A Proposed Framework for Big Data Analytics in Higher Education

Beenu Mago[1], Nasreen Khan[2]

Assistant Professor, School of Information Technology, Skyline University College, Sharjah, U.A.E[1]
Assistant professor, SZABIST University, Dubai, UAE[2]

Abstract—**Students, faculties, and other members of the higher education (HEd) system are increasingly reliant on various information technologies. Such a reliance results in a plethora of data that can be explored to obtain relevant statistics or insights. Another reason to explore the data is to acquire valuable insight regarding the novel unstructured forms of data that are discovered and often found to have a connection with elements of social media such as pictures, videos, Web pages, audio files, etc. Moreover, the data can bring additional valuable benefits when processed in the context of HEd. When used strategically, Big Data (BD) provides educational institutions with the chance to improve the quality of education from all the perspectives and steer students of HEd toward higher rates of completion. Further, this will improve student persistence and results, all of which are facilitated by technology. With this aim, the current research proposes a framework that analyzes the data collected from heterogeneous sources and analyzes using BD analytics tools to do various types of analysis that will be beneficial for different learners, faculties and other members of HEd system. Moreover, current research also focuses on the challenges of acquiring BD from various sources.**

Keywords—Big data analysis; higher education; learning analytics; academic analytics

I. INTRODUCTION

Many people in the twenty-first century are interested in pursuing careers as data scientists [1]. Large amounts of data are being collected and analyzed, which is the source of the inspiration for this newly growing demand. BD analytics is the act of analyzing large amounts of data to uncover hidden patterns using computational algorithms, programming, and statistical modeling tools in order to discover significant and timely correlations [2, 3]. This enables the extraction of actionable information that may be used to impact internal company decision-making. As well as this, it assists in the discovery of patterns and the detection of anomalies across a wide range of data sources. Education is not an exception to the growing demand for BD in a variety of industries [4], which is increasing all the time. Certain parties, on the other hand, are still unwilling to engage with BD as a result of the application's specific constraints. Example: BD sets are frequently too enormous to be collected, stored, and analyzed correctly using typical database approaches [5, 2], and there is frequently no consistent structure for the data that has been gathered. According to Reference [6], consumers have evolved into a "continuous initiator of both structured, transactional data as well as contemporary unstructured behavioral data." It is possible to quickly overcome the

difficulties to using BD by integrating novel technologies with freshly created approaches for data analysis [7]. Therefore, organizations that employ dynamic platforms for BD may profit through the enhanced corporate skills of critical thinking and taking appropriate decisions, which contributes to the broad usage of BD analytics across industries [8]. However, only a small number of papers and industrial accounts explicitly focus on the use of BD in tandem with HEd [2,10,11], despite the fact that several researchers have looked into the applicability, data availability, affordability, competence, privacy, relevance, and ownership of BD in many domains [9,10].

BD functions as a previously undiscovered opportunity for HEd institutions. The basic structural organization of the campus, the networks in use, servers, relevant apps, learning management systems (LMS), and other end-user systems generate data in huge volumes which, upon proper evaluation, may reveal substantial information and a new understanding concerning the various challenges that govern HEd [13, 14, 15]. As such, implementing BD in a HEd environment may result in enhanced skills involving the process of making decisions. It is the goal of this study to further add to the existing research on BD by analyzing the application of BD analytics in HEd institutions.

This paper is structured into different sections. First, the authors review the role of BD in HEd followed by importance of data analytics in HEd. Authors provide deep insight about the educational analytics. Post the review of literature, authors propose the conceptual framework followed by conclusion, future research and limitations of the current study.

II. BD AND EDUCATION

Graduating students' learning processes are frequently chastised for failing to adequately prepare different stakeholders by not providing them the knowledge, relevant tools and mindsets, and values which are essential in future, yet, the huge data produced by the infrastructure various systems of today's digital campus that includes data produced or collected from the network applications of campus, various other internal applications used for serving various needs of HEd. It is observed that HEd rarely use Learning Management System (LMS), mobile or other devices, which are used by different stakeholders of HEd, to gain important details about the issues that are affecting or can affect current educational institutions [16]. This proves that the use of promising techniques such as BD is still in its infancy in HEd. Owing to various educational concerns, there has been a rise in the

number of useful suggestions and propositions offered to ensure that HEd institutions can flourish in every way possible which can be summed up under the category of increased transparency and accountability involving all stakeholders including students, parents, taxpayers, and others [5] It is important to understand that using BD or LA in HEd is the best method of not only accumulating but also evaluating data regarding the advancement of the students and the educational environment. It not only offers educational content information to instructors, but identifies activities that improve teaching and evaluation procedures [2]. Much of the motivation for employing analytics in HEd stems from more stressed university finances and increased cost awareness. Students are also becoming more demanding as stakeholders, asking proof that their tuition is being spent wisely.

The advanced new integrative platforms have the ability to reduce the cost of education, improve operations, and frame more refined learning techniques [21]. This is possible because these platforms can access and interpret data that is collected from diversified sources that embraces log files, firewall data, remote sensors, legacy applications, structured databases, LMS, web servers, networks, student registration systems, and mobile and online learning apps, in addition to detecting problems, risks, and opportunities using untapped machine data [2]. A more convincing and engaging educational experience can be achieved by combining the digital and physical worlds. AA can help decision-makers make better decisions by focusing on trends that signal student performance and academic brilliance.

The following part will review the three primary research cores areas of data analytics in HEd. These First core area is Learning Analytics (LA). The second area is Academic Analytics (AA). Next, researcher has discussed Educational Data Mining used as EDM in short in this research. Next, it will present an overview of each research stream after which an integrated picture of the entire process will be presented.

III. DATA ANALYTICS IN HED

Educational analytics is the deep and detailed analysis of large amounts of educational data to gain valuable insights. Teaching Analytics, LA, AA, and EDM are some of the academic areas through which the power of analytics is utilized for this process [16].

A. Learning Analytics (LA)

Although a variety of terminology and concepts have been used to define the developing study subject of LA, it is best understood as the application of data diagnostic in the context of learning and teaching. Unlike AA and EDM, this technique is heavily focused on the learners and their learning process, gathering, integrating, and evaluating static and dynamic data on the learners' profiles, educational resources, and learning context. It hopes to achieve this by providing descriptive modeling and prediction of learning elements in real-time.

"The measurement, collection, analysis, and reporting of data about learners and their contexts for the purposes of understanding and optimizing learning and the environments in which it occurs," is what LA is according to the Society for LA Research. This interpretation has recently gained widespread acceptance in the scholarly community [11,18].

B. Content Analysis

Many people are interested in analyzing hypertexts discovered on websites because of the ongoing advancements in the techniques related to web analytics such as machine learning, algorithms, web crawling, etc. [10].

The contextualized interpretations of textual documents are dealt by a branch of LA which is known as Content Analysis [19]. This method analyzes texts to understand their underlying meanings, either manually or with the assistance of technology. Despite the availability of various textual educational sources, this technique divides the texts into five groups for analysis purposes: written, oral, iconic, audiovisual, and hypertexts.

C. Discourse Analytics

To extract the relevant information related to the features of the used languages from the learning outline. It includes the study of user engagement. From the empirical data, the Content Analysis reveals the meaningful data while Discourse Analytics mainly concentrates on the learners' language. Through the platforms where the employees communicate like, online learning pages, it addresses to the learners. The interaction will contribute the procedure of the knowledge production [20, 22]. Along with the platforms that connect the learners to the content items, the learning can be considered as socio-constructivist procedure [23]. The analytics of the discourse data and metadata is now possible because of log follow up and data mining. On the other hand, the Discourse Analytics focuses on the data and language that has been employed in learning process while the Social LA concentrate more on the involvement of the learners.

D. Social LA

Social LA is a unique branch of LA related to the cooperation and coordination between the learners [24]. In education, the Social LA studies a learning procedure from the social point of view and states that attaining newer information and skills are not only a person's achievement while the Discourse Analytics examine the learners' language. The way social interactions and platforms influence learning routine has been explained by many prior researches [25]. Like social network analysis (SNA) was carried out by [24] in order to determine the link between educational performance and contacts through social platforms. The outcomes of this research will help us to know whether we can depend on social network factors to act as indicators of student progress but on the other hand, for the calculations, this study discourages to entirely depend on social groups aspects. Also, for the social learning analytics, the data visualization is effective according to this study.

E. Disposition Analysis

The theoretical information about the learners' back ground and their involvement in learning is discussed by Disposition Analytics which will help to determine the character of students along with their connection with the education procedure [26]. That means, the aspects, brought by

the learners in the learning process to recognize the learning behaviors and anticipate the better learning attitudes for the better performance of both educating and learning, are studied by the LA technique. This is evident in [26], where the disposition analytics were used to study the influence of different features on the learner's progress and the teacher will be able to take informed decisions, regarding specific option and optimal instructing methods and approaches, by understanding this dynamic.

F. Academic Analytics (AA)

AA is precisely defined corporate intelligence in education and particularly, it refers to the procedure of determining the insightful patterns in educational data in order to maintain calculated management and to predict scholastic issues like drop out ratio [11]. The method mainly talks in the favor of organizational executive bodies and educationists but it has been assumed by the students that data analytics are used anticipate and laud their educational progress. In order to supervise and improve the academic the Key Performance Indicators (KPIs), like student retention, the organization's management will use the AA. It has been referred to, as a development for the provision of advanced educational institutes with the info that is required to support the functional an economic management [17].

Through the disclosing the important factors contributing to the learners' achievement, helping to understand the useful methods and increasing the scholarship of learning and educating, the AA support the staff. In the higher schooling, the student's accomplishment is considered to be among the primary key performance indicators (KPIs) so; majority of staff is interested in supervising and anticipating student success. Also, relevant and important information is collected from the academic data by AA in order to allow the teaching staff to make educational alterations to fulfill students' necessities and to examine the most useful methods.

To back their management, the executive bodies may extract required information from the AA. A unique set of KPIs is offered by AA which traditional educational systems do not possess them. For instance, the vice chancellor may force for the supervision of the learning and teaching methods upon knowing the ratio of at-risk students. The AA are also used maximize the employment of resources by the administrative team. The AA may play a part in development of the organization's answerability and improve the institute's image [30]. Regardless of many advantages, managements often highlight the expenses incurred in the AA plan [28,29], when the system is functioning, they might be worried about the security threat and privacy.

IV. EDUCATION DATA MINING (EDM)

The EDM comes with the target of solving academic crisis and exploring the hidden info, talks about the usage of the data mining methods on scholastic data [29]. This information can extract learners' demographic information and their attitude towards the educational activities like, quizzes, interactive class exercises/tasks in a learning space and use data from team of learners who work together in a project, discussion platform, teacher data, managerial information, demographic

and cognitive information. To attain the data regarding education of student, the EDM can be used which will help to improve the educational processes and monitor the leaning for the feedback. The information can be employed to give suggestions to adjust with the leaning attitude of the students depending on their learning patterns along with the unexpected leaning attitudes [31]. Like, the way of how the learning tasks of students were anticipated with the usage of k-means clustering algorithm according to Reference [30]. A machine-algorithm was used by [33] to know about the undergrads at engineering colleges who dropped out in their freshmen year. To conduct, student identification and to classify them according to their academic assessments that will comprise of quiz scores, assessment grades, exam scores and practical test scores, [34] used various data mining algorithms.

There are various kinds of EDM. All the types of academic data mining could be classified into five categories: prediction, clustering, relationship mining, discovery with models, and distillation of data for human judgment according to [35]. The initial three groups are generally considered by the EDM research team while, the last two groups are just dominant in the area of academic data mining [31]. The term prediction, in the EDM targets to show case the educational finding from the other information factors. Input factors are referred as predictor variable whereas the forecast factor is mentioned as predicted variable. The clustering is the second form of academic data mining that targets on grouping the unprocessed information into sets of clusters and identifying the borders in the groups. As said by [35], this category of academic info mining techniques can take predefined theories or no following studies into account.

The relationship mining is the third category of the academic data mining that aims to examine the existent connections in the datasets with different variables. It is divided in to four sub groups that are: association rule mining, causal data mining, correlation mining, and sequential pattern mining [35, 36]. The association rule mining will reveal the if-then rulings in the variables. To be specific, data mining techniques determine the links where any set of variables is described and other variable will have certain value. For example, [33] used shared the association rule mining as it was recommended to propose the guidelines like, the assignment score might be increased if the responded messages number in the discussion is high.

Approaches of Causal Data Mining are utilized to discover "casual relationships" in which one occurrence causes another. Both unidirectional and bidirectional causal relationships are possible [27]. Relationship Mining and Correlation Mining are remaining two types. They are looking for positive or negative linear correlations between variables whereas Sequential Pattern Mining is looking for temporal relationships. Reference [37], employed Sequential Pattern mining to integrate metacognitive judgments and eye movements to explore processes underpinning multimedia learning. Discovery with Models is deemed as the fourth sort of EDM; it builds a model of a phenomenon using existing EDM approaches or knowledge engineering and then using it as a component in another research. Discovery with Models,

according to [35], finding with frameworks usually implies the substantiated adaptation of a forecasting model throughout several environments. The principal application of this EDM class is the finding of relationships among student conduct and subjective variables in the teaching environment [32].

V. BD ANALYTICS IN HED

As a result of the massive amount of data being generated on a daily basis, contemporary system of integrated innovative technology demands has reached an all-time high. As a result, BD has emerged. There have been numerous studies conducted on the use of BD in a variety of sectors, with particular emphasis on the availability of data, cost of acquiring data, application, and importance of BD [8]. Despite this, only a few papers [9] [10] have documented the integrative application of BD in HEd. HEd institutions have access to an enormous amount of data that is stored in a variety of different sources and is managed by a variety of different processes. In this way, BD has the ability to harness institutional data that can improve the future of the education sector, particularly in terms of decision-making, and so improve the future of society. The evaluation of a diverse variety of organizational and functional information collected via the methods conducted to assess organizational efficiency and improvement in predicting long term growth and manage possible problems in academic programming, research, teaching, and learning is referred to as BD in HEd [13,20]. Many stated that HEd should employ insights as a mechanism to meet the demands of higher efficiency. A proportion of researchers have claimed that the BD paradigm is well suited to dealing with some of the core challenges that HEd faces as a new discipline within academia.

At this moment, much of the analytics work in HEd is integrative, involving the fields of Educational Technology, Statistics, Mathematics, Computer Science, and Information Science. The present work on analytics in education is centered on data mining.

In HEd, database systems that store vast volumes of longitudinal data about students, down to extremely specific transactions and activities relevant to learning and teaching, are referred to as BD [12]. Students leave data trails when they interact with learning tools, which can expose their attitudes, social connections, intentions, and aspirations. This data can be used by researchers to examine trends in student performance over time, such as from one semester to the next or from one year to the next.

On a higher level, the ability to find meaningful data and turn it into useable information by identifying patterns and deviations from patterns might be said to constitute BD's added value. According to [4], BD is now in a good position to begin tackling some of the major difficulties that HEd is currently experiencing. It could serve as the foundation for HEd to rethink both its financial model and the evidence used to make judgments about educational outcomes [2].

From the viewpoint of organizational learning, the analysis of appropriate data is well recognized, and today's technologies enable institutions to obtain insights from data at previously unimaginable levels of sophistication, speed, and accuracy [3, 7]. The production of essential information by students, computer programs, and systems is increasing as technology continues to permeate all sectors of HE [1].

BD Analytics may also be used to evaluate student replies to a course exam, discussion board postings, blog entries, and wiki activity, all of which can result in hundreds of transactions per student each course. These data would be collected in real-time or near-real-time as transactions occur, then evaluated to suggest actions. "[Learning] analytics are a foundational tool for informed change in education," according to [38], and they provide evidence on which to create knowledge and make informed (rather than impulsive) decisions.

With the widespread use of information and communication technology, the amount of data that may be transmitted on the smart education platform grows exponentially [39]. When data is scattered over numerous unconnected diverse data systems in institutions, iLMS works closely with the learners serving many learners at a time and tailored according to the individual needs [40]. With innovation in technology, the data becomes so vast and takes the form of BD. The data is definitely unstructured and semi-structured, petabytes are produced on a daily basis, and traditional relational databases cannot effectively scale to support real-time analytics based on the data [41]. Further, it is also observed that analysis of this BD can help with the issues of obtaining information at the correct time. BD can help increase decision-making capability by exploring ways to aggregate data across platforms.

Despite growing interest in investigating and unlocking the possibilities of growing data in the HEd environment, there is a lack of research on BD in HEd.

- This work adds to HEd's theoretical and conceptual understanding of BD and Analytics.

- It defines BD and discusses how it relates to HEd.

- It presents a conceptual framework based on this expanding field of study.

VI. PROPOSED FRAMEWORK, FUTURE RESEARCH AND LIMITATIONS

With the falling cost of BD storage, open source software such as Apache Hadoop and NoSQL databases, network bandwidth, and on-demand access to resources provided by cloud computing, complicated technologies are becoming more accessible to a broader range of people. To emphasize this, it should be noted that, as the cost of BD and Analytic tools continues to decline, they are becoming significantly easier to use [30]. This, in turn, opens the door to the use of these tools by businesses and educational institutions to achieve better outcomes while also making more efficient use of available resources. BD in HEd is being driven by a number of causes, including stakeholder expectations of good practices and accountability, as well as an increased desire for evidence-based guidelines to assist in decision-making. This means that the first stage in establishing a method to achieve this goal is to identify acceptable technology platforms that will enable BD analytics in addition to relevant technological

talents, such as data scientists. Due to the fact that data is no longer viewed as just a back-office accounts-settling tool, but rather as a real-time decision-making tool that can be used by data scientists to extract useful information from terabytes of data that would otherwise remain hidden [16], data is becoming increasingly valuable. By utilizing virtual and online learning systems, information technology is transforming education and broadening the scope of the educational environment beyond traditional lecture halls and classrooms to accommodate learners who require flexibility in terms of both time and distance [6] [4]. The educational system is progressively being pressed to respond to economic, political, and social developments, such as the need to expand the number of students enrolled in specific subjects and the need to teach graduates with the skills and traits demanded by industry, among other things. These schools must use information technologies such as BD, which have significant potential in HEd, if they are to be responsive to these shifts in the marketplace. The majority of educational data has traditionally been collected through traditional assessments, but it is now increasingly being collected through online educational systems, educational games, and simulations, resulting in a large amount of data, which in turn opens the door to BD analytics opportunities. In the context of today's technology-enabled learning, the student is at the heart of a network of information-rich relationships in the educational system, which is offering opportunities for BD analytics, according to [17]. The author in [14] argues that HEd institutions are functioning in increasingly complicated situations, and that, as a result, it is essential that they employ BD to provide the finest learning environments for the benefit of the entire society. There are opportunities to use BD to link research and education, by making effective use of the most recent research techniques and outcomes to guide teaching and to enable the conduct of research activities as part of educational activities, for example. Students generate massive volumes of potentially valuable information during the course of their learning processes in areas such as course selection and registration, financial information, class participation, online resource usage, and textbook purchases, all of which can be used as raw material for BD analysis and redeployment to assist administrators and students in making better decisions. As a result of the digital revolution, HEd institutions are gathering more data than they have ever collected before.

Many colleges and universities have proved that analytics may help an institution considerably advance in such important areas as resource allocation, student achievement, and financial planning and administration. When leaders of institutions of HEd read about innovations occurring at other institutions as a result of BD analytics, they begin looking for methods to implement analytic programs in their own institutions [31]. Several studies, including the McKinsey Global Institute Report on BD [33], have found that the most developed regions, such as Europe, have the most potential to create value through the use of BD. BD adoption and widespread use in other regions, particularly in poor countries, has been bolstered as a result of this development. In accordance with [32], the diffusion of BD gives an account of how BD goes from discovery to widespread use, and how this

is aided by steps taken by service providers of important technologies necessary to enhance the resources and capacities of academic institutions. [34] Stated that countries may take advantage of the numerous BD opportunities that are accessible to them in order to gain value from the huge volumes of data that are generated and, in the long term, aid in their development. BD and analytics in HEd have the potential to be transformative, affecting existing processes of administration, teaching, and learning, as well as contributing to policy outcomes by assisting in the resolution of existing difficulties facing educational institutions [4, 35]. The following are some examples of how successful institutions have made use of BD. First and foremost, we must cultivate a culture of completion and outplacement. 2) Decrease in the number of nonproductive credits. 3) Rethinking the way in which teaching is delivered. 4) Redesigning essential support services such as human resources, academic services, and finance in order to generate data that is valuable for strategic decision-making. 5) Improving the efficiency of non-core services and operations [17].

BD may have an impact on HEd practice in a multitude of ways, including improving student experience, academic programming, scientific proof judgement, and tactical responses to changing economic trends. BD has the ability to turn complex, often unstructured data into actionable information. Reference [14] claims that BD is a cost-effective approach to improve judgement. In order to organize the available literature and build a study technique to aid in the establishment of a set of choices for investigation, reference [2] presented a conceptual framework to describe BD in HEd along four components (Learning Analytics, Institutional Analytics, and AA and information Technology analytics). Following Fig. 1 shows the diagrammatic representation of it.

There is a wide range of administrative and operational data gathering processes that are used to assess institutional performance and progress in order to predict future performance. Reference [42] describe a wide range of administrative and operational data gathering processes that are used to identify potential issues related to academic programming, research, teaching, and learning. Data is rising in HEd institutions, but the majority of it is distributed among computers and departments, and it comes in a variety of formats, making it impossible to retrieve or consolidate the information. Having the ability to analyze diverse information sets, regardless of where they come from, is required for effective use of these data. Consolidating data stored in silos within institutions, managing and governing the data while protecting sensitive information across databases is a key requirement for implementation of BD in HEd [42]. A better basis for making better decisions about essential business and technological demands is provided by consolidated data from many sources across an institution, minimizing redundancy and the loss of valuable time gathering data from multiple sources. Though it can be incredibly challenging to combine data sets from across a number of disconnected systems, doing so results in more complete insights that inevitably lead to enhanced predictive modeling skills. Based on these existing deficiencies, the researcher of the current study develops a framework for utilizing large amounts of data from a variety

of sources. Following Fig. 2 shows the proposed framework for the current study.
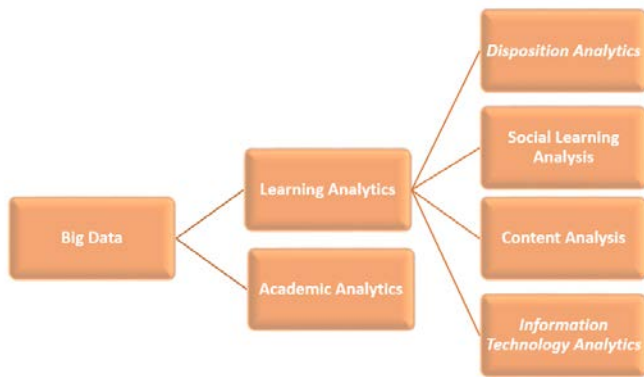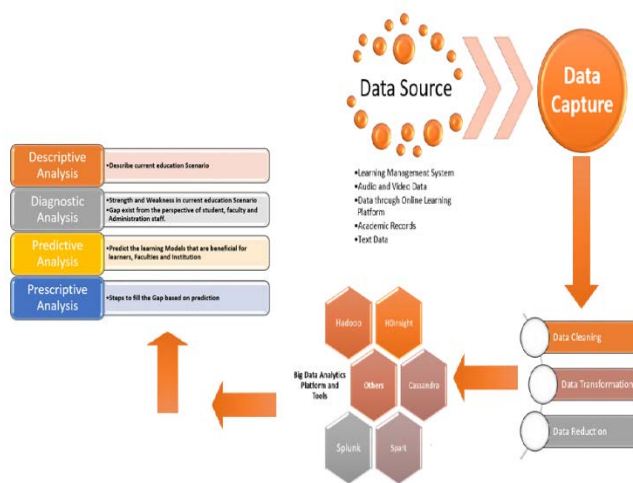

Fig. 1.   Components of BD.

Fig. 2.   Proposed Framework for Analyzing BD in HEd.

Based on the above framework, the BD which can be collected from various heterogeneous sources and can provide benefits which can be summarized as shown in Fig. 3.
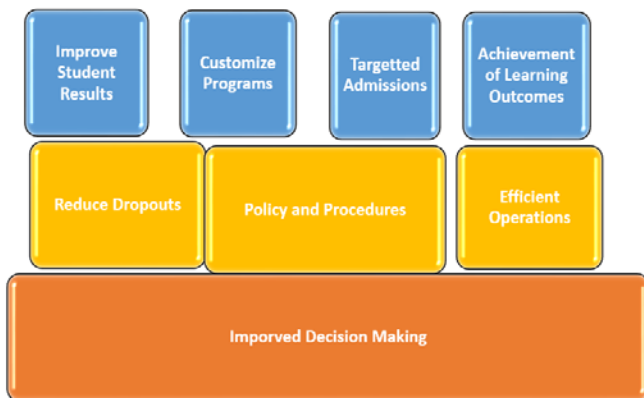

Fig. 3.   Benefits of Analysis of BD in HEd.

VII. CONCLUSION, FUTURE RESEARCH AND LIMITAIONS

As a result of the explosion of data, BD has evolved as a means of processing huge amounts of data from a number of sources. Universities with strong research and development capabilities can reap significant benefits from this technology. HEd institutions must consider the long-term implications of implementing BD analytics, which has the potential to broaden the scope of their teaching-learning orientation. Among the recommendations made by executives are that administrators and educators assist the system and gain value from practical applications, such as creating a culture of data use for educational decision-making, being keen on data users by asking critical questions about market deals and suggesting the most beneficial uses and features, involving IT departments in data collection, application planning, and starting with a pilot program, among others. As a result, institutions can save money while also making attempts to incorporate pedagogical abilities and make more informed decisions, which is beneficial. As a thoughtful contribution to curriculum design for instructors and learners in education in particular, and as a sustained competitive advantage for the education sector in general, this framework is intended to facilitate the development of innovative teaching-learning orientations for the performance of educational institutions, as well as for the performance of educational institutions as a whole.

Furthermore, the integration of BD analytics is not an easy process. The lack of skills and expertise in the HEd sector of will pose a challenge in data storage, collection and management of data [4]. Data privacy will be another concern along with the potential ethical concerns such as purpose of data collection and analysis. Hence, it is suggested that the key stakeholders of HEd sector should be involved to gain acceptance.

REFERENCES

[1] J. Gagliardi, A. Parnell, and J. Carpenter-Hubin, "The Analytics Revolution in Higher Education," Change: The Magazine of Higher Learning, vol. 50, no. 2, pp. 22–29, 2018.

[2] B. Daniel, "Big Data and analytics in higher education: Opportunities and challenges", vol. 46, no. 5, pp. 904–20. 2014.

[3] R. Kitchin, The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. London: Sage.2014.

[4] M. I. Baig, "Big Data in Education: A State of the Art, Limitations, and Future Research Directions." International Journal of Educational Technology in Higher Education, vol. 17, no. 1, 2020. Crossref, doi:10.1186/s41239-020-00223-0.

[5] M. Attaran, J. Stark, and D. Stotler, „Opportunities and challenges for big data analytics in US higher education. Industry and Higher Education", vol. 32, no.3, pp 169–182,2018. https://doi.org/10.1177/0950422218770937.

[6] S. Erevelles. "Big Data Consumer Analytics and the Transformation of Marketing." Journal of Business Research, vol. 69, no. 2, pp. 897–904,2016. Crossref, doi: 10.1016/j.jbusres.2015.07.001.

[7] H. Davenport, Thomas. "How Strategists Use 'Big Data' to Support Internal Business Decisions, Discovery and Production." Strategy & Leadership, vol. 42, no. 4, pp. 45–50,2014. Crossref, doi:10.1108/sl-05-2014-0034.

[8] D. Trabucchi, and B. Tommaso, "Data-Driven Innovation: Switching the Perspective on Big Data." European Journal of Innovation Management, vol. 22, no. 1, pp. 23–40,2019. Crossref, doi:10.1108/ejim-01-2018-0017.

[9] S. Zillner et al. "Big Data-Driven Innovation in Industrial Sectors". In: Cavanillas J., Curry E., Wahlster W. (eds) New Horizons for a Data-Driven Economy. 2016, Springer, Cham. https://doi.org/10.1007/978-3-319-21569-3_9.

[10] N. Nistor, and A. Hernández-Garcíac, "What Types of Data are used in Learning Analytics? An Overview of Six Cases". Computers in Human Behavior, vol.89, pp 335–338,2018.

[11] S. S. Chaurasia, D. Kodwani, H. Lachhwani, and M.A. Ketkar, "Big Data Academic and Learning Analytics". International Journal of Educational Management, vol.32, no.6, pp 1099–1117,2018.

[12] S. Parack, Z. Zahid, F. Merchant, Application of data mining in educational databases for predicting academic trends and patterns. In Proceedings of the 2012 IEEE International Conference on Technology Enhanced Education (ICTEE), Kerala, India, 3–5 January 2012; pp. 1–4.

[13] A. Nguyen, L. Gardner, and D. Sheridan, A Multilayered Taxonomy of Learning Analytics Applications. In Proceedings of the Pacific Asia Conference on Information Systems, Article 54.2017.

[14] M. Hilbert, "Big Data for Development: A Review of Promises and Challenges." Development Policy Review, vol. 34, no. 1, pp. 135–74,2015. Crossref, doi:10.1111/dpr.12142.

[15] U. Sivarajah, et al. "Critical Analysis of Big Data Challenges and Analytical Methods." Journal of Business Research, vol. 70, pp. 263–86, 2017. Crossref, doi: 10.1016/j.jbusres.2016.08.001.

[16] N. Bousbia, I. Belamri, "Which contribution does EDM provide to computer-based learning environments?" In Educational Data Mining; Springer: Berlin/Heidelberg, Germany, pp. 3–28,2014.

[17] Black, Paul, and Dylan William. "Classroom Assessment and Pedagogy." Assessment in Education: Principles, Policy & Practice, vol. 25, no. 6, 2018, pp. 551–75. Crossref, doi:10.1080/0969594x.2018.1441807.

[18] A. Nguyen, L.A. Gardner, and D. Sheridan, "A Framework for Applying Learning Analytics in Serious Games for People with Intellectual Disabilities". British Journal of Educational Technology, vol.49, no.4, pp 673–689,2018a.

[19] D. Clow, "An overview of learning analytics". Teaching in Higher Education, vol.18, no.6, pp. 683–695,2013.

[20] J.A. Howell, L.D, Roberts and V.O. Mancini," Learning Analytics Messages: Impact of Grade, Sender, Comparative Information and Message Style on Student Affect and Academic Resilience." Computers in Human Behavior, vol.89, pp 8–15,2018.

[21] S. Ayesha, T. Mustafa, A.R. Sattar, M.I. Khan, "Data mining model for higher education system." Eur. J. Sci. Res. vol.43, 24–29,2010.

[22] C. Kent, E. Laslo, and S. Rafaeli, "Interactivity in Online Discussions and Learning Outcomes. "Computers & Education, vol.97, pp 116–128,2016.

[23] D.T. Hickey, "Motivation and Contemporary Socio Constructivist Instructional Perspectives". Educational Psychologist, vol.32, no.3, pp. 175–193,1997.

[24] P. Vlachopoulos, S. K. Jan, and L. Lockyer. "A Comparative Study on the Traditional and Intensive Delivery of an Online Course: Design and Facilitation Recommendations". Research in Learning Technology, vol. 27, Mar. 2019, doi:10.25304/rlt. v27.2196.

[25] E. Vrieling, "Facilitating Social Learning in Teacher Education: A Case Study." Studies in Continuing Education, vol. 41, no. 1, pp. 76–93,2018. Crossref, doi:10.1080/0158037x.2018.1466779.

[26] S. Bharara, S. Sabitha, and A. Bansal, "Application of learning analytics using clustering data Mining for Students' disposition analysis". Education and Information Technologies, vol.23, no.2, pp. 957—984, 2018.

[27] P. Baepler, and C. Murdoch. "Academic Analytics and Data Mining in Higher Education." International Journal for the Scholarship of Teaching and Learning, vol. 4, no. 2, 2010. Crossref, doi:10.20429/ijsotl.2010.040217.

[28] D. Raju and S. Randall, "Comparing Data Mining Models in Academic Analytics." International Journal of Knowledge-Based Organizations, vol. 6, no. 2, pp. 38–54,2016. Crossref, doi:10.4018/ijkbo.2016040103.

[29] P. Mishra, et al. "Educational Data Mining and Learning Analytics in Higher Education." MERI-Journal of Management & IT, vol. 12, no. 1, p. 85,2018. Crossref, doi:10.25089/meri/2018/v12/i1/180120.

[30] S. Pal, "Mining educational data to reduce dropout rates of engineering students" Int. J. Inf. Eng. Electron. Bus. vol.4, no.1. 2012,

[31] F. Alshareef, et al. "Educational Data Mining Applications and Techniques." International Journal of Advanced Computer Science and Applications, vol. 11, no. 4, 2020. Crossref, doi:10.14569/ijacsa.2020.0110494.

[32] G. Ramaswami, et al. "Using Educational Data Mining Techniques to Increase the Prediction Accuracy of Student Academic Performance." Information and Learning Sciences, vol. 120, no. 7/8, pp. 451–67,2019 Crossref, doi:10.1108/ils-03-2019-0017.

[33] Y. Salal, S. Abdullaev, and M. Kumar, "Educational data mining: Student performance prediction in academic," IJ of Engineering and Advanced Tech, vol. 8, no. 4C, pp. 54–59, 2019.

[34] K. Agarwal, E. Maheshwari, C. Roy, M. Pandey, and S. S. Rautray, "Analyzing student performance in engineering placement using data mining," in Proceedings of International Conference on Computational Intelligence and Data Engineering, pp. 171–181, Springer, 2019.

[35] Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3-17.

[36] A. Peña-Ayala, Alejandro. "Educational Data Mining: A Survey and a Data Mining-Based Analysis of Recent Works." Expert Systems with Applications, vol. 41, no. 4, pp. 1432–62, 2014.Crossref, doi: 10.1016/j.eswa.2013.08.042.

[37] R. Azevedo, N.V. Mudrick, M. Taub, and A.E. Bradbury, Self-regulation in computer-assisted learning systems. In J. Dunlosky & K. A. Rawson (Eds.), The Cambridge handbook of cognition and education (pp. 587–618). Cambridge University, 2019.

[38] G. Siemens, "The Journal of Learning Analytics: Supporting and Promoting Learning Analytics Research." Journal of Learning Analytics, vol. 1, no. 1, 2014, pp. 3–5. Crossref, doi:10.18608/jla.2014.11.2.

[39] Ateeq, K., Mago, B. & Pradhan, M.R. "A novel flexible data analytics model for leveraging the efficiency of smart education." Soft Computing (2021). https://doi.org/10.1007/s00500-021-05925-9.

[40] Alkhatib, O., & Mago, B. (2017) "ilearning Management System: An Innovative Approach to E-Learning.", International Journal of Information Technology & Management Information System, 8(2).

[41] Mago B., "Analyzing Unstructured Data Using Mining Techniques", Proceedings of the The 7th International Research Symposium of the SGBED: Managing Business in a Digital Age: Opportunities and Challenges, 17 December–19 December, 2018, Skyline University College, UAE, pp. 241–251.2018.

[42] B. K. Daniel, and R. Butson, Technology enhanced analytics (TEA) in higher education, Proceedings of the International Conference on Educational Technologies, 29 November–1 December, 2013, Kuala Lumpur, Malaysia, pp. 89–96.2013.