TRENDS AND APPLICATIONS OF BIG DATA IN EDUCATION

M. A. Raza[1], H. U. R. Kayani[2], A. A. Malik[3], M. Gul[4] and A. Suleman[5]

[1,2,4,5]The University of Lahore, Sargodha Campus
Corresponding author Email: dr_aftab_malik@yahoo.com

ABSTRACT: Big data technologies have facilitated innovation across various industries, including healthcare, technology, and education. The utilization of this innovation is now required within the education sector at all levels. It is a progressing framework named as "Education 4.0" to accommodate the numerous needs of this sector. Specifically, this article first describes the architecture of big data and its properties compatible to Education sector . The process of utilizing large educational data and data mining techniques in education are explained. To find solutions to diverse educational problems, an ever-increasing variety of mining tools have been assembled. This study also provides insights into these techniques and other data sources for gathering educational data. Several applications have been developed to promote the use of big data techniques in higher education and highlight its benefits for the sector, but there are still a number of issues that need to be considered. We will highlight the obstacles of using big data techniques in higher education.

INTRODUCTION

Currently, the global number of internet users reached a staggering 5.16 billion, resulting in a massive rise in the quantity of data collected on an ongoing basis. However, the challenge lies in effectively harnessing this vast amount of data. The COVID-19 pandemic, with its transformative impact, has accelerated the digital transformation process in higher education, aligning it with other industries. Researchers have shown a growing interest in this field of education since 2016, as evident by the Scopus database. Over the past decade, 352 publications have been published, with 98 of them (more than 27%) released in 2021 alone.These findings were discovered by searching the Scopus database for research publications with at least the keywords "Big Data" and "Higher Education". All of these elements inspire the authors of this research to explore Big Data's potential in this crucial area.(Ahaidous, Tabaa, & Hachimi, 2023).

Our online actions generate a sizeable digital data warehouse as we digitise education. A network of computers is always required for data processing because the size of data is huge enough that it can exceed the processing capability of computer.Ninety percent information in utilization nowadays was created within the final two years due to the speed at which they are being processed (IBM, 2013). The data variation is so vast that data sources include digital device usage logs, browsing data from the web and social media, geo-located photos, and audio data.(Wang, 2016)

Methodology: This article will use a research methodology approach employing two methods. The initial method will be hermeneutic, aiming to clarify the importance of big data in high-level education, the interested parties involved, and the instruments employed. Meanwhile, the second technique will involve conducting a comprehensive review of scholarly articles with widely recognized systematic literature review methodology, encompassing the systematic exploration, selection, examination, and integration of pertinent academic publications. To provide an overview of the implementation in top level education, we investigate the highly referenced articles pertaining to the subject throughout various periods. Subsequently, we conduct another search to gather the most up-to-date and applicable data that utilize big data mining in the educational domain. We then proceed to screen the articles based on their titles and abstracts to filter out the relevant ones
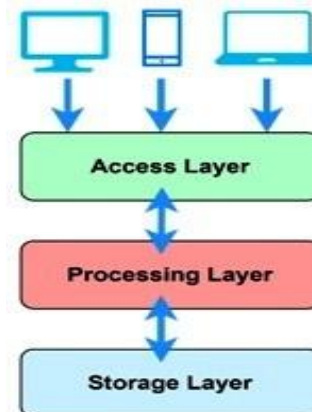

Figure 1 Big Data Framework

Characteristics of V's of Big Data: The 3 Vs (Volume, Variety & Velocity) are the first features related to big datadefinition that many academics pay close attention to, which prompts them to embed more Vs to definethe big data. Some researchers used the terms pillars as an alternative to the Big Data V's characteristics.Essential exploration into the realm of big data centers around a collection of 14V's, with the objective of effectively managing and harnessing the immense volumes of data at hand.(Kapil, Agrawal, & Khan, 2016)

The researcher continuously expands the scope of big data by incorporating additional V's into its definition. Eventually, the original three characteristics of big data evolve into a comprehensive set of fifty-six characteristics. (Hussein, 2020)(Muhammad Ali Raza, 2023)

Big Data Skeleton: The Big Data skeleton delineates the arrangement and structure of the systems employed to manage substantial massive size of data. The constituents of this structure based on the age of the data as its transitions from data sources to the final output for end users.(Ahaidous et al., 2023)

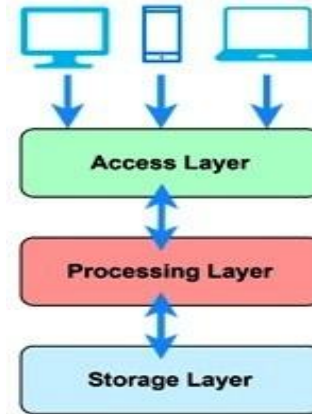The framework is categorised as a storage, processing, and Access layer. Presented in

Figure **1**

Storage Layer: Major concern to adopt this data in BDA is the sheer volume of data sources that need to be saved. Log files, SIS, and LMS can all provide data. For use when needed, each of these data kinds needs to be preserved and stored reliably. Hadoop is unquestionably still known as a Big Data solution for storing and analyzing data for a later time. Hadoop could be coupled with Business Intelligence tools or utilized solely as a storage system via HDFS.
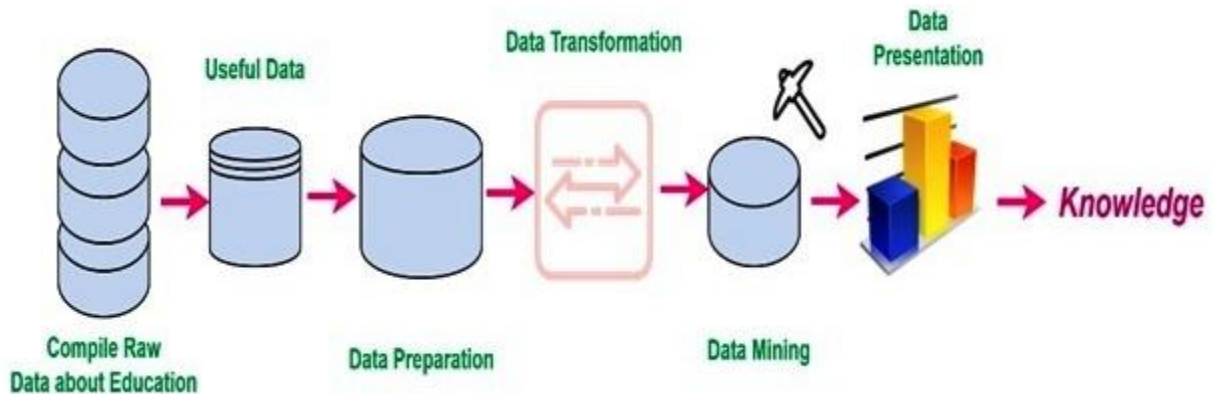

Figure 2:Process of Mining Educational Data

Processing Layer: The Big Data Analysis System (BDAS) second layer is their structural foundation. Batch processing and Stream processing are the two methods of processing.

Access Layer: Internet of Things (IoT) technology can help in analytical outcomes that can be observed on any web-connected device, encompassing laptops and intelligent timepieces.

Educational Aspects of Big Data: Big data technologies are used to obtain helpful information from enormous quantities of diverse, reliable, and increasing data.(Al-Kabi & Jirjees, 2019)Massive amount of data is being utilised to generate diverse applications for mining educational data and extracting insights from that

information, enhancing the intelligence of educational establishments such as schools and universities. The educational system data is considered as big, due to the volume and variety of information that is generated regularly about students attitude and interactions with learning platforms or systems, as well as learning activities, course information that varies from one another, and additional information that enhances the quality of educational processes.(Vaitsis, Hervatis, & Zary, 2016)

Big Opportunities in Education: Big data have much promise to change education. Our current generation of pupils has been using technology since their childhood. Their acquisition and everyday actions generate a

plethora of digital imprints, encompassing, but not restricted to, movements detected by motion detectors, keystrokes and mouse interactions conducted in front of computing devices, taps and gestures executed on mobile phones and tablets.

Although the massive data trend is still immature, butit has already demonstrated potential in education.(Wang, 2016)

Mining Process of Educational Data: Education-related data is collected from diverse origins across various educational settings, including conventional classrooms or alternative LMS. These sources include student records, behaviour logs, examination results, social media posts, administrative data, demographic data, and IOT data.(Khan, 2018) The education industry can generate a vast amount of data. As with any extensive data mining. The data mining of large amounts of educational data involves the following steps:
a) Compile the relevant raw data about education
b) Selection of valuable data.
c) Data preparation and cleansing.
d) Data transformation: removing anomalies and normalization.
e) Data mining: extraction of hidden patterns.
f) Data evaluation.
g) Acquire knowledge by analysing the outcome.

Contributors of Educational Data: Contributors areentities that generate the raw data and benefit from the knowledge extracted during the educational data mining process are the stockholders of educational data. Even if the data may be the same, each stakeholder can utilise it to obtain their goals(Khan & Alqahtani, 2020). The stakeholders can be classified as follows:
a) **Learners/Students:** According to them, most educational data is produced. They can also gain from data mining outcomes, such as personalised

online learning that suggests activities, resources, and courses help them learn more effectively.
b) **Educators/teachers:** These individuals may use data in a sense that will improve their methods of instruction. For example, they can analyze student behaviorsforecast students' academic performance, choose from various course organisation strategies, identify the common errors that students make and need to avoid, look for any unusual patterns, track students' learning progress, etc.(Shende, Thakare, Byagar, & Joshi, 2020)
c) **Course creators:** They can use it to assess the courses' outcomes, the students' satisfaction, and the course structures.
d) **Researchers:** Fact finders can assess the mining techniques, work with them to create tailored tools for mining, and contrast these tools to determine which is beneficial.
e) **Universities and Administrators:** They are in charge of allocating the necessary materials or human resources for execution. Additionally, they will gain from educational data mining in that they may improve decision-making, assist with the entire admissions process, review the teachers, etc.(Shende et al., 2020)

Tools for Mining Educational Data: Today, it is possible to mine datasets from any topic or field of study using a variety of commonopen-source and paid mining tools and frameworks.However, these tools were not created specifically to address pedagogical or educational issues.

Because they are made more for strength and versatility than for simplicity, they are difficult for a teacher to utilise. However, a galaxy of mining tools has been created that is primarily focused on finding solutions to various educational issues.(Cristobal Romero & Ventura, 2013)Table 1 shows the top-rated mining tools.

Table 1: Tools for mining educational data.

Tools	Goals	Reference
EPRules	This tool is Used for the identification of prediction criteria in order to give courseware authors feedback.	(Cristóbal Romero, Ventura, & Bra, 2004)
GISMO	It is used to see what is occurring in online courses	(Mazza & Milani, 2004)
LISTEN Mining tool	To examine extensive logs of student-tutor contact.	(Mostow et al., 2005)
MINEL	To evaluate the learner's performance and directional behaviour.	(Bellaachia, Vommina, & Berrada, 2006)
Data-Shop	A famous tool used for analysing click-stream and longitudinal data produced by educational systems.	(Koedinger, Cunningham, Skogsholm, & Leber, 2008)
SAMOS	To browse student activity with overview spreadsheets	(Juan, Daradoumis, Faulin, & Xhafa, 2009)
AHA! Mining Tool	To suggest the top websites that students should surf on the internet.	(Cristóbal Romero, Ventura, Zafra, & De Bra, 2009)

MMT tool	To make it simpler for new users to complete each phase of the Moodle data mining process	(Pedraza-Perez, Romero, & Ventura, 2011)
SNAPP	To chart the development of participant interactions in discussion forums	(Bakharia & Dawson, 2011)
E-learning Web Miner	Used to learn about the profiles related to attitude and working layout of students in virtual courses	(Zorrilla & García-Saiz, 2013)

Applications: The utilization of BDA in top-level education extends to various areas includingfollowing regulations, adhering to guidelines, managing student enrollments, processing admissions, planning finances, tracking donors, and monitoring student progress. These are just a few examples that illustrate the diverse scenarios where BDA can play a significant role and enhance the operational efficiency and decision-making processes within higher education institutions.(Picciano, 2012)

Student Success: MOOC are becoming more and more popular, presenting an opportunity for data collection and analysis. The data can be used to forecast student success or dropout rates by analysing the link between various variables. Additionally, the majority of institutions of higher education make use of ERP systems and LMS.(Kalota, 2015)

These two technologies both gather data that can be used for a variety of things.

By foreseeing the causes that could cause student dropout, analytics can be utilised to avoid dropout. For instance, using a sample of N = 14791, researchers used survival analysis to forecast student dropout rates. They discovered that a key predictor of student dropout was academic achievement.(Niemi & Gitin, 2012)

a) **Mobile Device & E-Book Reader:** E-textbooks and E-Book Readers are more widespread among students. The Horizon Report claims that it took less than a year for electronic books to become widely used. This presents a chance for further data mining. Publishers can collect and mine information about how books are used, course material, how content is presented, etc.

For instance, data related to education can be mined to track student patterns, such as the amount of time they spend on a specific website.(Consortium, 2011)

b) **Finance and Budgets:** Businesses employ market entry strategies to determine the best way to enter specific market sectors. They typically create various business models and evaluate the outcomes to see whether it is feasible to enter the market. Educational bodies can also gather data and use analytics to break into new sectors. They might also use these models to plan their subsequent operations.(Kalota, 2015)

c) **Performance Prediction:** Predicting performance, particularly for performances like course grades and test scores, is a typical use of big data in education. Reliable prediction results for each student can obtain and use for future applications by isolating correlations between elements (for example, student enrollment information) and specific performances.(Yu & Wu, 2015)

d) **Performance Presentation:** Big data is typically used in education to present activity performance, such as implementing new lesson plans. The presentation method involves gathering information from several sources and creating metrics like "students' reading proficiency" and "3- and 4-year-olds enrolled in preschool." Administrators can then assess the effectiveness of the activity and take appropriate action based on these metrics.(Yu & Wu, 2015)

e) **Understanding the Student's Learning Activity:** Finding trends in students' learning behaviours is another way big data is used in education. The pattern that was extracted can subsequently be used for other purposes, such as instruction adaption and cognitive process analysis. The system can determine the student's exploration approach and the relationship between the selected strategy and prior knowledge by analysing the data produced by the student's exploration actions.(Levy & Wilensky, 2009)

Data Sources from Instruction Frameworks Leading to Educational Data: Multiple sources, including student information systems, LMS, and library management systems, contain valuable data derived from educational frameworks. The emergence of the online instruction sector results from recent advancements in education, Information Technology applications, and web technology. With the growing number of higher education institutions embracing online instruction, there is a notable rise in accessibility of educational digital libraries, storage inventories and technological innovations.(Ang, Ge, & Seng, 2020).

Table 2: Source of educational data.

Data Source	Description of extracted data
Learning Management Systems.	Curriculum and course management, delivery, tracking, and reporting are all handled by Learning Management Systems (LMS), and education administration platforms. Considered as most widely used open-source LMS alternatives nowadays is Moodle.(Dougiamas & Taylor, 2003)
MOOC	Massive Open Online Courses isonline learning technology to attract many students worldwide. Text or video-based materials can be used to provide MOOC learning resources and information.(Ang et al., 2020)
OER	Open educational resources denote instructional resources that are readily available to the general public and fall under the purview of the public domain. These materials encompass authorised text, media, and digital resources employed for instruction, learning, and evaluation.(Ang et al., 2020)
Digital media platforms	Learners can express their academic background, feelings, and apprehensions regarding the educational journey via various social networking sites like Twitter, Facebook, and YouTube. Moreover, they can solicit social assistance from their fellow students within these online platforms. This wealth of digital data provides instructors with valuable knowledge and insights that can facilitate a deeper understanding of students' experiences beyond the confines of the traditional classroom setting.(Chen, Vorvoreanu, & Madhavan, 2014)
Linked data	Linked datais method of associating data that can be maintained in a distributed database in many locations using Internet technology. The authors of the cited study provided a methodical mapping of recommendations for using linked data to assist educational goals.(Pereira, Siqueira, Nunes, & Dietze, 2017)

Usefulness of Big Data in Education: Higher Education Institutions now have new duties and opportunities because of big data. Authors of cited articlesbelieves that big data presents best-researched foundation for managing enormous data collections and predicting subsequenthigher studies.(Siemens & Long, 2011) Big data deals with the investigation of a broad range of organisational data, gathered techniques intended to estimate formal performance and improve performance forecasting, as well as identify potential issues pertaining to learning, teaching, and research.(Hrabowski III & Suess, 2010)

Additionally, according to some academics, in order to increase productivity, higher education must include the analytics tools in their system. Big data is expected soon to benefit education in several helpful ways.(Hrabowski, Suess, & Fritz, 2011; Picciano, 2012)

a) **Assessment:** Having access to relevant details and being aware of the context can be beneficial when dealing with a large collection of learning data. A student may frequently perform poorly in a topic if they know the reason(s). When the student can examine oneself and other individuals who have gone through the same experience, it becomes beneficial. He or she may obtain insight into the situation, either to describe it and avoid becoming frustrated or to utilise it to repair it and achieve success once more.

The development of electronic learning modules encourages logical, on-the-spot evaluation of students. Various analytical software has the capability to provide students and educators with prompt feedback on their academic achievement.

b) **Motivation:** When big data analytics is implemented effectively, learners may demonstrate increased commitment to actively contributing data to the process driven by their recognition of potential impact and value.

c) **Tracking:** In order to fully grasp the actual learning patterns of students, Big data can be used by teachers to monitor a student's progress through an online course. They are observing the digital trails that students take too soon. Throughout the entire learning process, teachers can monitor their student's progress.

d) **Collaboration:** For a Learning Management System to continue to perform at its best, experts from many departments must collaborate. This promotes collaboration, teamwork, and cross-disciplinary thinking.

e) **Efficiency:** Large-scale data can conserve significant amounts of time and exertion when it pertains to achieving our objectives and the competencies, we must employ to accomplish them.

f) **Personalisation:** Big data can help us approach e-learning design successfully by enabling designers to customise courses to meet the unique demands of their learners. As a result, e-learning creators can promote the benchmark for excellent and effective e-learning courses.

g) **Perception of the learning process:** Teachers can identify whether portions of a course or examination were overly simple and which portions were so challenging that the student needed help to complete them with the assistance of big data in e-learning. After that, teachers can examine other aspects of the learner's journey, considering frequently revisited pages, preferred learning approaches, areas that peers have advised and the time of day when learning is most effective.

Implementation of Challenges in Higher Education: Although, applications are developing rapidly that support practical data approaches in Higher Education, a few concerns must be considered. The author of the cited article recommended looking at the three co-operating factors of timing, population and place, while gathering data for analytics(Becker, 2013). Any interval of time can be referred to as a timing element. The population is referenced by the attributes of students participating in the learning atmosphere. The location could be specified based on the learning atmosphere in which students recover knowledge.

It has been recognised that certain scholars are essential for students to possess the capacity to assess sources of information critically. Innovative facilities and an inspired attitude have also been recognised as crucial in a setting marked by change and complexity.(Sahlberg, 2009)

The financial implications could potentially emerge as a significant barrier hindering the widespread adoption of BDA in the realm of top level education.

Many organisations consider analytics to be more of an expense than an investment. Concern over affordability often centres on the alleged need for expensive data collection techniques.(Campbell, DeBlois, & Oblinger, 2007)

More money needs to be spent on hiring analytics experts, who efficiently utilize BDA correctly. Experts should be capable of monitoring the whole process, from identifying crucial inquiries to creating data models for creating and distributing warnings, suggestions, and reports.

The main issues associated with utilising large-scale data in education are data profiling, confidentiality and students' entitlements concerning documenting their particular behavior (Boyd, 2010) While learning analytics are essential totrack students' behaviour on a new level and scale and should be evaluated, the traditional classroom approach constantly analyses learners' performances and academic behaviour.

Implementing data masking techniques at the source level holds promising potential for addressing these challenges effectively. One innovative strategy that will enable large-scale data applications while yet preserving the privacy of student and teacher data is masking. New ETL software application capabilities and performances allow for its database-level masking when bringing sensitive data into a data warehouse.(Barlow, 2013) Big data can resolve major issues that higher education is currently dealing with and provide solid justification for anticipating, converting complex andfuzzy data into fruitful information. Big data also gives researchers a chance to understand the implications of data and provides them with a platform to analyse them in a substantial, efficient, and consistent way that advances theory and practice. Because of this, every educational institution needs to have a plan in place for utilizing big data.(Klašnja- Milićević, Ivanović, & Budimac, 2017)

Conclusion: Academic education and practice are significantly impacted by big data, as it contributes to uplift student experiences and understanding through enhanced academic studies, as well as enabling effective decision-making and scripted responses to changing trends. Big data has capability to deal with major challenges of Higher Education. It provides a robust justification for transforming complex data into valuable information.

In this article, we discussed the big data approaches, trends, applications, and tools that can be utilized to enhance the functionality of educational frameworks. The article underscored the importance of adopting a big data methodology in Higher Education, accentuating two primary implementations: managing reform initiatives in education & helping instructors in improving teaching skills.Nevertheless, the present era of expansive data systems faces multiple obstacles.

Moreover, educational institutions should strive to find a balance between their concept of student development and various government privacy regulations. Organizations need to acknowledge the dynamic nature of educational performance and retention, foster open and honest discussions, and enhance procedures and methods to address these issues. Educators, policymakers, and researchers will continue to make progress in their understanding of large-scale data. This understanding will enable the exploration of untapped opportunities for leveraging big data in education, as educational technologies are poised to expand. The analytical methodologies for large-scale data are continuously evolving and improving.

The findings presented in this study are expected to make a substantial contribution to the ongoing discourse on the development of a robust learning framework for students, teachers, curriculum designers, and educational institutions. By harnessing learning analytics, these research findings and their corresponding conclusions aim to drive the progress of effective educational practices.

REFERENCES

Ahaidous, K., Tabaa, M., & Hachimi, H. (2023). Towards IoT-Big Data architecture for future education. *Procedia Computer Science, 220*, 348-355.

Al-Kabi, M. N., & Jirjees, J. M. (2019). Survey of Big Data applications: health, education, business & finance, and security & privacy. *Journal of Information Studies and Technology, 2018*(2), 12.

Ang, K. L.-M., Ge, F. L., & Seng, K. P. (2020). Big educational data & analytics: Survey, architecture and challenges. *IEEE access, 8*, 116392-116414.

Bakharia, A., & Dawson, S. (2011). *SNAPP: a bird's-eye view of temporal participant interaction.* Paper presented at the Proceedings of the 1st international conference on learning analytics and knowledge.

Barlow, M. (2013). *Real-time big data analytics: Emerging architecture*: " O'Reilly Media, Inc.".

Becker, B. (2013). Learning analytics: Insights into the natural learning behavior of our students. *Behavioral & Social Sciences Librarian, 32*(1), 63-67.

Bellaachia, A., Vommina, E., & Berrada, B. (2006). *Minel: A framework for mining e-learning logs.* Paper presented at the Proceedings of the 5th IASTED International Conference on Web-based Education.

Boyd, D. (2010). *Privacy and publicity in the context of big data.* Paper presented at the Keynote Talk of The 19th Int'l Conf. on World Wide Web.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE review, 42*(4), 40.

Chen, X., Vorvoreanu, M., & Madhavan, K. (2014). Mining social media data for understanding students' learning experiences. *IEEE Transactions on learning technologies, 7*(3), 246-259.

Consortium, N. M. (2011). The 2011 Horizon Report.

Dougiamas, M., & Taylor, P. (2003). *Moodle: Using learning communities to create an open source course management system.* Paper presented at the EdMedia+ innovate learning.

Hrabowski, F., Suess, J., & Fritz, J. (2011). Assessment and analytics in institutional transformation. *UMBC Office of the Vice President of Information Technology*.

Hrabowski III, F. A., & Suess, J. (2010). Reclaiming the lead: higher education's future and implications for technology. *UMBC Office of the Vice President of Information Technology*.

Hussein, A. e. e. A. (2020). Fifty-six big data V's characteristics and proposed strategies to overcome security and privacy challenges (BD2). *Journal of Information Security, 11*(4), 304-328.

Juan, A. A., Daradoumis, T., Faulin, J., & Xhafa, F. (2009). SAMOS: a model for monitoring students' and groups' activities in collaborative e-learning. *International Journal of Learning Technology, 4*(1-2), 53-72.

Kalota, F. (2015). Applications of Big Data in education. *International Journal of Educational and Pedagogical Sciences, 9*(5), 1607-1612.

Kapil, G., Agrawal, A., & Khan, R. (2016). *A study of big data characteristics.* Paper presented at the 2016 International Conference on Communication and Electronics Systems (ICCES).

Khan, S. (2018). Modern Internet of Things as a challenge for higher education. *International Journal of Computer Science and Network Security, 18*(12), 34-41.

Khan, S., & Alqahtani, S. (2020). Big data application and its impact on education. *International Journal of Emerging Technologies in Learning (iJET), 15*(17), 36-46.

Klašnja- Milićević, A., Ivanović, M., & Budimac, Z. (2017). Data science in education: Big data and learning analytics. *Computer Applications in Engineering Education, 25*(6), 1066-1078.

Koedinger, K., Cunningham, K., Skogsholm, A., & Leber, B. (2008). *An open repository and analysis tools for fine-grained, longitudinal learner data.* Paper presented at the Educational data mining 2008.

Levy, S. T., & Wilensky, U. (2009). Crossing levels and representations: The Connected Chemistry (CC1) curriculum. *Journal of Science Education and Technology, 18*, 224-242.

Mazza, R., & Milani, C. (2004). *Gismo: a graphical interactive student monitoring tool for course management systems.* Paper presented at the International Conference on Technology Enhanced Learning, Milan.

Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., & Heiner, C. (2005). *An educational data mining tool to browse tutor-student interactions: Time will tell.* Paper presented at the Proceedings of the Workshop on Educational Data Mining, National Conference on Artificial Intelligence.

Muhammad Ali Raza, H. U. R. K. (2023). Review: Big Data V's Models, Challenges, Hadoop Ecosystem, Issues, Uses, Benefits and Applications.

Niemi, D., & Gitin, E. (2012). Using Big Data to Predict Student Dropouts: Technology Affordances for

Research. *International Association for Development of the Information Society*.

Pedraza-Perez, R., Romero, C., & Ventura, S. (2011). *A Java desktop tool for mining Moodle data*. Paper presented at the Proceedings of the 3rd conference on educational data mining.

Pereira, C. K., Siqueira, S. W. M., Nunes, B. P., & Dietze, S. (2017). Linked data in education: A survey and a synthesis of actual research and future challenges. *IEEE Transactions on learning technologies, 11*(3), 400-412.

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of asynchronous learning networks, 16*(3), 9-20.

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data mining and knowledge discovery, 3*(1), 12-27.

Romero, C., Ventura, S., & Bra, P. D. (2004). Knowledge discovery with genetic programming for providing feedback to courseware authors. *User Modeling and User-Adapted Interaction, 14*, 425-464.

Romero, C., Ventura, S., Zafra, A., & De Bra, P. (2009). Applying Web usage mining for personalizing hyperlinks in Web-based adaptive educational systems. *Computers & Education, 53*(3), 828-840.

Sahlberg, P. (2009). lifelong learning. *Lifelong learning in Europe, 1*, 53.

Shende, A., Thakare, M., Byagar, S., & Joshi, M. (2020). The review of different data mining tools, techniques and algorithms for the data mining in education. *Our Heritage, 68*(27), 290-296.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review, 46*(5), 30.

Vaitsis, C., Hervatis, V., & Zary, N. (2016). Introduction to big data in education and its contribution to the quality improvement processes. *Big Data on Real-World Applications, 113*, 58.

Wang, Y. (2016). Big opportunities and big concerns of big data in education. *TechTrends, 60*, 381-384.

Yu, X., & Wu, S. (2015). *Typical applications of big data in education*. Paper presented at the 2015 International Conference of Educational Innovation through Technology (EITT).

Zorrilla, M., & García-Saiz, D. (2013). A service oriented architecture to provide data mining services for non-expert data miners. *Decision Support Systems, 55*(1), 399-411.