International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2017): 7.296

Reconnoitring the Higher Education Systems using Big Data Analytics

D. Naga Jyothi[1], Dr. Y. Rama Devi[2]

[1]Assistant Prof., CSE, CBIT, Telangana, India

[2]Professor, CSE, CBIT, Telangana, India

Abstract: ***Now-a-days large amounts of data are needed to be gathered due to the increased competition and thus many companies are having lots of terabytes of data to be stored and analysed. There is lots of data generated at high volume, velocity and variety based on the sources it is coming from. This is the scenario where we need complex analytics to deal with Big data. So far, most database management innovation has not kept pace. Performing ad hoc queries on such large data volumes does not work naturally for existing database management systems (DBMS), which use a row-oriented design for write-intensive transaction processing rather than for read-intensive analytics. The traditional view of big data is not enough. Rather than focusing exclusively on what technology big data brings, it has to be looked as to what value it can create. Big data is insignificant in a vacuum. Its potential is unlocked only when leveraged to drive decision making. To enable such evidence- based decision making, organizations need efficient processes to turn high volumes of fast- moving and diverse data into meaningful insights. Thus, the current research explores mainly on big data analytics, the main opportunities it gives rise to, and how big data should be expanded to support analytics. This research paper presents application of big data analytics on Higher Education aspects to choose more education models, which can enhance the performance of the instructors and students. In spite of the increasing opportunities for instructors and students, online learning also brings challenges due to the absence of direct human contact. Online environments allow the generation of large amounts of data related to learning/teaching processes, which offers the possibility of extracting valuable information that may be employed to improve students' performance. This paper includes various analytics' approaches and the educational data mining and learning analytics.***

Keywords: Big data, Big data analytics, educational data mining, learning analytics

1. Introduction

The overall process of extracting intuitions from big data is divided into Data Management and Analytics. **Data Management** involves processes and supporting technologies to acquire and store data and to prepare and retrieve it for analysis. **Analytics** refers to the techniques used to analyse and acquire intelligence from big data. The user organizations are implementing definite forms of analytics, particularly what is sometimes called advanced analytics. This is a collection of associated techniques and tool types, usually including predictive analytics, data mining, statistical analysis, and complex SQL. The list can be extended to cover data visualization, artificial intelligence, natural language processing, and database capabilities that support analytics (such as MapReduce, in-database analytics, in-memory databases, columnar data stores).

All these techniques have been around for years, many of them appearing in the 1990s. The difference today is that many user organizations are using them. That's because most of these techniques adapt well to very large, multi-terabyte data sets with minimal data preparation that brings us to big data (fig.1).

1.1 Defining Big Data Via the five Vs

Most definitions of big data focus on the size of data(volume) in storage. Size matters, but there are other important attributes of big data, namely data variety, velocity, value and veracity.

Velocity denotes the speed at which vast amounts of data are being generated, collected and analyzed. **Volume** refers to the incredible amounts of data generated each second from social media, cell phones, cars, credit cards, M2M sensors, photographs, video, etc (fig.3). Collecting and analyzing this data is clearly an engineering challenge of immensely vast proportions. **Value** refers to the worth of the data being extracted. Having endless amounts of data is one thing, but unless it can be turned into value it is useless. **Variety** is defined as the diverse types of data that can be used. **Veracity** is the quality or trustworthiness of the data, as to know how accurate is all this data.

The popular discourse on big data, which is dominated and influenced by the marketing efforts of large software and hardware developers, focuses on predictive analytics and structured data. It ignores the largest component of big data, which is unstructured and is available as audio, images, video, and unstructured text (fig. 2). It is estimated that the analytics-ready structured data forms only a small subset of big data. The unstructured data, especially data in video format, is the largest component of big data that is only partially archived.

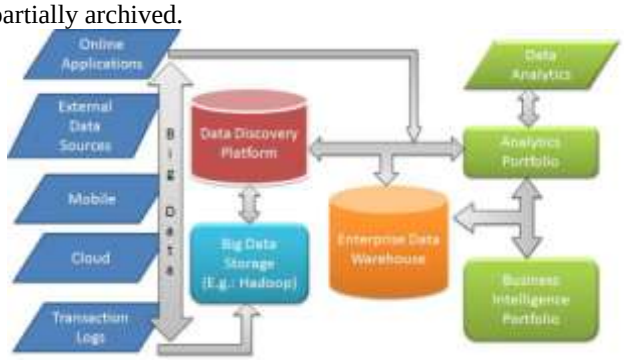

Figure 1: Big data related

1.2 Big data analytics

In this section, big data analytical techniques are briefly reviewed for structured and unstructured data and few of the analytic techniques being used.

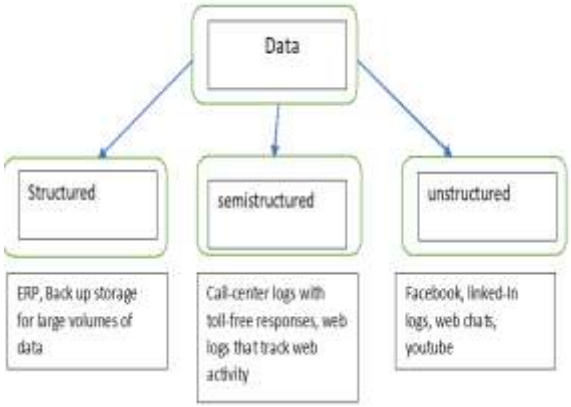

Figure 2: Different types of data

Text analytics
Text analytics (text mining) refers to techniques that extract information from textual data. Social network feeds, emails, blogs, online forums, survey responses, corporate documents, news, and call center logs are examples of textual data held by organizations. Text analytics involve statistical analysis, computational linguistics, and machine learning. Text analytics enable businesses to convert large volumes of human generated text into meaningful summaries, which support evidence-based decision-making. For instance, text analytics can be used to predict stock market based on information extracted from financial news (Fig.3).

Figure 3: Big data Analytic Techniques

Audio analytics
Audio analytics analyse and extract information from unstructured audio data. When applied to human spoken language, audio analytics is also referred to as speech analytics. analytics is also referred to as speech analytics (Fig.3).

Video analytics
Video analytics, also known as video content analysis (VCA), involves a variety of techniques to monitor, analyse, and extract meaningful information from video streams. Although video analytics is still in its infancy compared to other types of data mining, various techniques have already been developed for processing real-time as well as pre-recorded videos. The increasing prevalence of closed-circuit television (CCTV) cameras and the booming popularity of video sharing websites are the two leading contributors to the growth of computerized video analysis (Fig.3).

2. Big Data Analytics in Educational Systems

Big Data Analytics is relevant in addressing a significant number of pressing issues for **education systems** [8], key among them are 1) increasing educator effectiveness; 2) harnessing insights from learning experiences; 3) delivering education for all that may also be tailored for individual learners needs and 4) equipping students with relevant skills for their future careers. Institutions of higher learning are faced with pressure to improve quality of learning outcomes while cutting costs and to build new value relationships with students. This is achievable when these institutions have the capability of leveraging the data collected during the enrolment stages, finance sourcing and instructional process to support greater learning outcomes.

The Hadoop ecosystem which includes Pig, Hive, Mahout, and RHadoop is a good example of the Big Data technologies that can be deployed in educational systems. Hadoop and NoSQL databases are capable of enabling analysis of large heterogeneous datasets at unprecedented speeds.

The **Education sector** is nowadays becoming more technology oriented. The digital revolution in developing countries is leading to development of new technologies such as ubiquitous computing devices and the **Massive Open Online Courses(MOOC)** all of which are radically transforming the mode and accessibility to teaching and learning [3]. These Massive open online courses are generating huge amounts of data which are relevant for Big Data Analytics [9]. The **virtual classroom** is another important tool on which distance learning programs are based upon, and on which the instructors can conduct the webinars to share the learning material. Big Data analytics has the potential to positively impact all the major areas that are of importance for an institution of learning; in areas such as student enrolment and retention, integrated information management and reporting, operational cost management, regulatory compliance and research. According to report by [11], through analytics learning, Institutions can perform thorough analyses of student and learning data to make informed decisions on future course offerings in order to cater for the needs of potential and existing students.

The decreasing costs of big data storage, open source software such as Apache Hadoop, NoSQL databases, network bandwidth and on-demand access to resources through cloud computing are bringing these complex technologies close to nearly everyone [10].

Many colleges and universities have demonstrated that analytics can help significantly develop an institution in such strategic areas as resource allocation, student success, and finance. Leaders of Institutions of higher education are looking into ways of initiating analytic programs in their institutions when they learn about transformations occurring at other institutions through big data analytics [12].

Big data and analytics in higher education can be transformative, altering the existing process of administration, teaching and learning and contribute to the policy outcomes by helping to the address existing challenges faced by the educational institutions.

Educational Data Mining is concerned with developing, applying and exploring computerized methods to detect patterns in large collections of educational data. Data of interest is not restricted to communication of individual students with an educational system (e.g., navigation behaviour, input to quizzes and interactive exercises) but might also include data from collaborating students (e.g., text chat), administrative data (e.g., university, university location, instructors), and demographic data (e.g., gender, age, university grades).

Academic analytics describes the combination of technology, information, culture and information to manage the academic enterprise. The term applies to all aspects of handling the business of an academic institution, from enrolment management to finance and budgeting as well as everything associated with the student experience of learning.

Learning analytics, is related to educational data mining, and is an emerging area with a series of computational and psychological methods which can identify and understand individual learning styles and hook the content that better matches the learner's learning style, abilities and interests. It is related to academic analytics, action analytics and predictive analytics which have implication to the academic systems.

Learners and faculty are benefitted by **Learning Analytics** by analysing

1) Course-level like social networks, conceptual development, discourse analysis
2) Departmental level like on predictive modelling on the patterns of success/failure

Administrators and funders are benefitted by **Academic Analytics** by analysing

1) Institutional level like learner profiles, performance of academics, knowledge flow
2) Regional level like comparison between Systems
3) National governments and education authorities get benefitted by National and international level Academic Analytics.

Successful institutions have used Big data in the following ways

1) Creating a culture of completion and out placement.
2) Reducing the non-productive credits
3) Redesigning delivery of instructions
4) Redesigning core support services such as human resources, academic services and finance to produce strategically useful data.
5) Optimizing non-core services and operations

3. Big Data Challenges in Educational Institutions

The widespread analytics use is limited mainly to the areas of enrolment management, student progress and resource optimization. This limited scope of analytics is mainly due to barriers of affordability, data, culture, expertise and communication. In fact, most of the literature surrounding the data analytics is surrounded around data itself.

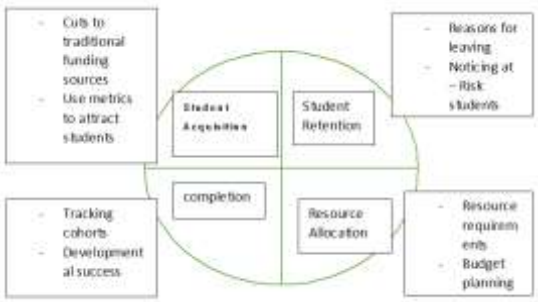

Figure 4: Challenges in Educational System

Various other challenges coming ahead in the educational systems are Decentralised Education System, disparate stake holders, Lack of infrastructure in educational institutions for data collection and analysis, lack of teacher training and significant start-up costs.

3.1 The Process or Workflow

This process can work using a learning management system. When the students learn and study online, they are likely using an online learning system. These learning systems can interact with a student by delivering individualized subject content and assessments.

i) Detailed data about the student's experience is collected and stored in the database.
ii) That data is used to make the predictions about the students' future performance.
iii) Those predictions and feedbacks are displayed on visual dashboards
iv) Students receive learning material appropriate for their performance level and interests
v) Teachers, administrators and developers can intervene and help as necessary.

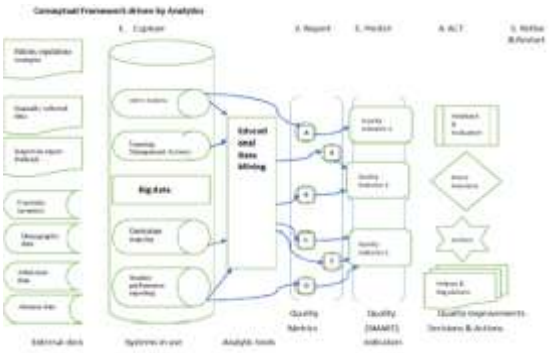

Figure 5: Conceptual Framework driven by Analytics

3.2 Educational data mining and learning analytics can help to handle the above challenges

Educational Data Mining focuses on developing the new tools and algorithms for discovering the data patterns
It can answer the questions like:

i) What sequence of topic is most effective for a specific student?
ii) Which student actions are associated with better learning and higher grades?
iii) Which actions indicate satisfaction and engagement?
iv) What features of an online learning environment lead to better learning?

Learning Analytics focuses on applying tools and techniques at larger scales in instructional systems. It can answer questions like:

i) When are students ready to move on to the next topic?
ii) When is a student at risk for not completing a course?
iii) What grade is a student likely to receive without intervention?
iv) Should a student be referred to a counsellor for help?

4. Conclusions

It is found that data and analytics can help institutions to better understand themselves – 'know thyself', and having timely, up-to-date and accurate knowledge of their own business. It is essential to shape the successful university of the future. There is potential for learning analytics to be used to help institutions to support their students through their educational journey. Thus, all institutions should consider introducing an appropriate learning analytics system to improve student support and performance. There are many instances where universities are already sharing data and intelligence to the benefit of the sector as a whole. This experience can be used to further foster high quality teaching through the higher education system and to continue to shape and build a highly successful sector. National bodies can help lead institutions through the next few years of change in order to build a more dynamic, better informed and more knowledgeable sector.

Leveraging analytics to improve operations and enrich learner experiences will unlock a world of opportunity for institutions of higher learning. Analytics improves operational decision making while helping to measure institutional performance against goals. Analytics can help institutions craft learning experiences according to ability, learning style and personal preference, increasing student success. Teaching and learning will become agile and deeper when the learning platform adapts to and engages with students as individuals.

References

[1] Julius Murumba , Elyjoy Micheni, 'Big Data Analytics in Higher Education: A Review', The International Journal of Engineering and Science (IJES) || Volume 6 || Issue 6 || Pages PP 14-21 || 2017 || ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805
[2] Amir Gandomi , Murtaza Haider ,'Beyond the hype: Big data concepts, methods, and analytics', International Journal of Information Management Volume 35, Issue 2, April 2015, Pages 137-144
[3] Uthayasankar Sivarajah , Muhammad Mustafa Kamal, Zahir Irani, Vishanth Weerakkody, 'Critical analysis of Big Data challenges and analytical methods', Journal of Business Research, Volume 70, January 2017, Pages 263-286
[4] Zachary A Pardos , 'Big data in education and the models that love them' , Current Opinion in Behavioral Sciences 2017, 18:107–113, www.sciencedirect.com
[5] C.L.Philip ChenChun-YangZhang, 'Data-intensive applications, challenges, techniques and technologies: A survey on Big Data', Information Sciences,Volume 275, 10 August 2014, Pages 314-347
[6] Darrell M. West , 'Big Data for Education: Data Mining, Data Analytics, and Web Dashboards', September 2012
[7] Ben Daniel, ' Big Data and analytics in higher education: Opportunities and challenges', British Journal of Educational Technology (2014)
[8] Marsh, O., Maurovich-Horvat, L., & Stevenson, O. (2014). Big Data and Education: What's the Big Idea. Big Data and Education conference. UCL .
[9] Ray, S. (2013). Big Data In Education. Gravity (20).
[10] Ohri, A. (2015, March 6). Big Data Initiatives in Developing Nations. Retrieved April 10, 2017, from Big Data & Analytics Hub: http://www.ibmbigdatahub.com/blog/big-data-initiatives-developing-nations
[11] Oracle Corporation . (2015). Improving Higher Education Performance With Big Data. Redwood Shores, CA 94065, USA: Oracle Enterprise Architecture White Paper
[12] Bichsel, J. (2012). Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations . Louisville, CO: EDUCAUSE Center for Applied Research.
[13] Chris Dede, Andrew Ho, and Piotr Mitros, 'Big Data Analysis in Higher Education: Promises and Pitfalls', SEPTEMBER/OCTOBER 2016 EDUCAUSE review.