Big Data Analytics Applicability in Higher Learning Educational System

Tasmin, R. ¹, Muhammad, R. N ², and A. H. Nor Aziati ¹

¹ Department of Technology Management, Universiti Tun Hussein Onn Malaysia, 86400, Parit Raja, Johor.
² Midland School of Business and Finance, Abuja, Nigeria.
Corresponding author: rosmaini@uthm.edu.my

Abstract. The application of big data analytics is significantly gaining awareness in the educational sector as universities are operating under pressure in an increasingly competitive environment. Quite a number of researchers have reported the use and application of big data in different fields with few publications addressing the integrative application of big data analytics in institutions of higher education. The purpose of this paper is to explore the applicability of big data analytics in higher institutions. This paper outlines the major areas in which big data analytics can be of use in the higher institutions such as in academic and management levels. The ultimate goal is to optimize using big data analytics in learning and decision-making.

1. Introduction

Over the years with the advancement in areas of information technology, artificial intelligence, and machine learning, the amount of information been processed is continually increasing at an unprecedented rate [1]. Due to this huge deposit of information and the difficulty in processing, the technique of big data was developed. The technique involves the process of analyzing large data in a timely, simplified, and efficient manner [2]. Today, the use of big data technique has been gaining attention in several application domains. IT companies, small and large scale business enterprise, health care systems, sports, securities, among others, have been adapting the concept and techniques of big data in enhancing their activities. For this purpose, data science has become one of the most sort profession globally [3]. Big data analytics is defined as the processing of vast amount of data using mathematics and statistical modeling, programming and computing algorithm techniques for finding actionable value [4]. Big data applications are meant to gain insights, discover patterns, and identify abnormalities across numerous dataset within an organization [5]. Most importantly, big data analytics use in organizations will improve decision making for managers, and provide better customer services in business, optimize resources, and better understanding of customer engagement patterns [6].
Many types of research have been carried out with big data in several fields such as business companies tracing their contents on several social media networks to perform analysis, public-sector organizations, healthcare systems monitoring various networks and research to evaluate and treat various epidemics [7]. Some of the reports on big data use in various areas include the availability of data, its relevance, benefit, opportunities, cost, and ownership and also, its applicability [8]. However, there are few reports and publications on the use of big data in the higher education system (HEIs) [9].

Decision-making is a fundamental feature in every organization, which can be viewed based on organization’s perception. Generally, data-driven decision-making is the primary measure considered as a good management by organization. In higher education sector, every institution has their unique way towards decision-making approach aiming at upholding its operations. Typically, it is the responsibility of the executives of higher education to make decision, which is usually based on the executives’ intuition and experience [10]. [11] carried out a survey on 380 senior managers in IT departments with the objective of understanding the use of data and how it improves the processes of decision-making. The result shows that data-based decision-making in higher education is a significant basis of sustainable competitive advantage.

Similarly, [11] conducted multiple research, in which the first study investigated how administrators of faculty at a higher institution in California used student data in decision-making. The second study was the use and influence of data at large at the institution. Both studies indicated that although institutions have data management approach, however, they are yet to attain its full gain. Therefore, with the potential of big data technologies, organizational tasks, importantly in the HEIs can optimize some of the challenges especially in the area of strategic management and policy. HEIs is a good application area for big data analytics as there are untapped potentials where the technique could be applicable. The universities have large amount of data such as registration data, alumni data, course outline data, students’ data, assessment data, tutor data, among others. Applying big data in mining could help provide easy access and processing which will then help to enhance the workflow in the educational system. The analytics of big data in HEIs is broadly categorized into two: academic and learning.

Academic analytics is broadly referring to as data-driven decision-making practices of analysing institutional data (large data sets, statistical techniques, and predictive modeling) to generate actionable intelligence for operational purposes in an education institution [12]. It has the potential of cost-benefit to the institution, the increment in the student recruitment and a better quality of education which will provide a competitive advantage for institutions. On the other hand, learning analytics is the analysis of a large data sets of student’s data which is generated and collected to measure academic progress, predict future performance, and site possible difficulties [13]. It has significant positive influence on tutoring, curriculum, assessment, and decisions which are usually made at the classroom, university, or policy level by academicians, administrators, and policymakers. With the emergence of big data and its applications, collection and storage of data are held in incredible magnitudes. The application of big data has had massive positive impact on solving real-world scenarios. [14] presented in their findings that Facebook in about two years has around 500 million users. On the social media, information is being processed almost every seconds. [15] pointed out that within a short time, companies globally will have stored data exceeding seven exabytes.

Despite the benefits of big data, there are some challenges concerning its usage and application. [16] highlighted in his study some significant preliminary obstacles of big data which includes; workers who are from technological or business fields but do not have adequate skills of big data, inadequate infrastructure, and handling of categories of data from diverse sources (unstructured and semi-structured data). Others include an indecisive choice of technique and methodologies to analyze data to be able to overcome these obstacles and the problem of agreeing and understanding between an organization and a third party (the hiring of expert outside an organization to manage big data). Also, the problem within a collaboration between stakeholders, security and privacy among others. In this paper, the applicability and impact of big data analytics in institution of higher education is explored. Therefore, the objective of this paper is to fill the gap of big data uses in universities by exploring the applicability of Big Data analytics in institutions of higher education. The remainder of this paper is organized as follows. Section 2 presents the works related to Quranic text classification, Section 3
presents the methodology with detailed account of the dataset, classification experiment as well as the evaluation metrics, Section 4 reports the experiment results, and finally Section 5 concludes with some directions for future work.

2. **Big Data**

Big data is based on the concept that voluminous data can be processed, analyzed and treated using big data analytics. There is no specific definition of big data yet, as many experts define big data as its suite their purpose. [17] stated that big data is an asset of Information (data) considered by its high volume, variety, and velocity which require specific technology and analytical approaches for its revolution into value. Likewise, [2] reviewed that big data refer to the concept that the volume of data requires much more robust technologies, techniques and people with new skills to treat, process and analyze large dataset in a simplified, timely way. Therefore, the technology of big data can help organizations to sustain and gain competitive advantage internationally. Big data analytics is correlated with business intelligence and is supported in statistical analysis and data mining [17]. It is on the basis of data-driven decision making [18].

Recently, more than 75% of companies exceptionally private organizations have already started implementing the concept of big data in order to optimize their strategies and support improvement of new products and services [19]. Big data analytics enable organizations to manage large amount of data within a short time. This helps to gain competitive advantage by allowing management to have better handling of data processing [20]. Big data analytics have the potential to advance the performance, provides decision making support to facilitate innovation in the products and services of business standards [21]. Furthermore, the concept involves data discovery, visualization, and advanced analytics [3]. There are several standard techniques organizations use for big data analytics such as: MapReduce, Scalable Hadoop, Spark, YARN (for batch processing engine), Apache Kafka, Apache Samsa, and Apache Storm (for distributed stream processing engine) [22].

3. **Application of Big Data Analytics in Higher Education Sector**

Big data Application of Big Data in HEIs is still in its early stage. The universities have a large amount of data that could be mined using big data applications. The pressure on HEIs concerning graduating students is a global challenge [23]. However, with the techniques of big data, the analytics of large data has the potential of creating actionable value that could improve the success of students, improve the quality of teaching, and also ease the learning process for students. In addition, the application of big data analytics to HEIs also help promote awareness for parents, tax payers, and stakeholders, whom are integral components of an educational system [12], [8].

Big data can be mapped to higher institutional data as it can be applied to learner, tutors, department, and faculty depending on the aim and purpose in which it needs to be analyzed. Analytics in higher education system have various descriptions parallel to where the analytics are purposeful [13]. Big data analytics techniques can be used to analyze, manipulate, store and find meaningful information for decision making in higher institutions. The use of big data in HEIs as shown in Figure 1 and further illustrated in Table 1 is mainly categorized as academic and learning analytics [20].
Figure. 1: Schema Big data mapping in HEIs.

A. Academic Analytics
Big data analytics in academics has to do with the development of resources, improving processes and workflows of the academic institution by use of the institutional, academic and student data. [24] regards academic analytics as the combination of institutional data, arithmetical analysis, and predictive modeling to create intelligence with which students, instructors, or administrators can change the academic pattern of its institution. Big data analytics can be applied in academics as follows:

- Students admission by use of predictive modeling to allow the offices of admission and administrative units predict better and manage the size and admission pattern for new students to be enrolled.
- Decision making for the administrative officers and the stakeholders using analytical models and data mining approaches to find, understand and reoccurrences of patterns in data [12]. Also, to identify new trends of data in which significant value can be attained from the data such that the right decision can be taken according to the situation for the betterment of the institution.
- Alumni data can be analyzed by big data analytics to statistically forecast graduate employability in society. Big data analytics can be applied for fundraising by using predictive models to identify and examine the alumni information to predict those most likely to donate [12]. Big data analytics can also help students to get to know about the skills that they need that can be suitable for their dream jobs [25]. Therefore, making graduates get appropriate jobs.
- Other areas where academic analytics of big data could be applied include resource allocation, budget (finance), strategic planning, staff-centric services, Attrition patterns, industrial collaborations and linkages, and transportation management.

B. Learning Analytics
Learning analytics focuses on the success of the students. This involves the use of predictive analytics and processes to collect, analyze, use, integrate generated and actionable student data with the purpose to improve the performance of students and teaching of teachers. Learning analytics can also be used to substitute tables of data with dashboards that give an instant response about academic goals, student needs, and targets [8]. Additionally, big data is also said to have replace human decision-making by using automated algorithms [26]. The fundamental of learning analytics involves capturing and analysis of data to improve learning for students. [12] stated that the analysis process of learning analytics has five phases. The steps include (a) Capturing; data collected in real-time from various sources such as course management systems (CMS), virtual learning environments (VLE), personal learning environment, forums, web portals, chat rooms are combined with student data [27]; (b) Reporting; the collected data is used to generate precise patterns for recognizing and measuring the
student’s progress. Afterward, learning analytics dashboards use visualization for a better understanding of the data [28]; (c) Predicting: the visualized data is used for predicting the performance of student, success and also for identifying students at risk. Also, the predictive models can be used for decision making and policies about courses and resource allocation and the decision-makers of the institutions [29], [30]; (d) Acting: the information obtained from the processing and analysis of data is used to set appropriate interventions especially with regard to students at risk of failure or dropping out [31]; (e) Refining: the gathered data is used in a recurrent process for constant developments of the operated model in teaching and learning for both tutors and students [32]. Big data analytics can be applied to learning analytics as follows:

- Tracking student performance: Analytics of big data such as analytical dashboard can be applied to mine data stored in its cloud-based platform using real-time analysis to predict which students are struggling, under pressure and at risk of dropping out, understand and improve the performance of the students.
- Retention rate can be potentially improved by applying big data analytics such as adaptive learning analytics with the use of course management system data to build a predictive model that will identify and examine the students who are academically struggling. This way, early warning, and counseling can be proactively provided to such students.
- Scholarship can be awarded to students in need by using predictive analytics to create a model utilizing the student profile to identify students who are have not paid their bills or are struggling to pay.
- Student clubs and activities help the student to develop a healthy mental system. Students can learn a lot from programs available such as academic tutoring and advising, fitness educational programs, social activities and services (student organization memberships), sessions advising, career for both staff and students in the institution. All these programs can be created and managed with the application of big data analytics such as social network analysis for using data from both manual and online records [33]. Other applicability of big data analytics in learning analytics includes admission pattern, talent, and leadership.

The integrative platforms of big data analytics can use various machine language and statistical analysis to identify difficulties, threats, and opportunities. Consequently, as shown in Figure 2, education can become active, cheaper, operations can be provided, and learning strategies can be enhanced [34], which is also presented in Table 1. Therefore, the quality of education can be improved.

![Analytics in Higher Education Institutions](image_url)
Table 1. Summary of the Applicability of Big Data Analytics in Education.

| Big Data Analytics   | Descriptions                                                                                                                                                                                                 | Applicability Level                                                                 | Beneficiaries                      |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|-----------------------------------|
| Analytics            | Data evaluation and analysis that enables the assessment, improvement of tutors, programs, departments, institutions and as well as predicting and comparing the learner’s performance.                       | Can be applied to all levels of the institution for better data-driven decision making. | Learner, Tutor, Department, Institution |
| Academic Analytics   | Decision making primarily on the operation and financial of the institution (academic difficulties) by using the support of information technology.                                                              | The Institutional students’ profiles, students’ enrollment, academicians’ profiles, alumni and research profiles | Learner, Tutor, Faculty, Administrator, and Stakeholder |
| Learning Analytics   | It can also be referred to as the interpretation of an institutional data generated and collected regarding the learners to measure academic progress, predict future performance to improve the decision making of the institution. | Students’ recruitment, course management, performance, retention                  | Learner, Tutor, and Department     |

4. Successful Case Examples of Analytics Application in universities

Some higher education institutions are progressively implementing analytics especially the development of curriculum design for the student to create a centre of attention for intending students, enhance both student retention and graduation rates and gain funding from the government [5]. For examples, some of the institution already using predictive analytics in USA universities and colleges to optimize central units of the institution, such as student behaviour and activities and to support their funds with their institutional objectives. Arizona State University has implemented analytics using College Scheduler as an analytics solution into their system using Civitas Learning as the service provider [1]. Similarly, Sinclair Community College, University of Oregon have also implemented analytics using Visual Analytics solution and SAAS as the service provider [4].

Arizona State University implemented analytics to enhance academic experience students’, and as a result of this, the graduation rate has increased by 20% [33]. The university is applied analytics software known as College Scheduler. It is a software design which allows students to input individual and confidential information into a dashboard. The platform considers students’ academic and personal requirements and auto-fills the mandatory courses they must take [35]. The software is beneficial as it prevents students from taking courses that are not relevant to their majors, hence helps in managing their time and financial funds [36]. [37] ascertained that the College Scheduler software could improve the rate of college completion by more than 3%. Another university that has applied analytics is the University of Maryland – College Park. The college targeted management of students’ performance, by using analytics to predict student success or failure and forward advising and with the implementation, the college has gained by supporting limited success gaps for low-income students, shortened graduation time and thereby improved graduation rates [38].
5. Challenges of Big Data Analytics in Higher Education Sector

Big data analytics offers numerous benefits that are of great impacts in resolving critical issues of the educational system. However, accompanied with the benefits are some challenges. Some of the critical challenges of big data management are analytical and technical challenges. Analytical challenges cover the volume, storage, analysis and quality of the data. [39] state that available data in an organization may be in multiple format such as structured, semi-structured or unstructured. With the continuous increase of data, managing it would be a challenge. Similarly, [40] explained that working with heterogeneous data is time consuming and very expensive. To add, there are also issues such as data collection, management, storage, data protection, privacy, and data ownership—which may be caused by the teachers and students to question who get access to information and security of the data [38]. Another major challenge of applying big data in aspect of learning analytics is that the analytics may focus more intensely on the technological context particularly in collection and analysis of data, rather than showing what data should be collected and analyzed and for what purpose such data could be used [41].

6. Conclusion

With high demand on the education system to improve the quality of education, there is no doubt, that application of big data analytics has the potential to impact, transform and produce a better outcome especially regarding strategic planning and policy in the education system. This will enable the education system to develop new ways of attaining excellence in both teaching and learning. Therefore, enabling continual development on the activities of student data which can be aggregated and collectively supported with other educational data to give a better descriptive of the effectiveness and successively of learning and teaching at institutions. Big data analytics provides for students meaningful which helps them towards achieving academic success. Also, adapting learning analytics in institutions could help students improve performance, skills, and knowledge in a more personalized and self-paced way. Big data analytics in education can be ascertained to provide a better understanding of educators, of how effective their subjects are being understood and applied by students. Hence, the application of big data analytics in higher education system provides opportunities and supports for students, educators, faculty, deans, senior management, education authorities and ultimately the government. However, identifying or choosing the right analytics to apply, integration of those databases and the selection of appropriate data to be used to get the best insight especially in decision-making are still significant challenges in the education industry.

Acknowledgement

The authors highly appreciate Universiti Tun Hussein Onn Malaysia (UTHM) for funding this research work under the Fundamental Research Grant Scheme FRGS vot K20

References

[1] Adeleke A O, Samsudin N A, Mustapha A, Nawi N M 2018 A Group-Based Feature Selection Approach to Improve Classification of Holy Quran Verses, International Conference on Soft Computing and Data Mining 282-297
[2] Fisher D, DeLine R, Czerwinski M, Drucker S 2012 Interactions with big data analytics, interactions 19(3) 50-59
[3] Davenport T H, Patil D J 2012 Data scientist, Harvard business review 90(5) 70-76
[4] Elgendy N, Elragal A 2014 Big data analytics: a literature review paper, Industrial Conference on Data Mining 214-227
[5] Crawford K, Schultz J 2014 Big data and due process: Toward a framework to redress predictive privacy harms, BCL Rev 55(93)
[6] Bates D W, Saria S, Ohno-Machado L, Shah A, Escobar G 2014 Big data in health care: using analytics to identify and manage high-risk and high-cost patients, Health Affairs 33(7) 1123-1131
[7] Chandarana P, Vijayalakshmi M 2014 Big data analytics frameworks, International Conference on Circuits, Systems, Communication and Information Technology Applications
[8] Chaurasia S S, Rosin A F 2017 From Big Data to Big Impact: analytics for teaching and learning in higher education, *Industrial and Commercial Training*

[9] Bienkowski M, Feng M 2012 Enhancing teaching and learning through educational data mining and learning analytics: An issue, *Proceedings of conference on advanced technology for education* 1-64

[10] Daniel B 2015 Big data and analytics in higher education: Opportunities and challenges, *British journal of educational technology*, 46(5) 904-920

[11] Molina H M 2019 Big-Data Readiness of Four-Year Public and Private North Carolina Higher Education Institutions, *ProQuest LLC*

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

[13] Van Barneveld A, Arnold K E, Campbell J P 2012 Analytics in higher education: Establishing a common language, *EDUCAUSE learning initiative* 1(1) 1-11

[14] Bughin J, Chui M, Manyika J 2013 Ten IT-enabled business trends for the decade ahead, *McKinsey Quarterly*

[15] Manyika J 2011 Big data: The next frontier for innovation, competition, and productivity, [http://www.mckinsey.com/Insights/MGI/Research/](http://www.mckinsey.com/Insights/MGI/Research/)

[16] Russom P 2013 Managing big data, *TDWI Best Practices Report, TDWI Research* 1-40

[17] Adnan K, Akbar R, Wang K S 2019 Information Extraction from Multifaceted Unstructured Big Data, *International Journal of Recent Technology and Engineering (IJRTE)* 8 1398-1404.

[18] Manohar A, Gupta P, Priyanka V, Uddin M F 2016 Utilizing big data analytics to improve education, *ASEE*

[19] Gantz J, Reinsel D 2011 Extracting value from chaos, *IDC iview* 1142 1-12

[20] Patel M R, Desai T 2016 Big Data Analytics in Optimizing the Quality of Education: Challenges, *International Journal for Innovative Research in Science & Technology* 3(6) 165-167

[21] Khan N, Yaqoob I, Hashem I A T, Inayat Z, Ali M, Kamaleldin W, Gani A 2014 Big data: survey, technologies, opportunities, and challenges, *The scientific world journal* 2014

[22] Halaweh M, Massry A E 2015 Conceptual model for successful implementation of big data in organizations, *Journal of International Technology and Information Management* 24(2) 2

[23] Neubaum D O, Pagell M, Drexler Jr J A, Mckee-Ryan F M, Larson E 2009 Business education and its relationship to student personal moral philosophies and attitudes toward profits: An empirical response to critics, *Academy of Management Learning & Education* 8(1) 9-24

[24] Baepler P, Murdoch C J 2010 Academic analytics and data mining in higher education, *International Journal for the Scholarship of Teaching and Learning* 4(2) 1-9

[25] 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

[26] Jagadish A C, Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M 2014 Big data and technical challenges, *57*(7) 86-94

[27] Tseng S F, Tsao Y W, Yu L C, Chan C L, Lai K R 2016 Who will pass? Analyzing learner behaviors in MOOCs, *Research and Practice in Technology Enhanced Learning* 11(1) 8

[28] Ruipérez-Valiente J A, Muñoz-Merino P J, Leony D, Kloos C D 2015 ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform, *Computers in Human Behavior* 47 139-148

[29] Ain N, Kaur K, Waheed M 2016 The influence of learning value on management system use: An extension of UTAUT2, *Information Development* 32(5) 1306-1321

[30] Akhtar S, Warburton S, Xu W 2017 The use of an online learning and teaching system for monitoring computer aided design student participation and predicting student success, *International Journal of Technology and Design Education* 27(2) 251-270

[31] De Freitas S, Gibson D, Du Plessis C, Halloran P, Williams E, Ambrose M, Arnab S 2015 Foundations of dynamic learning analytics: Using university student data to increase retention, *British journal of educational technology* 46(6) 1175-1188
[32] Nam S, Lonn S, Brown T, Davis C S, Koch D 2014 Customized course advising: investigating engineering student success with incoming profiles and patterns of concurrent course enrollment, *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* 16-25

[33] Reid-Martinez K, Michael M 2015 Big data in education: Harnessing Data for Better Education Outcome, *3* 33

[34] Ward H, Brown R, Hyde-Dryden G. *Assessing Parental Capacity to Change when Children are on the Edge of Care: an overview of current research evidence*. London, United Kingdom: Department for Education 2014

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

[36] Siemens G 2012 Learning analytics: envisioning a research discipline and a domain of practice, *Proceedings of the 2nd international conference on learning analytics and knowledge* 4-8

[37] De Mauro A, Greco M, Grimaldi M 2016 A formal definition of Big Data based on its essential features, *Library Review* 65(3) 122-135

[38] Meenakumari J, Kudari J M 2015 Learning Analytics and its challenges in Education Sector a Survey, *Int. J. Comput. Appl* 0975-8887.

[39] Jagadish H V, Gehrke J, Labrinidis A, Papakonstantinou Y, Patel J M, Ramakrishnan R, Shahabi C 2014 Big data and its technical challenges, *Communications of the ACM* 57(7) 86-94

[40] Labrinidis A, Jagadish H V 2012 Challenges and opportunities with big data, *Proceedings of VLDB Endowment* 5(12) 2032-2033

[41] Sclater N, Peasgood A, Mullan J 2016 Learning analytics in higher education, *London: Jisc. 8*(2017) 176