

## A Review of Big Data Analytics in Sector of Higher Education

Vatsala<sup>\*</sup>, Rutuja Jadhav<sup>\*\*</sup>, Sathyaraj R<sup>\*\*\*</sup>

<sup>\*</sup>(School of Computer Science and Engineering, VIT University, Vellore-632014, India

<sup>\*\*</sup>(School of Computer Science and Engineering, VIT University, Vellore-632014, India

<sup>\*\*\*</sup>(School of Computer Science and Engineering, VIT University, Vellore-632014, India

### ABSTRACT

This paper is about the use of big data analytics in higher education. In this paper, we see what the big data is and where does it come from. We will also try to find why the big data analytics has become a buzzword in almost every sector today through our literature review on the big data analytics and its applications in higher education sector. Then we see what the big educational data is, how it is generated and analyzed. We found that the two most important types of analytics are- Learning and academic analytics which will be discussed. Several papers describe the benefits of implementation of analytics in the education sector and the opportunities provided which will be discussed in this paper. We also found that the basic characteristics such as size, speed, variety and some other factors are responsible for some issues and challenges to the use of analytics in this sector. We will discuss those issues and challenges and discuss some proposed solutions to address them.

**Keywords:** big data, analytics, higher education, learning analytics, academic analytics.

### I. INTRODUCTION

Higher education is one of the major necessities for the development and growth of the society and the nation. A college student is considered more skilled, both professionally and personally. A highly-educated person gets to enjoy higher incomes and improved social conditions. Today the education sector has become technology oriented to a considerable extent [1]. In many institutes, the students upload their assignments online through their moodle accounts instead of directly submitting to their faculty. Students also write their exams online. Teachers upload their prepared notes on some online portal and students download them from the sources. In several institutes, whiteboards have been replaced by smart boards. They also have digital libraries where issuing and returning of books is without any need of human interference. There are several metric systems at different places in the campus for student monitoring and attendance. This interaction with technology generates a huge amount of data and has transformed the modern world into an era of big data. Many students graduate every year. Each of them is responsible for bringing another tsunami of data in this sector. Hence, this sector needs to become data-driven in to use data as an asset to it.

Big data analytics can address some issues such as retention of students, improving their progression, improving curriculum and teaching quality, graduation rates etc. This sector needs big data to retain a large student population, admissions, funds, and organizational efficiency. Implementation of analytics can help the institutes to enhance their

students' success rate. So far, big data has not been much successful in having an impact on higher education. The campus administration needs to pursue some objectives to harness the power of big data analytics. They need to understand the advantages and the importance of their data. They need to discover new paths to the next generation of analytics. They should identify proper tools and models. They must establish data-driven decision making. There are some ethical challenges as well as the challenge of applying analytics on academics. But, a good data management system and some guideline principles can help the institutes overcome those challenges.

In this paper, section 2. gives an introduction about the big data. Section 3. gives its introduction in higher education. In section 4. we will see the types of analytics and analytics approaches that can be used. Later in section 5. we see the benefits provided by the big data followed by the opportunities it creates in section 6. In section7, we will see the various issues challenges of implementing big data analytics in higher education sector and try to come up with the solutions to these challenges in section 8. Section 9. concludes this paper.

### II. WHAT IS BIG DATA

Some few years ago, even a single terabyte was considered as a huge amount of data. But, today data is being described in petabytes, exabytes, and even yottabytes. Big data, as the name suggests is a huge amount of data that generally refers to the data size in terms of exabytes and beyond. It describes the data that is so big in volume that could not be

gathered, stored, managed and processed by traditional database software traditional tools. Today the big data has taken many sectors by storm. Data has become an important factor for decision-making. The reason for this can be easily found in the definition of the Big Data Analytics- It is a process to examine large data sets for identifying the hidden pattern and other necessary information. This information can be used to make important decisions thus, boosting the system and saving money in the concerned sector. Due to the technological advancements, different public, commercial and social sectors are generating as well as exchanging a huge amount of data from different sources and in different formats. More data cross the internet every second than were stored in the entire internet just 20 years ago. Earlier the characteristics of the big data could be described using 3 Vs -Volume, Velocity and Variety. But, today these can be described more broadly using 7 Vs which are *Volume*- It is related to the amount of data to be analyzed. *Velocity*-It is related to the speed at which data flows in from various sources. The flow of data can be very continuous and hence, it needs to be handled carefully. It is used to describe speed and frequency. *Variety*- Data is received from the various sources and can be in different formats. Some of them can be structured while some can be unstructured. The data in the education sector can in the form of e-mails, audio, video, images, PDFs and monitoring devices [2]. *Veracity*- It deals with the trustworthiness of data. Data may have some noise and abnormalities. So, our data should be clean and free of noise. *Validity*-It deals with the correctness and accuracy of data. Only valid data can lead to right decisions. Therefore, this validity is very important characteristic. *Volatility*-As data generation is a continuous process, some data may no longer remain valid after a certain point of time. A data must be checked for its validity and how longer it should be stored. *Value*-It is the ability to have insights from the data [3]. The current areas where the analytics is applied and experimented involve government services and jobs, healthcare and business data since these streams generate voluminous data which can be studied to understand patterns and behavior. An ideal big data analytic model consists of two stages. *Strategic question*: The question around which the whole idea revolves. *Data analysis and prediction*: Studying and interpreting the data pattern and applying predictions over the same Insight and action: After a keen understanding of the pattern, required and necessary actions to be taken. It's the implementing phase of the analytics.

### III. BIG EDUCATIONAL DATA

In the higher education sector, both the college and its students are equally important. A college cannot exist without students. Similarly, the students do need colleges for their higher education. The institute provides its resources such as faculty, technology (such as labs and software), infrastructure facilities, knowledge, course materials and a degree to its students. In return, students provide the college finances, their preparations, and participation ensuring smooth running and functioning of the system. So, we can say that college is co-created. So, the colleges need to deliver a good service to their students in return for high tuition fees. Big data analytics can help the colleges to achieve so. Big data has a huge potential to impact the higher education sector. It can bring accountability and transparency in the education sector. It can help both learners and organizations to identify their achievements and areas of weaknesses and compare with other such organizations. Faculty and students can track their academics and behavioral progress through big data analytics. The educational systems and organizations differ from each other in terms of their source, size, format etc. Some examples of educational data are data from students' interaction with learning management system, learning activities, syllabi, course materials, curriculum, examination patterns, attendance, results, timetables, data related to administrations, infrastructure facilities provided, information about staffs and teachers, data from social media, campus sensor data, data from log in details to the online course page etc. [4]. Most of the data generated in this sector is of video and audio type followed by text and image. Educational data is classified into two categories - static data and fluid data. The institutes collect, record and store some information. This is called static data and includes *Student record data* such as name, address, age, school, course, grades, contact details, attendance records. *Staff data* includes details about the number of working staffs, their records, which department they belong to etc. *Admission and application data* is about the number of applications for the institute in an academic year, the number of students who could qualify for admission and those who acknowledged their acceptance. *Financial data* includes incomes, expenditures, budget, taxes, profits and losses in a particular year. *Alumni data* is about those who have already completed their education from the institute. It includes information about their current contact details, job positions, workplace etc. *Course data* includes data about each course undertaken by the students and the students enrolled in each course. *Facilities data* includes data about infrastructure facilities like the number of classrooms and labs, equipment, and software provided in the labs etc. The

students' digital interaction with the university generates flow data. This includes login details to the online course page, notes downloaded, biometric given at library, the time they stayed online for, the number of books issued in his name, dates of issuing and returning the books etc.

#### IV. ANALYTICS IN HIGHER EDUCATION

Per the definition given by the Educause, *Analytics* is an overarching concept that that is defined as data-driven decision-making [5]. The data is gathered. It is then turned into information. Finally, this information is used to help learners for progress and betterment. Through analytics in higher education sector, we can see a complete view of everything. It allows both the institute, and its students, make some good data-driven decisions [6]. Analytics can bring innovation which can be of two types -*Sustainable*- It improves the existing system. *Disruptive*- It creates new systems different from the existing system. These innovations separate systems from the mainstream [7]. Analytics in Education sector is classified into the following six levels. *Personal* level indicates the analysis of personal performance in terms of the goals and labels. *Course* level is related to conceptual development. *Departmental* level deals with predictive modelling, patterns of success/failure *Institutional* level deals with the performance of academics, resource allocation etc. *Regional* level involves a comparison between systems and standards. Several other attempts have been made to define the different levels more specifically. *Nanolevel*- It relates to activities in a course. *Microlevel*- It indicates an entire course in the education program. *Mesolevel* - It describes many courses in an academic year. *Macrolevel* - It describes several programs in the institution [8]. The analytics in the higher education sector can be of two types. *Learning analytics* aims at improving learner's success. Learning analytics helps in making decisions that are required to achieve some specific learning outcomes. It refers to the operations at the micro level and the nano level. Another type of analytics is *Academic Analytics* which is concerned with the operations at mesolevel and macro levels. It helps in making decisions that concern with management and other institutional operations. Through learning analytics, patterns in the behavior of students can be identified. Teachers can use this input to identify students at risk and thus help the institutions to take appropriate preventive measures. Nowadays the number of traditional students is decreasing at a very fast rate i.e., the number of dropouts is increasing very tremendously. Thus, learning analytics can be very helpful here. Academic analytics can be defined as the combination of information, technology, culture, management for

managing the academic enterprise [9]. It is concerned with decisions related to the campus services. Financing, administration, academic counselling etc. can be considered as an example of academic analytics. It aims at improving the effectiveness at an organizational level. It is, generally concerned with the working of an academic institution by using learners' academic data as well as institutional data. Thus, in order to improve our higher education sector, we need to analyze data for both learning and academic activities. Learning analytics benefits learners, education and teaching staffs whereas, academic analytics benefits administrators, funders, marketers, the National government and other big organizations such as UNESCO. Earlier learning analytics domain has six dimensions. *Stakeholders* describe the various stakeholders that benefit from the analytics. These stakeholders can be classified into three categories. Those who are directly involved with teaching form primary group (e.g. students, professors etc.) whereas those involved indirectly form secondary group (e.g. parents, alumni etc.). involved with decision making processes form the hybrid group of stakeholders. *Objectives* describe the objectives of analytics at the institute. These can be academic as well as administrative. *Data* describes the data that is being collected for evaluation. *Instruments* describe what analytics systems and tools are available (e.g. Oracle data mining, SPSS clementine, CART etc.). *Constraints* describe the processes that deal with any legal or ethical issues such as privacy, norms, timescale etc. It also deals with those who have access to the data. *Limitations* deal with the fact whether the institutes have skilled and trained staffs for analytics and the clients will understand the visualization of information and be able to use it efficiently [10] [11] [12]. Now, we have a four-dimensional model for analytics. *Data and Environment* relate to 'what' data is available. *Stakeholders* deal with 'who' is targeted. *Objectives* is about 'why' do we need to do the analysis. *Methods* deal with the question 'how' to perform the analysis. There are some other analytics as well. *Predictive analytics* deals with the extraction of information and is an area of statistics analytics. It can be used in prediction of behavior and events [13]. *Visual analytics* is also used for manipulation of big data by combining data analysis, information visualization and human visual perception. *Business Analytics*- This analytics uses data and its analysis to have effective business [5]. It is helpful for institutional enterprise.

We can adopt two types of analytical approaches. The analytics which begins at data and ends on a decision is *data-driven analytics approach*. In this approach, the importance is given to the data and the methods used for cleaning, collecting, storing

and processing these data. Another type of analytical approach is *context or need-driven analytics approach*. This approach starts with a need for a decision. After that, the required analysis is done to support the decisions. [8].

The potential of analytics has still not been realized and explored in areas like optimization of resources, multiple administrative functions and total cost to complete a degree. It is significant to realize analytics of the big data generated in this field as an investment rather than a costly affair.

In big data analytics, the data is acquired first then, information is extracted and cleaned. The data is then represented in the proper structure. On this data, analysis is done and finally the results are interpreted [14]. As we have large data sets to be processed in real time, big data analytics is often associated with cloud computing due to enormous volume.

## V. BENEFITS OF BIG DATA ANALYTICS IN HIGHER EDUCATION

Big data holds great opportunities for higher education some of which are attempted to be implemented while rest are the areas to be explored. It is because of the several benefits provided by the big data analytics. Big data technology can be used to both the institutes and their students' data. Both educational institutes and their students co-exist with each other and are largely inter-dependent analogous to the functioning of an eco-system. So, both have something to give to each other consequently, take from each other. Right from management of enrolment of students, finance and marketing to faculty teaching and research performances, human resources and facilities are the wide areas of applications which can make an impact with the analytics. Some of the significant promising advantages that big data analytics can offer in this territory are described in the next few sub-sections.

### 5.1. Student acquisition

Big data can help the institutes to acquire those students who could stand up to the expectations of the institute from its students and help them fulfil their objectives. To achieve this, we can apply big data to applications, test scores and face to face interviews. This will ensure better and more reliable filter in screening of students. Because of online application process, a student can apply for multiple colleges at a time. This makes it difficult for the institute to predict whether a student will accept its offer or not. Big data can be useful for the college in such a situation. Institutes can take decisions about enrolments using demographics and historical performance data of students [15]. The institutes will be able to know about the demands of the students. Hence, they can improve their existing curricula and

infrastructure facilities, attracting more number of elite students to apply. This in return, will help the institutes generate high revenues and with good students will also generate a wide fame. Thus, ensures a well-being and sound functioning of the educational eco-system. Students can also take better decision about the institute they should apply for. This is possible because of various organizations that rank institutes using big data analytics.

### 5.2. Improving student retention

Many students drop out of the college every year and this number is increasing rapidly for several reasons. Losing a student by an institute affects its income as tuition fee forms a significant part of the institutes' income. Institutes can identify those students who could possibly drop out and take the necessary steps. Data about a student's interaction with his course and his institute can tell about his engagement and involvement with the college [13]. For example, a student who has continuous falling grades and does not spend much time on the campus, or does not log in to the online portal and does not interact with his fellows is more likely to drop out. Online education systems have dashboards to measure the engagement of students [16]. Big data can also be used to identify depressed student with interaction quotients and thus can be very helpful in suicide prevention on the campus. Thus, the interpretation of students' behavior can help the colleges and university to monitor and regulate them well.

### 5.3. Helping students' progress

Through the big data, students can be provided with better feedback on their progress [17]. Through analytics, teachers can examine each student's performance better and thus, provide a more efficient, reliable and immediate feedback. The feedbacks can be given regularly. This system of providing feedback is better than the traditional methods of taking exams. We have this feature enabled in some online education systems. It will also help in the development of a healthy competitive environment among the students which is important for their progress and growth. Several factors such as a student's past performances, his demographic outlook, choice of subject can also be used to predict a student's performance [18] [19].

### 5.4. Enhancement in teaching

If we could identify some patterns regarding how the students study, which patterns yield better results we can also develop an effective feedback for the teachers. For example, if the analysis is made of the most referred learning material can give us an idea about its quality. Thus, feedback can be given to

the concerned faculty about his notes. Several factors such as the use of whiteboard/projector to teach, knowledge about the practical applications given by the faculty, encouragement given to the students to ask their doubts etc. can help institutes develop their feedbacks about the teachers. The lectures can also be recorded and this can be used to analyze and study their teaching strategies and give feedbacks accordingly. They can use data like number of students, students' demographics, students' behavior, students' ability etc. to match the perfect teacher for the class.

### 5.5 Matching students to programs and employment

Big data can help students decide about their major and utilize their college investments to the fullest [20]. Today many students and their parents do not get which institute to choose. This is due to the lack of transparency in the information provided by the institutes regarding the courses offered, placements offered and rules and regulations of the institutes. Different institutes make different promises and offer different resources and so, the students get confused about which institute to enrol in. Big data can be of rescue here as well. The big data can help the students in making such decisions. Data for this purpose can be collected from various sources such as the institute itself and the feedbacks given by alumni. Also, the data can be collected from students and parents to survey about what they expect from the education institutes. Big data analytics can also help students to get their dream jobs. Students can get to know about the skillset that they need to have to qualify for their dream companies and earn a living. Also, on the other hand, companies can identify the candidates who would be the most suitable for them. In this way, the placements rates on the campus can also be improved and the companies can also end up getting the best and most deserving candidates for their vacancies [16].

### 5.6. Improving student's experience

A student's performance depends on several other factors besides his study such as eating, sleeping and social habits and routines. Big data can be used to better understand a student's behaviour pattern and sentiments based on the data collected from social media and their experiences and interactions within the campus. Some data like the amount of time spent in classrooms, hostel, library, restaurants, cafeterias etc. can help in such analysis after understanding his behavioural pattern. Data like correctness of the response to a question, response time, mistakes made also help analyse a student's behaviour[21]. The institutes can also identify if anyone needs some extra attention, motivation, and

support. A student's passion can also be identified and he can be helped by the institute to pursue his passion as well. The institutes can use historical data of students to create a more personalized experience for them [21].

### 5.7. To develop an effective administrative system

Big data analytics can help the administrators to identify needs of the students at administrative and management level and to implement the needful for student and university welfare [16].

### 5.8. To enhance research effectiveness

Big data can be very useful in advanced research projects. The students can apply analytics to a very wide variety of data thus, improving the quality of the research. It can help them in storing and processing a very high volume of research data. The students will get an idea about the feasibility of the projects, what time it should be continued for and what should be the required pace [4].

### 5.9. Better data transparency

With a proper management and sharing of data the redundancy can be avoided to a considerable extent.

## VI. OPPORTUNITIES PROVIDED BY BIG DATA ANALYTICS IN HIGHER EDUCATION

We know that the big data analytics provide many benefits. Hence, it offers several opportunities for innovation in the future.

### 6.1. Motivation

We have seen that how the analytics could do wonders in improving students' performance. If students could realize its impact and if they themselves witness its benefits, they will become more interested in inputting data for the process and betterment.

### 6.2. Personalization

Big data analytics could help developers to develop personalized courses according to cope up and match with the learning ability of each student. Personalization could bring a significant revolution in the e-learning systems and improvise the results [22]. It will ensure higher rates of classroom success and higher engagement of students [23]

### 6.3. Efficiency

It helps us save a lot of time and efforts. For example, if a student is preparing for a job, big data can help him prepare thoroughly by giving him knowledge about who had earlier got qualified for the job and what preparation they had thus, understanding what patterns need to good results.

#### **6.4. Collaboration**

Big data analytics can promote collaboration among various institute at regional, national and international levels. It will also help institutes make comparison among them [10].

#### **6.5. Finance**

It will help institutes in a reduction of their expenditures by managing their finance efficiently. Analysis of what patterned expenditure would be the most sustainable help in having efficient finance management [10].

#### **6.6. Effectiveness in learning**

Self-measurement by learners and teachers can improve learning. We can achieve higher student success rate and knowledge flow [10] [19].

### **VII. ISSUES AND CHALLENGES**

There are several issues regarding the big data analytics that has been identified for this sector. Those issues can be broadly classified as storage issues, management issues and processing issues [24]. There are some other ethical and practical challenges as well such as privacy, security, consent, availability of data etc. [25].

#### **7.1. Issues**

##### **7.1.1. Storage issues**

A tsunami of data is being generated everywhere, every day and by everyone. Hence, the storage of data is a great issue as the storage space provided by the available storage devices might not be sufficient for the data generated. The data is available from different sources. The data collected may have different formats such as text, image, video, audio etc. Some data are available in structured form while some are not and they need to be transformed into a structured form for storage [14]. The Institute must make proper decisions regarding what to discard and what to keep for further analysis.

##### **7.1.2. Management issues and affordability**

It involves the issues pertaining to access, allegation, and governance. Data is gathered from different sources and that too in different formats. This is known as heterogeneity of data. In most of the institutes, data is managed by different departments. While implementing the analytics, bringing those data together and relating them adds up more challenges to its management. Analytics algorithms accept only structured data. Even after cleaning the data, there may exist some errors and incompleteness that can make its management more difficult [2] [27]. This issue is one of the major areas of concern.

##### **7.1.3. Processing issues**

It is due to the size of the big data. It may take a lot of time for its processing if we do not use appropriate analytics tools and algorithms. This is also called timeliness [2]. There are many patterns that cannot be found by the computer easily but, humans can detect them. The volume of data is increasing tremendously in terms of size. This is known as scalability. In the past, Moore's Law helped us deal with this challenge. But, today data has grown in volumes and the traditional computer resources cannot help in storing, managing and processing them [27].

#### **7.2. Challenges**

There are various challenges in the application of big data analytics in this sector. To utilize the benefits provided by the big data analytics to the fullest, the institutes must address those challenges. Of course, this sector also faces those basic issues that other sectors do due to its size, speed, and variety. There are some challenges specific to this sector.

##### **7.2.1. Student consent**

It is related to the ethical concerns and is about students' consent to the use of their personal data. Sometimes a student might not be comfortable with the usage of some of his personal data for analytics. In such case, the student's decision must be respected.

##### **7.2.2. Privacy and security**

The data used for analytics must not be too personal. It should be related to his academics. Sometimes, the institute might want some personal data such as his socio-economic status, physical/mental health state. Such data must be handled carefully so that the students' privacy is not intervened. A single student can generate a very large amount of data that can reveal a lot about him. The institute must store these data securely and prevent them from being misused. The fear of misuse of the data generated and analyzed is another major prominent hindrance. Data like grades and evaluations must be kept private so that students can learn to gain knowledge and not just reputation. Sometimes research, unpublished work, theory etc. need to be kept private for some time so that they can be developed properly [28].

##### **7.2.3. Fooling the system**

Students can easily game and manipulate the system. For example, to manipulate the analysis about his progress and performance, a student can issue many books from the library or spend hours in the library just killing the time and pretending to

learn. Security- Some data sets may be very precious and can be targeted for cybercrimes. It is the responsibility of the institute to keep the data secured.

#### 7.2.4. Institutional capability

The quality of data available can affect the results of analytics. The data must be consistent and accurate. The institutes should be capable enough of collecting and accessing data. They should have a strong management system to link the data available from different departments. Some data can be expensive to gather.

#### 7.2.5. Staff capability

Big data is new for the higher education sector. So, many employees are unaware of the trends. Many of them are unfamiliar with analytics. They need to know about the advantages of learning analytics. So, the institutes need to make huge investments for training their employees. They also need to hire new analysts for this purpose.

#### 7.2.6. Absence of common language

To ensure clear data governance and a good communication, we need a common language across various platforms. This sector is lagging in this context.

#### 7.2.7. Correlation vs Causation

The institutes should not directly jump to some conclusion on identifying any pattern. A correlation between two variables does not always imply that one of them is the cause for another [29].

#### 7.2.8. Claims beyond data

A researcher may fake the analytics results [30]. An example for this issue can be University rankings. One institute can claim itself the best on the basis of academics, while the other on the basis of infrastructure and so on.

### VIII. SOLUTIONS TO THE ISSUES AND CHALLENGES

To deal with the ethical challenges such as students' consent and privacy, the institutes should make sure that the data is processed. The data should be gathered only for specific purposes. It should be adequate and relevant. It should be stored and used only for the required time. Mapping out the strategies and effective planning around the strategic questions impacts a wide area. Self-correcting analytics systems can be used to avoid any manipulation of the system. For example, if the system finds out that a student spends more hours on the online study portal than that is required, this system will gradually stop considering this factor for evaluation. The issue of data security can be dealt with using good data management systems. The system should identify the sensitive data and ensure its encryption to protect it.

To improve Institutional capability and resourcing the institutes need a good management system. To improve staff capability the institutes, need to improve the analytics and data skills for its staffs through training. The data needs to promote as an important asset to the vector. Digital literacy needs to be improved across the institutes. The introduction of some sensors and matrices in the campus can make the task of data acquisition easier. Sharing and development of open-access analytic software tools must be promoted. Institutions and industries can join to create big data collaboration [26].

### IX. CONCLUSION

We saw how the big data can improve the education. We saw where the big educational data comes from. The power of big data can be harnessed to revolutionize our higher education system completely. We can come up with new goals and methods. Our higher education sector can be enriched with new teaching and learning methods. Big data can improve not only the quality of learning process, teaching process, and campus services but also help plot out an effective and sustainable management model. Big data can improve enrolment, student performance and optimization of resources. The institutes can be unified based on common analytics goals and the data can also be centralized. In institutes, big data can ensure better learning in very short time. But, the era of big data has just started and so, there are numerous challenges in front of us. These challenges are not just the obvious issues of scale, but also heterogeneity, incompleteness, lack of structure, error-handling, timeliness, privacy, and security. Due to our growing interest in data and analytics, there is a need for improved research techniques, methods, and tools. Big Data requires smart policies that address the issues and challenges it poses.

### REFERENCES

- [1] Saptarshi Ray. Big Data in Education. *Research Gate*, 2013.
- [2] Jaseena K.U., Julie M. David. Issues, Challenges and Solutions: Big Data Mining. *Computer Science & Information Technology (CS & IT.)* 2014, 10.5121/csit.2014.41311.
- [3] Krishna Madhavan, Michael C. Richey. Problems in Big Data Analytics in Learning. *Journal of Engineering Education*, 2016, 10.1002/jee.20113.
- [4] Improving Higher Education Performance with Big Data. *Oracle Enterprise Architecture White Paper*, 2015.
- [5] Angela van Barneveld, Kimberly E. Arnold, and John P. Campbell. Analytics in Higher

- Education: Establishing a Common Language. *Educause*.
- [6] Dr. J Meenakumari, Jayashree M. Kudari. Learning Analytics and its challenges in Education Sector a Survey. *International Journal of Computer Applications*, 2015, 0975 – 8887.
- [7] B. Tulasi. Significance of Big Data and Analytics in Higher Education. *International Journal of Computer Applications*, 2013, 0975 – 8887.
- [8] Christos Vaitzis, Vasilis Hervatis, Nabil Zary. Introduction to Big Data in Education and Its Contribution to the Quality Improvement Processes. *Intech* 2016, 10.5772/63896.
- [9] Philip J. Goldstein. Academic Analytics-The Use Of Management Information And Technologyin Higher Education. *ECAR*, 2005.
- [10] Shikha Anirban. Big Data Analytics in the Education Sector: Needs, Opportunities and Challenges. *International Journal of Research in Computer and Communication Technology*, 2014.
- [11] Rajni Jindal, Malaya Dutta Borah. A Survey On Educational Data Mining And Research Trends. *International Journal of Database Management Systems (IJDMS)* 2013, 10.5121/ijdms.2013.5304.
- [12] Thilini Ariyachandra, Joseph Landers, Xavier University, Mark Frolick. Analytics in Behavioral Intervention Education. *Issues in Information Systems*, 2016, pp. 236-243.
- [13] Predictive Analytics in Higher Education, Data-Driven Decision- Making for the Student Life Cycle . *EDUCAUSE*, 2013.
- [14] Divyakant Agrawal, Philip Bernstein, Elisa Bertino, Susan Davidson, Umeshwas Dayal, Michael Franklin, Johannes Gehrke, Laura Haas, Alon Halevy, Jiawei Han, H.V. Jagadish, Alexandros Labrinidis, Sam Madden, Yannis Papakonstantinou, Jignesh Patel, Raghu Ramakrishnan, Kenneth Ross, Cyrus Shahabi, Dan Suciu, Shiv Vaithyanathan, and Jennifer WidomChallenges and Opportunities with Big Data. *Cyber Center Technical Reports*.
- [15] [https://infocus.emc.com/william\\_schmarzo/what-universities-can-learn-from-big-data-higher-education-analytics/](https://infocus.emc.com/william_schmarzo/what-universities-can-learn-from-big-data-higher-education-analytics/)
- [16] Anthony G. Picciano. The Evolution Of Big Data And Learning Analytics In American Higher Education. *Journal of Asynchronous Learning Networks*.
- [17] Athanasios S. Drigas, Panagiotis Leliopoulos. The Use of Big Data in Education. *IJCSI, International Journal of Computer Science*, 2014.
- [18] Annapoorna Manohar, Pooja Gupta, Veena Priyanka, Muhammad Fahim Uddin. Utilizing Big Data Analytics to Improve Education.
- [19] Pooja Thakar, Anil Mehta, Manisha. Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue. *International Journal of Computer Applications*, 2015, 10.5120/19412-1007.
- [20] <https://blog.electriclightwave.com/big-data-is-improving-college-student-outcomes-and-more/>
- [21] B R Prakash, Dr.M. Hanumanthappa, Vasantha Kavitha. Educational Data Mining and Learning Analytics Applications. *International Journal of Innovative Research in Computer and Communication Engineering*, 2014, 2320-9801.
- [22] Ben Daniel. Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology. Research Gate*, 2014.
- [23] <http://er.educause.edu/blogs/2017/3/can-making-college-completion-paths-more-personalized-increase-student-success>
- [24] [Stephen Kaisler, Frank Armour, J. Alberto Espinosa, William Money. Big Data: Issues and Challenges Moving Forward. *46th Hawaii International Conference on System Sciences*, 2014.
- [25] Rebecca Eynon. The rise of Big Data: what does it mean for education, technology, and media research? *Learning, Media and Technology*, 2013.
- [26] The Implications of Big Data for Graduate Education, *Council Of Graduate Schools*.
- [27] <http://cra.org/ccc/wp-content/uploads/sites/2/2015/05/bigdatawhitpaper.pdf>
- [28] <http://er.educause.edu/blogs/2017/2/i-have-nothing-to-hide-why-should-i-care-about-privacy>
- [29] [https://en.wikipedia.org/wiki/Correlation\\_does\\_not\\_imply\\_causation](https://en.wikipedia.org/wiki/Correlation_does_not_imply_causation)
- [30] <https://www.wired.com/2013/02/big-data-means-big-errors-people/>