Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2449

Survey on Big Data Analytics for Educational Sector: Abiding Challenges and Contributions

Umang Garg
Asst. Professor, Department of Comp. Sc. & Info. Tech., Graphic Era Hill University, Dehradun, Uttarakhand India 248002

Article Info
Page Number:1591 - 1599
Publication Issue:
Vol 70 No. 2 (2021)

Abstract

Due to technology advancements, the education sector has undergone major changes, particularly with the adoption of e-learning platforms and web services for student interactions. As a result, as many people log into these programmes, a large amount of data is being generated. Most academic institutions, however, have not yet completely tapped into this data, frequently merely handling the information connected to user queries while ignoring the crucial web behavioural inputs. This research attempts to provide a paradigm for using big data technologies to improve teaching and learning in educational institutions in order to address this issue.The document has three distinct goals. To begin with, list the advantages of big data technology in education, emphasising the potential gains for educational institutions. Second, to outline several big data applications in education and show how they might be applied to enhance educational procedures and results. The paper's final goal is to present specialised big data models that are designed for the educational system and provide a useful implementation framework.The research uses both survey and modelling methodologies to accomplish these goals. The survey's component uses review papers to describe the advantages and uses of big data in education while drawing conclusions from the body of existing research. The modelling approach, on the other hand, uses the Unified Modelling Language (UML) to offer a descriptive framework for comprehending the connections between little data and big data in the educational system. Big data covers the full learning process, including both students and teachers, while small data concentrates on conventional input about student performance alone.

Article History
Article Received*: 05 September 2021*
Revised*: 09 October 2021*
Accepted*: 22 November 2021*
Publication*: 26 December 2021*

Keywords: Big Data, Educational Institutes, LearningAnalytics, Teaching and Learning

I. Introduction

A considerable shift in teaching and learning approaches has resulted from the recent quick infusion of technological improvements in the educational sector. The use of big data analytics is one of the most promising breakthroughs in this field. Educational institutions now have access to an unprecedented amount of data produced by students, teachers, and the learning environment itself because to the growing adoption of e-learning platforms and web services. Nevertheless, despite the data's enormous potential, harnessing its value and using it to improve educational outcomes remain very difficult.The purpose of this survey is to examine the difficulties and benefits of using big data analytics in the educational sector. This study aims to shed light on the possible advantages of using big data analytics in education by assessing the current environment and comprehending the

challenges faced by educational institutions.When using big data analytics, educational institutions face a variety of difficulties. The first major obstacle is the enormous volume of data produced from numerous sources. Institutions struggle with data collecting, storage, and administration because they frequently lack the tools and knowledge necessary to handle such big datasets. Due to the sensitive information about students and teachers that is contained in educational data, privacy issues and data security also present significant difficulties. The protection of data privacy and adherence to legal obligations become top priorities.

Despite these difficulties, big data analytics has a lot to offer the field of education. Institutions can improve curriculum design, identify areas for development, and personalise learning experiences by analysing large-scale statistics to get deeper insights into student learning patterns. By anticipating student performance, predictive analytics can offer effective interventions at the right time. Big data analytics also enables educational institutions to optimise resource allocation, assess the efficacy of instructional approaches, and make data-driven decisions.With the goal of highlighting both the difficulties and the potential benefits of big data analytics in the educational sector, this study will look at previous studies, case studies, and best practises. Educational institutions may create strategies to deal with problems and use big data analytics to enhance teaching and learning results by understanding these constraints and opportunities.

Today, big data is required for both an increase in processing power and storage capacities. Additionally, the availability of data necessitates it. When using big data in the context of big data technology, there are a few considerations to keep in mind [1].

a) Additional concerns include data utilisation, data quality, context in which the data can be used, data dynamism, and data scalability or accessibility.
b) The final point relates to data modalities, which include structuring data, building networks, using multimedia, transmitting data signals, etc.
c) Data operators work with those in charge of data collecting, data preparation, data presentation and representation, data model creation, and data visualisation.

Big data's introduction has drastically changed the focus of analytics, moving it away from descriptive analytics and towards predictive and prescriptive analytics. Based on patterns in previous data, predictive analytics use cutting-edge approaches to produce insights and forecasts about upcoming events or outcomes. Prescriptive analytics, on the other hand, takes the outcomes of predictive analytics and goes a step further by offering suggestions or solutions to solve problems or improve decision-making based on those predictions. Organisations may now forecast future trends and make proactive, data-driven decisions thanks to the shift towards predictive and prescriptive analytics [2].

II. Related Work

There is a growing body of related work that explores the application of big data analytics in the educational sector, focusing on data analysis to drive insights and improvements. Several studies have highlighted the potential contributions and benefits of utilizing big data analytics in education. Here are some key areas of research and findings [3-8]:

Learning Management and Optimisation: Analysing data from learning management systems (LMS) can provide information on how well students are doing and interacting. In order to improve learning outcomes, discover effective teaching strategies, and improve course design, researchers have investigated the use of LMS data. Resource allocation, instructional methodologies, and assessment techniques can all benefit from data-driven approaches.

Making Educational Decisions and Policies: Big data analytics can assist in the decision-making and formulation of educational policies. Researchers have investigated the variables influencing student success, dropout rates, and achievement gaps by analysing extensive educational data. This data aids in the development of evidence-based policies, efficient resource allocation, and implementation of targeted initiatives by policymakers[9].

The massive-scale educational databases that store and manage massive volumes of educational data have been created as a result of the usage of big data and data analysis in the education sector. These databases act as central repositories for gathering, archiving, and analysing various kinds of information about students, teachers, curriculum, assessments, and other topics. In the context of big data and data analysis, the following are some crucial components of educational databases:

a. Data Collection: Educational databases gather data from multiple sources, such as student information systems, learning management systems, assessment platforms, attendance records, and demographic information. These databases consolidate data from various systems and provide a comprehensive view of the educational ecosystem[10].
b. Data Storage and Management: Educational databases utilize robust storage and management systems to handle the vast amount of data generated in the education sector. They employ scalable and secure infrastructure to ensure data integrity, accessibility, and privacy. Data is organized and categorized to facilitate efficient retrieval and analysis.
c. Data Integration: Educational databases integrate data from different sources to create a holistic view of the educational landscape[11]. By combining data from student records, assessments, attendance, and more, educational databases enable cross-referencing and correlation analysis to derive meaningful insights.
d. Data Analysis and Reporting: Educational databases support data analysis and reporting capabilities. Data analytics tools and techniques are applied to the database to extract valuable insights, identify patterns, trends, and correlations. These insights can be used to generate reports, visualizations, and dashboards that provide stakeholders with actionable information for decision-making.
e. Predictive Analytics: Educational databases facilitate the application of predictive analytics models. By analyzing historical data and patterns, predictive models can forecast student performance, identify at-risk students, and provide early intervention strategies. These predictive insights enable educational institutions to take proactive measures to support student success[12].
f. Data Privacy and Security: Educational databases prioritize data privacy and security. Stringent measures are implemented to safeguard sensitive student information and comply with data protection regulations. Access controls, encryption, and data anonymization techniques are employed to ensure data confidentiality.

g. Data-Driven Decision Making: Educational databases empower educational stakeholders, including administrators, teachers, and policymakers, to make informed decisions based on data insights. By analyzing data related to student performance, curriculum effectiveness, and teaching methodologies, data-driven decision making becomes an integral part of educational practices.
h. Educational databases powered by big data and data analysis offer a wealth of opportunities to enhance educational processes, personalize learning experiences, improve student outcomes, and inform policy-making. By effectively leveraging these databases, educational institutions can unlock the potential of data-driven insights to drive innovation and continuous improvement in the education sector.

III. Methodology

There are several different methods and approaches to applying big data analytics in the educational sector. Here are some commonly used methods:

1. Predictive Analytics: This [13] method involves utilizing historical data and statistical algorithms to make predictions about future educational outcomes. It helps institutions identify at-risk students, forecast student performance, and predict trends in educational processes. Predictive analytics can also be used for early intervention and personalized learning strategies.
2. Machine Learning: Machine learning algorithms enable the analysis of large datasets to identify patterns, correlations, and anomalies. By training models on historical data, machine learning can provide insights into student behavior, learning preferences, and instructional effectiveness. It can also be used for automated grading, adaptive learning systems, and personalized recommendations.
3. Natural Language Processing (NLP): NLP [14] techniques are used to analyze and understand human language data, including student essays, forum discussions, and feedback. It enables sentiment analysis, topic modeling, and text classification, facilitating automated assessment, feedback generation, and sentiment-based interventions.
4. Social Network Analysis (SNA): SNA focuses on analyzing the relationships and interactions among individuals within educational settings. By examining social network data, institutions can identify influential students, study collaboration patterns, and detect communities of practice. SNA can inform the design of collaborative learning environments and enhance social learning experiences [15].
5. Data Visualization: Data visualization techniques transform complex educational data into intuitive visual representations. This enables stakeholders to gain insights quickly and understand trends, patterns, and relationships in the data. Interactive dashboards, charts, and graphs facilitate data exploration and decision-making.
6. Text Mining: Text mining involves extracting valuable information and knowledge from unstructured text data, such as research articles, textbooks, and student notes. It can assist in curriculum development, identifying emerging topics, and generating personalized recommendations based on textual content analysis.
7. Learning Analytics: Learning analytics combines educational data and analytics techniques to understand and improve the learning process. It involves tracking student engagement, performance, and behavior to provide real-time feedback, personalized interventions, and

adaptive learning experiences. Learning analytics can optimize instructional design and inform pedagogical strategies.

These methods, among others, contribute to the wide range of approaches used in applying big data analytics in the educational sector. Institutions can adopt a combination of these methods based on their specific goals, available data, and technological capabilities to leverage the potential of big data for enhancing teaching and learning outcomes.

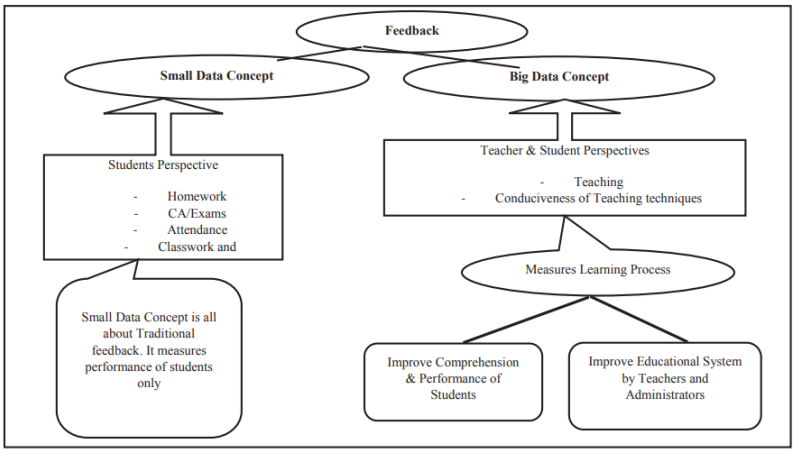

Figure 1: Education System Model presenting feedback systems for both the concepts of little data and big data

The idea of big data in education is shown in Figure 1, which focuses on both students and teachers through comprehending the learning process. By [16] assessing unique student requirements, selecting appropriate learning resources, and modifying teaching strategies and styles in accordance, this strategy personalises education. It includes learning analytics for forecasting student performance, stepping in when problems develop, and customising the learning experience to each student's skills. It also incorporates feedback mechanisms for continual improvement. The idea of "one size fits one" prevents personalised learning environments for each student, encourages personalised interventions for underachievers, and enables students to focus on their own areas of growth.

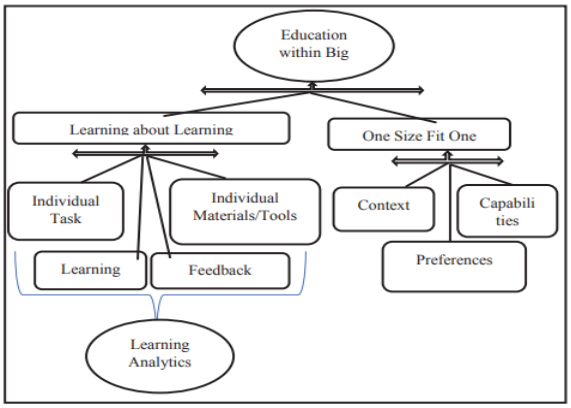

Figure 2: Education System Model in the Big Data Concept

Figure 2 illustrates the conventional educational idea outside of the big data domain, where students are treated equally and given the same learning resources and examinations. Individual variations in performance and comprehension levels are ignored by this "one size fits all" strategy. The requirements of students, especially those who perform averagely, are sometimes overlooked as the emphasis of this notion is primarily on the interests of teachers and the system.

IV. Discussion

This section describes the suggested structure that software developers might adopt and implement while creating systems related to the educational system.

Table I: Survey And Benefit Of Big Data To Education System

Sr.No.	Advantages	Finding
1	Instruction and teaching are improved	By enabling teachers to individually tailor their courses to each student through the use of data analytics, big data has the potential to improve students' performance and learning capacities. In order to assist students enhance their learning capacities and academic performance, teachers can modify their instructional approaches and offer individualised guidance when they have access to relevant data [17].
2	Match Job to Educator	The abundant availability of data allows industries and job seekers to match job requirements with desired skills and tools, enabling students to easily discover and apply for jobs that align with their capabilities and abilities[16].
3	Technology Transparent system	Big data technology makes it possible for students to participate in educational activities and programmes on an affordable budget, and it also gives educational administrators and policymakers the ability to manage educational budgets more effectively and transparently [18].
4	Sentiment Analysis	This relates to the viewpoints of those who have an interest in educational institutions, including students, teachers, etc. What they are saying, thinking, or feeling makes up this. For instance, what the students are saying about the management of the school in terms of creating and building the schools, what they are considering regarding the teaching materials or techniques employed by the teachers, or what are their emotional reactions to the learning process and processes[19]
5	Operational System Analytics	By analysing data from a variety of stakeholders, including potential and current students and teachers, embedding analytics in academic processes automates decision-making. This makes it possible to evaluate course applications, student performance and progress, and teacher effectiveness in-depth, providing data-driven insights and better academic results.[20]

Advantages using Big data in Education Sector

1. Personalized Learning: Learning experiences may be tailored to meet the needs of each individual student thanks to big data analytics. Educational institutions can offer personalised suggestions, adaptive learning paths, and customised interventions to improve

student engagement and achievement by analysing huge volumes of data on student performance, preferences, and learning styles.

2. Decision-Making Based on Data: Big data analytics equips administrators and educators to make decisions based on data. Institutions can find trends, patterns, and correlations that influence instruction, curriculum creation, and resource allocation by analysing a variety of educational data sources, including student assessments, attendance records, and demographic data.
3. Enhancing the curriculum: Big data analytics assist institutions in analysing student performance data to learn more about the efficacy of curricula and instructional methods. Institutions can improve student learning outcomes by identifying improvement areas and making data-driven changes to curriculum design, pedagogical techniques, and teaching methodologies.
4. Resource Allocation Optimisation: Big data analytics supports the efficient allocation of resources in educational institutions. Institutions can pinpoint areas where resources are being misused or underutilised by analysing data on student enrolment, facility usage, and budgeting information. This makes it possible to allocate resources effectively, which reduces costs and boosts operational effectiveness.

V. Challenges and Directives

Big data analytics in the educational sector comes with its own set of challenges that need to be addressed to fully harness the potential benefits. Some of the key challenges include:

a) Data Quality and Integration: One of the primary challenges is ensuring the quality and integration of data from various sources such as student information systems, learning management systems, assessment platforms, and other educational applications. Data inconsistency, incompleteness, and compatibility issues can hinder effective analysis and interpretation.

b) Data Privacy and Security: Educational institutions handle sensitive student data, and ensuring data privacy and security is crucial. Protecting student privacy, complying with data protection regulations, and implementing robust security measures are essential considerations. Balancing data access for analysis with safeguarding personal information poses a challenge.

c) Ethical Use of Data: Using big data analytics in education raises ethical considerations. It is essential to ensure that data is used in an ethical manner, respecting student rights and privacy. Care must be taken to avoid biases, discrimination, or unfair profiling when using data for decision-making or interventions.

d) Resource Constraints: Implementing big data analytics requires adequate technological infrastructure, skilled personnel, and financial resources. Educational institutions may face limitations in terms of funding, access to advanced analytics tools, and expertise to effectively utilize big data analytics.

e) Data Governance and Policies: Establishing clear data governance frameworks, policies, and protocols is essential for effective data management. Educational institutions need to define data ownership, establish data sharing agreements, and address issues related to data access, retention, and data sharing across different stakeholders.

f) Change Management and Adoption: Integrating big data analytics into educational practices requires a cultural shift and change management efforts. Building data-driven decision-making capabilities, promoting data literacy among teachers and administrators, and fostering a data-driven mindset can be a significant challenge.

g) Interpretation and Actionability: Extracting meaningful insights from large volumes of data and translating them into actionable strategies can be complex. Education professionals need to possess the analytical skills to interpret and effectively utilize data insights to drive improvements in teaching and learning practices.

h) Addressing these challenges requires collaborative efforts among educational institutions, policymakers, and stakeholders. Implementing data governance frameworks, investing in infrastructure and resources, providing training and professional development opportunities, and ensuring compliance with data privacy regulations are essential steps toward overcoming these challenges.

i) By addressing these challenges, the educational sector can unlock the full potential of big data analytics and leverage its contributions to enhance teaching and learning outcomes, personalize education, improve decision-making processes, and drive continuous improvement in the educational landscape.

VI. Conclusion

The survey on big data analytics for the educational sector has, in conclusion, highlighted the key contributions and issues connected with exploiting big data in education. Big data technology has advantages that have been identified by the investigation, including personalised learning, data-driven decision-making, early intervention, and resource optimisation. Additionally, it has noted important difficulties with data utilisation, data quality, contextual factors, roles of data operators, and data modalities. Educational institutions may fully utilise the power of big data to improve teaching and learning results by addressing these issues.Big data analytics can also be used in broader contexts including curriculum creation, instructional design, and educational policy-making in addition to just student performance analysis. Educational institutions can make evidence-based decisions, carry out focused interventions, and constantly enhance their instructional practises by utilising big data insights.The poll shows how powerful big data analytics can be in altering education. Educational institutions may improve the teaching and learning experience, student outcomes, and educational innovation by overcoming obstacles and utilising the potential of data-driven insights. A more individualised, effective, and efficient educational environment may be shaped in the future via big data analytics in education.

References:

[1] M. Grobelnik, Big – Data Tutorial, March 2018, [online] Available: https://www.planet-data.eu/sites/default/files/presentations/Big_Data_Tutorial_part4.pdfo.

[2] S. Xiaomeng, *Introduction to Big Data Institutt for informatikkog e-laeringved NTNU Learning Materials Course IINI3012 Big Data*, March 2018

[3] A. S. Drigas and P. Leliopoulos, "The Use of Big Data in Education", *International Journal of Computer Science Issues (IJCSI)*, vol. 11, no. 5, September 2014

[4] R. Scapin, *Learning Analytics in Education: Using Student's Big Data to Improve Teaching. IT Rep Meeting*, April 2012

[5] G.R. Bian, "Application of military Internet of things in ammunition support of ordnance warehouse [J]", *Packaging engineering*, 2017.
[6] Y. JZhang, "Review on information security and privacy protection of Internet of things [J]", *Logistics technology*, 2017.
[7] M. Lenzerini, "Data integration: a theoretical perspective", *Proc. of the 21st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of database systems*, pp. 233-246, 2002.
[8] LvTeng and Wang Qunshan, *Journal of Artillery Academy*, vol. 30, no. 6, pp. 101-103, 2010.
[9] X. Dong, A. Halvey and C. Yu, "Data Integration with Uncertainty", *The VLDB Journal*, vol. 18, no. 2, pp. 469-500, 2009.
[10] Chen Min, Shiwen Mao and Yunhao Liu, "Big data: A survey", *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171-209, 2014.
[11] Siemens George and Phil Long, "Penetrating the Fog: Analytics in Learning and Education", *EDUCAUSE review*, vol. 46, no. 5, pp. 30, 2011.
[12] Heiner Cecily, Neil Heffernan and Tiffany Barnes, "Educational data mining", *Supplementary Proceedings of the 12th International Conference of Artificial Intelligence in Education*, 2007.
[13] R. S. J. D. Baker, "Data mining for education", *International encyclopedia of education*, vol. 7, pp. 112-118, 2010.
[14] Anthony G. Picciano, "The Evolution of Big Data and Learning Analytics in American Higher Education", *Journal of Asynchronous Learning Networks*, vol. 16, no. 3, pp. 9-20, 2012.
[15] Shum Simon Buckingham and Rebecca Ferguson, "Social Learning Analytics", *Educational technology and society*, vol. 15, no. 3, pp. 3-26, 2012.
[16] Bach Craig, "Learning analytics: Targeting instruction curricula and student support", *Office of the Provost Drexel University*, 2010.
[17] BKP Conrad Tucker and A. Divinsky, "Mining student-generated textual data in moocs and quantifying their effects on student performance and learning outcomes", *2014 ASEE Annual Conference Indianapolis*, 2014.
[18] Kennedy Gregor et al., "Predicting success: how learners' prior knowledge skills and activities predict MOOC performance", *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. ACM*, 2015.
[19] El Ahrache, Sara Ibn et al., "Massive Open Online Courses: a new dawn for higher education?", *International Journal on Computer Science and Engineering*, vol. 5, no. 5, pp. 323, 2013.
[20] BlagojeviÄĞMarija and DanijelaMiloÅąeviÄĞ, "Massive Open Online Courses: edXvs Moodle MOOC", *5th international Conference on Information Society and technology*, 2015.