Data-Driven Decision Making In Education

1Tejasvani Sharma, 2Dr Sherry Verma, 3Dr Dinesh Rai
1,2,3 School of Engineering & Technology, 1Sushant University, Gurugram, India

Abstract

This paper explores the role of data analytics and artificial intelligence (AI) in supporting data-driven decision-making in education. It analyzes how educational institutions can leverage student data, including academic performance, behavior, and socio- economic factors, to gain valuable insights and make informed decisions. The paper discusses the potential of data-driven decision- making in improving educational outcomes, enhancing student support systems, and optimizing resource allocation. Additionally, it examines the ethical considerations associated with data collection, analysis, and usage in the education sector, emphasizing the importance of privacy, data security, transparency, and equity. By reviewing relevant literature and case studies, this paper provides a comprehensive understanding of the benefits, challenges, and best practices in implementing data-driven decision-making in education. It concludes by highlighting the potential impact of data analytics and AI on shaping the future of education and suggests recommendations for educational institutions to effectively harness the power of data-driven decision-making.

Keywords - Academic performance, informed decisions, Ethical considerations, Data Security.

I.INTRODUCTION

The most crucial element in the present day is data, and despite the competition for control over the majority of the world's data appears to have been won by giants like Google, Amazon, Facebook, and Apple, it is being consolidated at a rapid pace. In the coming years, businesses will be highly dependent on their ability to analyze data in real-time and predict future outcomes and trends. Even a slight advantage in this capability by a matter of minutes could give a company a competitive edge. At present, the use of big data technology is common among large companies that have access to vast amounts of information and data.

Big data analytics refers to the process of analyzing large, complex sets of data to uncover valuable insights, patterns, and relationships. It involves utilizing advanced techniques and technologies to extract meaning from vast amounts of structured and unstructured data, including things like customer preferences, market trends, and hidden correlations. This can help businesses make better decisions, gain a competitive edge, and identify new opportunities.

Utilizing the results of analysis can result in various business advantages, such as more efficient marketing strategies, increased revenue generation, superior customer service, streamlined operations, and a competitive edge over other companies [13]. Currently, there is a significant amount of research being conducted in the field of big data, which involves studying extremely large datasets that can be analyzed using computational methods to uncover patterns, trends, and associations, particularly in regards to human behavior and interactions. A significant investment in information technology is currently directed towards managing and maintaining large data sets. This could bring advantages for companies that can collect, store, process and analyze large data sets to gain insights and make better decision making.

In recent years, and particularly today, we are constantly overwhelmed by an excessive amount of data. With the growth of technology, the internet, and the proliferation of connected devices, data is generated at an unprecedented rate, and we are currently in the era of big data. It became important for organizations and researchers to be able to collect, store, and analyze this data to extract insights and make more informed decisions. The Internet, which has its roots in the early 1960s, has grown tremendously over the years and has become a vast source of data [33]. Today, data is generated from a wide range of applications

and is collected at an unprecedented scale and speed. It ranges from a simple text, images, videos, and audio to complex multimedia, geospatial data and IoT sensor data. With the rapid growth of technology, the volume, variety and velocity of data continues to increase and it became vital for organizations and researchers to be able to process and analyze this data to extract insights and make more informed decisions. Big data analysis enables organizations to make decisions that were previously based on guesswork or on complex models, instead of on data itself. Big data is now the driving force behind many of the economic and social activities in modern society, such as mobile services (data, video, and financial), retail, engineering and manufacturing, finance, life sciences, and physical sciences [27]. In the field of life sciences, there is a well-established practice of depositing scientific data into public repositories and creating databases for use by other researchers. With the advent of Next Generation Sequencing (NGS) technology, the amount and number of experimental datasets have grown exponentially, increasing the need for big data analytics to manage, store and analyze these datasets. Furthermore, the field of bioinformatics has emerged which focuses specifically on the collection, organization, and analysis of life sciences data. This field is of great importance as it enables researchers to analyze and understand the vast amounts of data being generated by modern life science experiments and technological advancements, such as next-generation sequencing, which allows for the rapid and cost-effective generation of large amounts of genomic data. This discipline plays a vital role in the advancement of modern biotechnology and medicine, by providing valuable insights and new understanding of biological systems and potential new drug development.

Other scientific fields have also been revolutionized by big data, such as the Digital Sky Survey which has become a crucial resource for astronomers worldwide [31]. Astronomy, which once mainly involved taking pictures of the sky, has been transformed. Now the pictures are stored in a database, and the astronomer's task is to find interesting objects and phenomena in the database using big data analytics tools.

Currently, the potential benefits of data-driven decision making using big data are being increasingly recognized as a major contributor to economic growth. For instance, companies like Google are seen as engines of economic growth, as evidenced by Google's 2015 Economic Impact report, which found that the company contributed $165 billion in economic activity for 1.5 million businesses and non-profits across the country, an increase from $131 billion the previous year. The use of Google's big data, for instance through targeted ads and search results, leads to more clicks and ultimately more business. This in turn leads to more revenue and job creation. It's a clear indication that the use of big data can bring significant benefits to companies and the economy as a whole.

Despite the potential benefits that big data can bring, there is still a significant gap between the potential and its realization. Some of the major challenges in making progress with big data include dealing with heterogeneous data types, dealing with the scale of the data, ensuring that the data is timely and relevant, dealing with the complexity of the data, and addressing privacy concerns. These challenges affect all phases of the process of creating value from data, from data acquisition to analysis and decision making. One of the key issues is acquiring data in an organized and structured way, which currently happens in an ad-hoc manner.

A big portion of the data is not natively stored in structured format, for instance, text from blogs and tweets are unstructured and images and videos are structured for storage and display but not for semantic content and search. It poses a huge challenge in converting such data into structured format for analysis. Thus, effectively dealing with these challenges is crucial for effectively realizing the potential of big data.

One of the most important aspects of big data is that the value of data increases significantly when it can be linked with other data. Data integration is therefore a major contributor to

SUSHANT UNIVERSITY ORGANIZED INTERNATIONAL CONFERENCE ON
"ADVANCES IN MULTIDISCIPLINARY RESEARCH AND INNVOATION" ICAMRI-2023
ON 28-29TH OCTOBER 2023
INTERNATIONAL ADVANCE JOURNAL OF ENGINEERING, SCIENCE AND MANAGEMENT (IAJESM)
July-December 2023, Submitted in October 2023, iajesm2014@gmail.com, ISSN -2393-8048
Multidisciplinary Indexed/Peer Reviewed Journal. SJIF Impact Factor 2023 =6.753

value creation. Since most data is generated in digital format today, this presents both an opportunity and a challenge to influence data creation to facilitate later linkage, and to automatically link previously created data. However, there are also other foundational challenges such as data analysis, organization, retrieval, and modeling. Among these, data analysis is often considered a major bottleneck in many applications. This is due to the lack of scalability of the underlying algorithms and the complexity of the data that needs to be analyzed. Additionally, the presentation of the results and its interpretation by non-technical domain experts is crucial in order to extract actionable knowledge.

It's worth noting that over the last four decades, data management principles such as physical and logical independence, declarative querying, and cost-based optimization have formed the basis of a multi-billion-dollar industry. These technical advancements have allowed for the successful implementation of business intelligence applications and have laid the foundation for managing and analyzing big data. As big data poses new challenges and opportunities, it is important to rethink many aspects of these data management platforms while retaining their desirable features. It is believed that appropriate investment in big data will lead to a new wave of fundamental technological advances, which will be seen in the next generations of big data management and analysis platforms, products, and systems.

It is believed that the research problems associated with big data are not only important in the current time, but they also have the potential to create significant economic value in the US and around the world for years to come. However, these problems are also complex and difficult, which means that we need to rethink data analysis systems in fundamental ways. Nevertheless, a well- directed significant investment in big data can lead not only to major scientific advances but also lay the foundation for the next generation of advancements in science, medicine and business.

As Edwards Deming once stated, "without data, you are just another person with an opinion". But Deming couldn't have imagined the size and speed of data systems we have today. Automation that relies on continuously gathered data is changing our daily lives. For example, drivers don't need to know how to use maps anymore, when they can use smart navigators that find them the best routes and similarly, airline pilots spend more time flying on autopilot than by hand. Similar trends are happening in education systems with countless reformers trying to "disrupt" schools as they are known today. The current big demand is not only for increased access to valuable data but also for tools to automate data collection, archiving, and analysis. Most industry leaders, academicians, and other prominent stakeholders are in agreement that big data has become a major game changer in most, if not all, types of modern industries and how we do things over the last few years.

II. ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) or Machine Learning technology, which has been around for several decades, is once again gaining popularity, as can be seen in the widespread use of bots on the internet. Today, AI is becoming increasingly practical and is becoming an integral part of everyday life, becoming relevant across almost all sectors of the economy such as financial and banking services, healthcare, transportation, defense, and many others. Additionally, with the wide range and almost endless variety of AI applications, we are at the beginning of a new era of automation powered by machine-to-machine interconnection or a machine economy.

The use of data in education started to gain traction around 8 years ago, with the rise of the field of learning analytics. This field focuses on the measurement and analysis of student data in order to improve the learning experience and the learning environment. With the help of big data, education is becoming more personalized, more efficient, more affordable and more accessible for students. The utilization of large-scale student data allows for more efficient tracking of student progress, more accurate identification of at-risk students, and more

SUSHANT UNIVERSITY ORGANIZED INTERNATIONAL CONFERENCE ON
"ADVANCES IN MULTIDISCIPLINARY RESEARCH AND INNVOATION" ICAMRI-2023
ON 28-29TH OCTOBER 2023
INTERNATIONAL ADVANCE JOURNAL OF ENGINEERING, SCIENCE AND MANAGEMENT (IAJESM)
July-December 2023, Submitted in October 2023, iajesm2014@gmail.com, ISSN -2393-8048
Multidisciplinary Indexed/Peer Reviewed Journal. SJIF Impact Factor 2023 =6.753

effective and efficient use of resources, as well as allows for development of more tailored curricula, and so on. This is likely to bring significant improvements to the education system, and open up new opportunities for learners and educators alike.

A study found that the use of data to guide instruction is one of the top five policies correlated with measurable academic effectiveness, based on a quantitative comparison of different approaches taken by 35 charter schools in New York City. According to the study, collecting and analyzing data on student performance, and using it to inform instructional decisions and adjustments, is a key strategy for improving student outcomes [28]. The study suggests that this approach can be highly effective, and emphasizes the importance of data-driven instruction in achieving academic success. If we had access to a vast educational database that collected detailed measures of every student's academic performance, we could use that data to design the most effective approaches to education. The collected data could be used to improve instruction at all levels, from reading and writing to STEM (science, technology, engineering and mathematics) and advanced college-level courses. It is likely that in the future, we will have access to such data, as there is a growing trend towards the massive deployment of educational activities on the web, which is expected to generate an increasingly large amount of detailed data about student performance. This will allow school administration to proactively manage students' wellbeing, and improve the overall educational experience.

In the higher education sector, the big data revolution has provided universities with large amounts of data on student behavior and performance, which can be tracked and analyzed through online learning systems. However, it is important for the education sector to keep in mind the ultimate goal: to actually help students learn and improve their performance outcomes.

Big Data and predictive analytics are now being used by educational institutions and related companies to gain insights they previously did not have. In terms of student enrollment, these technologies allow for the creation of analytical databases that provide college administrators with fast and actionable information. This information can be used to make smart enrollment decisions and to allocate resources to increase enrollment from existing markets as well as to expand into new ones. From the perspective of students, big data and predictive analytics can also be used to help them find colleges that are the best fit for their profile, by allowing them to shortlist institutions based on various factors such as academic programs and financial aid. This usage of big data and predictive analytics in the education field has the potential to benefit both the institutions and the students by providing more accurate and efficient enrollment management and finding the best fit college for students. Online learning systems generate vast amounts of data, which can be overwhelming and difficult to analyze. However, mining this data can provide valuable insights into how to improve student learning outcomes. But before that, it's important to understand what the data is being collected for, which means having a better understanding of how to improve student learning outcomes. The question that needs to be asked is: what specific information or insights are we trying to gain from this data? Without a clear understanding of this, it can be difficult to determine where to focus our analysis and what data is most relevant. Therefore, it's important to have a clear research question and goals in mind when collecting and analyzing data in education to make sure the data is useful and actionable.

Most educational institutions use big data to address specific issues such as student retention, providing students with feedback on their learning, and identifying students who are at risk of falling behind or dropping out as early as possible so that proactive interventions can be taken to help them stay on track. To make the most of big data, it's crucial for institutions to have clear student key performance indicators. This helps in aligning the outcomes of the analysis of big data with broader institutional priorities. Overall, the key challenge when using big

SUSHANT UNIVERSITY ORGANIZED INTERNATIONAL CONFERENCE ON
"ADVANCES IN MULTIDISCIPLINARY RESEARCH AND INNVOATION" ICAMRI-2023
ON 28-29TH OCTOBER 2023
INTERNATIONAL ADVANCE JOURNAL OF ENGINEERING, SCIENCE AND MANAGEMENT (IAJESM)
July-December 2023, Submitted in October 2023, iajesm2014@gmail.com, ISSN -2393-8048
Multidisciplinary Indexed/Peer Reviewed Journal. SJIF Impact Factor 2023 =6.753

data in education is how to effectively use it to improve learning. This can be achieved by focusing on the specific issues that institutions want to address, setting clear goals, and selecting appropriate key performance indicators.

III. ETHICAL CONSIDERATIONS

When working with data in the real world, it is crucial to address the ethical considerations associated with the collection, analysis, and usage of student data. Three key ethical considerations include privacy and data protection, transparency and explain ability in AI-based decision-making, and addressing equity concerns. Privacy and data protection are of utmost importance when dealing with student data. As we implement the code and work with real-world datasets, ensuring that it complies with data privacy regulations and institutional policies. Take necessary precautions to protect student data, including securing storage systems, anonymizing data where possible, and limiting access to authorized personnel. Implement strong data encryption methods and regularly audit data security practices to mitigate the risk of data breaches.

Transparency and explain ability in AI-based decision-making are essential for maintaining trust and ensuring accountability. As we develop and deploy AI models, document the entire process, including data preprocessing, feature engineering, model selection, and evaluation metrics. Provide clear explanations of how the models make predictions or decisions based on the data. This transparency helps stakeholders understand the rationale behind decisions made using the models and guards against the potential biases or unfair treatment that AI systems may inadvertently introduce.

Addressing equity concerns is crucial to avoid perpetuating disparities in educational opportunities. Be mindful of potential biases in the data and models to be used. Perform regular audits and assessments to identify and mitigate biases that may affect decision-making processes. Ensure that data analysis and decision-making processes promote fairness, inclusivity, and equal opportunities for all students, regardless of their background or characteristics.

IV. BEST PRACTICES AND CHALLENGES

Implementing data-driven decision-making in educational institutions requires careful consideration of best practices and an awareness of the challenges that may arise. This section explores case studies highlighting successful implementations, identifies common challenges and barriers, and presents strategies and best practices for overcoming these challenges.

Case study:

Several educational institutions can implement data-driven decision-making to enhance student outcomes. For instance, School District A can utilize predictive analytics models to identify students at risk of dropping out. This can enable them to implement proactive interventions, such as personalized counseling and targeted academic support, resulting in a significant decrease in dropout rates. Similarly, College B ca leveraged data analytics to optimize course offerings based on student preferences, leading to increased student enrollment, higher satisfaction rates, and improved retention.

Challenges and barriers:

Implementing data-driven decision-making in education can face various challenges. Limited data quality and availability may pose challenges in accessing accurate and comprehensive student data. Insufficient data literacy and analytical skills among educators and administrators can hinder effective data utilization. Additionally, data silos and lack of interoperability across systems can impede the integration and analysis of diverse datasets. Ethical considerations, such as ensuring privacy and data protection, may present challenges in collecting and using student data responsibly. Resistance to change and a culture that does not prioritize data-driven decision-making can also be significant barriers.

Strategies and best practices for overcoming challenges:

To overcome challenges and maximize the effectiveness, educational institutions can adopt several strategies. First, it is crucial to establish a strong data governance framework, encompassing policies, procedures, and guidelines for data collection, management, and usage. This includes ensuring data privacy and security, promoting data quality and standardization, and fostering data sharing and collaboration across stakeholders.

Investing in professional development and training programs can enhance data literacy and analytical skills among educators and administrators. Providing opportunities for learning data analysis techniques and tools enables staff to effectively analyze and interpret data, empowering them to make informed decisions. Collaborative partnerships with external experts, such as data scientists or researchers, can also supplement internal capabilities and enhance data-driven decision-making.

To address data silos and interoperability challenges, educational institutions should prioritize data integration efforts. Implementing robust data systems that facilitate the seamless flow of information across different platforms and departments enables comprehensive data analysis. This integration should be accompanied by the use of appropriate data management tools and technologies to ensure data integrity and accessibility. Continuous monitoring, evaluation, and refinement of data practices ensure the long-term success and sustainability of data-driven initiatives in education.

V. FUTURE ASPECT

As the field of education continues to evolve, the potential impact of data analytics and artificial intelligence (AI) is significant. Leveraging these technologies through data-driven decision-making can revolutionize educational practices and outcomes. This section explores the potential impact of data analytics and AI on education, provides recommendations for educational institutions to effectively harness data-driven decision-making, and summarizes the key points discussed.

To effectively harness data-driven decision-making, educational institutions can consider the following recommendations:

1. Establish robust data governance frameworks: Develop clear policies for data collection, storage, and usage. Ensure compliance with privacy regulations, data security measures, and ethical considerations. Promote transparency and accountability in data practices.

2. Foster data literacy and skills development: Provide professional development opportunities for educators and administrators to enhance data literacy and analytical skills. Encourage continuous learning and upskilling in data analysis techniques and tools.

3. Invest in data infrastructure and systems: Implement data systems that facilitate data integration, interoperability, and data sharing across platforms and departments. Ensure data quality, standardization, and accessibility for effective analysis and decision-making.

4. Promote collaboration and interdisciplinary approaches: Foster partnerships between educators, policymakers, data scientists, and researchers to leverage expertise and insights. Encourage interdisciplinary collaboration to address complex educational challenges and promote innovation.

5. Embrace a culture of data-driven decision-making: Foster a culture that values and prioritizes data-driven insights. Provide leadership support, communicate the benefits of data-driven approaches, and involve stakeholders in decision-making processes. Encourage experimentation and learning from data-driven initiatives.

VI. CONCLUSION

The potential impact of data analytics and AI in education is substantial. Through data-driven decision-making, educational institutions can improve educational outcomes, enhance student

support systems, and optimize resource allocation. By effectively harnessing student data, educational institutions can personalize instruction, provide targeted interventions, and create inclusive learning environments. However, it is crucial to address ethical considerations, such as privacy, transparency, and equity, to ensure responsible implementation. By adopting the recommended strategies and practices, educational institutions can embrace the future of education and leverage the power of data analytics and AI to promote student success and shape the educational landscape.

In summary, data analytics and AI offer immense potential to revolutionize education. With careful implementation, educational institutions can leverage data-driven decision-making to enhance learning experiences, improve student outcomes, and foster a data- driven culture that prepares students for the future.

REFERENCES

[1] "How secure is blockchain technology," Blockchain & AI Technology, https://www.blockchaincommunity.in/blog/how-secure- is-blockchain-technology (accessed May 29, 2023).

[2] Admin, "Indian Education System," Competition Success Blogs, https://www.competitionreview.in/blogs/2023/05/12/indian- education-system-2/ (accessed May 29, 2023).

[3] "The benefits of applying blockchain technology in any industry," BLOCKCHAIN & AI TECHNOLOGY, https://www.blockchaincommunity.in/blog/the-benefits-of-applying-blockchain-technology-in-any-industry (accessed May 29, 2023).

[4] "A comprehensive guide to understanding blockchain technology: Skill-Lync Blog," Skill, https://skill-lync.com/blogs/a- comprehensive-guide-to-understanding-blockchain-technology (accessed May 29, 2023).

[5] T. R. Reddy, P. V. G. D. P. Reddy, R. Srinivas, Ch. V. Raghavendran, R. V. S. Lalitha, and B. Annapurna, "Proposing a reliable method of securing and verifying the credentials of graduates through blockchain," EURASIP Journal on Information Security, vol. 2021, no. 1, pp. 1-15, 2021.

[6] D. Lee and N. Park, "A Blockchain-based Untact Education System for the Post-COVID-19 Era," Ilkogretim Online - Elementary Education Online, vol. 20, no. 3, pp. 2035-2049, 2021.

[7] A. Amitkumar, M. I. Sanni, and D. Apriliasari, "Blockchain Technology Application: Authentication System in Digital Education," Aptisi Transactions on Technopreneurship, vol. 3, no. 2, pp. 112-117, 2021.

[8] "Ai in 2020: From experimentation to adoption," THINK Blog, https://www.ibm.com/blogs/think/2020/01/ai-in-2020-from- experimentation-to-adoption/ (accessed May 29, 2023).

[9] I. Yakovenko, L. Kulumbetova, I. Subbotina, G. Zhanibekova, and K. Bizhanova, "The blockchain technology as a catalyst for digital transformation of education," International Journal of Mechanical Engineering and Technology (IJMET), vol. 10, no. 3, pp. 791-802, 2019.

[10] K. Rabah, "Overview of Distributed Ledger Technology: A Review," _Realm International Journal of Information Security_, vol. 1, no. 1, pp. 1-9, 2018.

[11] K. Gusak, "The blockchain confusion: issues hindering blockchain adoption," Linnaeus University Institutional Repository, 2018.

[12] M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić, "EduCTX: A Blockchain-Based Higher Education Credit Platform," IEEE, 2018.

[13] S. Vaidyanathan,"What are the industry applications of big data analytics?"
[Online].
Available:https://www.researchgate.net/post/What_are_the_industry_applications_of_big_data _analytics. [Accessed: Jul. 3, 2017].
S. Vaidyanathan,"What are the industry applications of big data analytics?"

[14] J.Epstein,"When blockchain meets big data, the payoff will be huge," [Online]. Available: https://venturebeat.com/2017/07/30/when-blockchain-meets-big-data-the-payoff-will-be-huge/.

SUSHANT UNIVERSITY ORGANIZED INTERNATIONAL CONFERENCE ON
"ADVANCES IN MULTIDISCIPLINARY RESEARCH AND INNVOATION" ICAMRI-2023
ON 28-29TH OCTOBER 2023
INTERNATIONAL ADVANCE JOURNAL OF ENGINEERING, SCIENCE AND MANAGEMENT (IAJESM)
July-December 2023, Submitted in October 2023, iajesm2014@gmail.com, ISSN -2393-8048
Multidisciplinary Indexed/Peer Reviewed Journal. SJIF Impact Factor 2023 =6.753

[15] K. Rabah, "Agricultural Food Supply Chain Powered by Big Data and Blockchain: A Review," _Maraint. j. sci. res. publ._, vol. 1, no. 1, pp. 55-64, 2017.

[16] K. Rabah, "Challenges & Opportunities for Blockchain Powered Healthcare Systems: A Review," _Mara res. j. med. health sci._, vol. 1, no. 1, pp. 45-52, 2017.

[17] K. Rabah, "Challenges & Opportunities of Implementing Blockchain Security & Privacy," _Mara International Journal of Computer Science & Information Security_, vol. 1, no. 1, pp. 1-11, 2017.

[18] K. Rabah, "Blockchain and the Law: A Review," _Mara Research Journal of Law_, vol. 1, no. 1, pp. 22-30, 2017.

[19] K. Rabah and J. Oguta, "Integrated RFID-IoT Powered Car Parking Management System for Smart Cities," _Mara Res. J. inf. Sci. Technol._, vol. 2, no. 1, pp. 10-16, 2017.

[20] M. Iansiti and K. R. Lakhani, "The Truth About Blockchain," _Harvard Business Review_, vol. 95, no. 1, pp. 118-127, 2017.

[21] K. Rabah, "Overview of Blockchain as the Engine of the 4th Industrial Revolution," _Mara Res. J. Bus. Manag._, vol. 1, no. 1, pp. 125-135, 2016.

[22] K. Rabah, "Digital Cryptoeconomics Powered by Digital Cryptocurrency," _Mara Res. J. Comput. Sci. Inf. Secur._, vol. 1, no. 1, pp. 107-131, 2016.

[23] K. Lofgren, "Big data and public service delivery - big hype?" [Online]. Available: http://www.victoria.ac.nz/news/2016/09/big-data-and-public-service-delivery-big-hype.

[24] P. Sahlberg and J. Hasak, "'Big data' was supposed to fix education. It didn't. It's time for 'small data'," [Online]. Available: https://www.washingtonpost.com/news/answer-sheet/wp/2016/05/09/big-data-was-supposed-to-fix-education-it-didnt-its-time- for-small-data/.

[25] Hei and Songtao, "A Decision Tree Based Approach to Filter Candidates for Software Engineering Jobs Using GitHub Data," Vanderbilt University Institutional Repository, 2015.

[26] M. Gao, "Application of machine learning techniques to recommendation system," University of California Institutional Repository, 2015.

[27] H. V. Jagadish et al., "Big data and its technical challenges," _Communications of the ACM_, vol. 57, no. 7, pp. 86-94, 2014.

[28] W. Dobbie and R. G. Fryer Jr, "Getting beneath the veil of effective schools: Evidence from New York City," _American Economic Journal: Applied Economics_, vol. 5, no. 4, pp. 28-60, 2013.

[29] S. Borkar and K. Rajeswari, "Predicting Students Academic Performance Using Education Data Mining," International Journal of Computer Science and Mobile Computing, vol. 2, no. 9, pp. 398-404, 2013.

[30] J. Manyika et al., "Big data: The next frontier for innovation, competition, and productivity," _McKinsey Global Institute_, May 2011.

[31] D. J. Eisenstein et al., "SDSS-III: Massive spectroscopic surveys of the distant universe, the Milky Way, and extra-solar planetary systems," _The Astronomical Journal_, vol. 142, no. 3, p. 72, 2011.

[32] K. Rabah, "Review of Method for Integer Factorization Applied to Cryptography," _JAS_, vol. 6, no. 2, pp. 458-481, 2006.

[33] R. Kahn et al., "The evolution of the Internet as a global information system," _The International Information & Library Review_, vol. 29, no. 2, pp. 129-151, 1997.

[34] L. Corrin, G. Kennedy, and P. de Barba, "Asking the Right Questions of Big Data in Education," [Online]. Available: https://pursuit.unimelb.edu.au/articles/asking-the-right-questions-of-big-data-in-education.

[35] T. McConaghy, "Blockchains for Big Data from Data Audit Trails to a Universal Data Exchange," [Online]. Available: https://blog.bigchaindb.com/blockchains-for-big-data-from-data-audit-trails-to-a-universal-data-exchange-cf9956ec58ea.

[36] D. Smyth, "Why blockchain? What can it do for big data?" [Online]. Available: http://bigdata-madesimple.com/why- blockchain-what-can-it-do-for-big-data-2/.