensable in this chain of events for perceiving and interacting with graphical interfaces and doing high-level analysis. In summary, VA enables many stakeholders to readily understand the structure of the studied data, explain how each component coexists as part of a network, and justify its usage and value in the data. It also aids in a better understanding of stakeholders' particular roles in the educational process and the repercussions of supplying their pieces without knowing how they may be harmonized with other parts of the data. Stakeholders must also determine how to deal with inconsistencies and structural abnormalities discovered by gap analysis and the presence or absence of constructive alignment in the data.

***References :***


1. EMC Education Services' "Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing, and Presenting Data."
2. “Big Data: Does Size Matter?” by Timandra Harkness
3. “Statistics of Big Data for Dummies” by Alan Anderson and David Semmelroth
4. Michael Minelli's "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business."
5. “Big Data, Data Mining and Machine Learning(Wiley Big Data Series)” by Jared Dean
6. “Analytics: Data Science, Data Analysis and Predictive Analysis for Business” by Daniel Covington