

Future trends and scopes of Big Data Analytics in the field of Education

Jaimin N Undavia¹, Sheenal Patel², Atul Patel³

¹⁻³Assistant Professor, CHARUSAT University, India,

¹jaiminundavia.mca@charusat.ac.in,

²sheenalxpatel.mca@charusat.ac.in,

³atulpatel.mca@charusat.ac.in

Abstract— Big data is one of the most popular research trends in the current time. Popularity and usability of Big Data make it most promising research area in past decade. Learning analytics using Big Data is the emerging field and most of the areas are taking use of these results and outcomes as a yield of this analytics. It becomes important to learn and understand the learning analytics and outcome of the Big Data. Increasing data and demand in versatility in results have created new challenges in almost every field. Higher education institutions are also one of them as they have to face new challenges to maintain their reputation and students' growth. Institutes have responded by adopting analytics based approach to empower organizational and educational effectiveness. Making analytics actionable in the field of IT becomes mandatory for the higher educational institutes. This paper has identified different areas of education which can take advent use of such analytics and higher education can sense their growth in terms of quantity and quality. Quality assurance is the motive of the research paper so that we can put the results of analytics actionable.

Keywords: Education, Big Data, Learning Analytics, Academic Analytics

I. INTRODUCTION

In recent years, availability of data among all the business becomes easy and growing at a high speed. In this circumstances, such organizations or business with excessive availability of data are tend to become data centric in all aspect of their business. Big Data a recent buzzword, proved very much useful in decision making where data is responsible and organization growth and future aspects are more clear and specific as compared with previous methods and tools. Solely Big Data is not targeted or used in the business but it is being joined or integrated with another striking and most popular term analytics. Recent time is the witness of the emergence of terms like Business Analytics, Risk Analytics, Learning Analytics etc., are used with Big Data to produce some wonderful and comprehensive outcomes out of available data. This technology proves much beneficial especially in strategic decision making of various business [1].

As mentioned above, Big Data in the recent buzzword in the fields of the Information Technology, it becomes equally important to know the reason for the same. What such recent technologies are responsible for or what is the impact of such technology in current trends or businesses? Recent technologies which can discover the new capabilities at rapid development speed and which can open horizon of new challenges and opportunities to global issues [2]. For example, in the game of cricket, teams utilize the data and prepared their game plan accordingly[3]. Big Data analytics can also be helpful to protect the environment by monitoring biodiversity of rainforest[4]. In this paper, we have tried to highlight the history of Big Data, its current and future trends and its proposed implications in the field of education.

II. INTRODUCTION TO BIG DATA

After the abstract information of the Big Data technology, now it's time to define Big Data in the systematic way. It is difficult to define big data in general because the definition changes from requirements to requirements. Basically, it focuses on the data to be examined, ingested or processed. Amount of available data in different business is also different, like one company may have 35TB data whereas another may have 55PB. Big Data can be defined as "Big Data is an evolving and emerging term that describes any voluminous amount of data in structured, semi structured or unstructured form which has the predicted power and potential to be mined for knowledge and information against future aspects.

Basically, Big Data can be elaborated with 4 Vs [5]. Most big data community have explored these 4 Vs as Volume, Variety, Velocity and Veracity. One can add another V as Value but it is not accepted by global groups of big data.

A. Volume

Volume is the most important characteristics of Big Data which supports the term "Big" in the Big Data. This characteristic highlights the fact that the big data is not only high volume of available data but it is also increasing at a specific speed. So these volume of data is increasing at rapid speed. Sources of such data may be

smart phones, business transactions, social media, study patterns, behavioral patterns, travel pattern of specific category people, super stall surveys, etc., in digitized format [5].

B. Variety

It is another focused development in the technology as more and more information is digitized. The term variety refers the nature of data that may be in structured, unstructured and semi structured form. Structured data like date, time, name, age, contact number, etc., are of this category. Such structured data is augmented by unstructured data which turns into the data like MRI images, web pages, web logs, audio files etc. Unstructured data is a fundamental concept in big data and which can be defined and understand by comparing it with structured data.

At a glance, structured data can be well defined with a précised set of rules whereas unstructured data is having no protocols. In recent times, there are many ways to express ideas and thoughts through a comment, a like button, a voice or a picture too.

Semi structured data is the data which has some formal set of rules which are not as strict as in structured data and not as liberal as in unstructured data.

C. Velocity

This term of Big Data is highlighting the speed of processing of frequently incoming data. To imagine such huge and constant data updates and incoming, we can think of traffic of Facebook server which is responsible to record each activity of any of the user across the globe. We can also think of card swipes from various banks, restaurants, shopping malls which gives us details of every minute pings to telecom carrier server. Best example of streaming application can be Amazon Web Services which captures streaming data of Amazon users.

D. Veracity

Veracity concerns with the reliability of the available data. It is the sense of reliability of data that organization or business rely upon. Can the ownerof the business rely on the fact that the data is in presentable form? Every clever owner knows that the data which is collected is always heterogeneous.

E. Value

The most debated term, present to describe Big Data is “Value”. This value refers not the cost to handle or maintain such huge amount of data but it refers the decisions, judgments or knowledge that are coming out of these data. Those outcomes from the data produces some value to the organization or the business.

III. BIG DATA – AN ANALYTICS FOR EDUCATION SECTOR

Analytics through big data has many dimensions and areas as well for the better outcome and research proposals. Education sector is one the prominent area where Big Data analytics found much helpful and effective. The constant evaluation in data variety and volume forces management to look upon to it and have some betterment in their work and process [6].

Education institutes produces large volume of data which need to be analyzed to sustain in such a competitive environment. These institutes needs to have defined process to analyze their data for the development of their learning and academic activities. Education institutes may have different dimension of Big Data analytics as shown in figure.

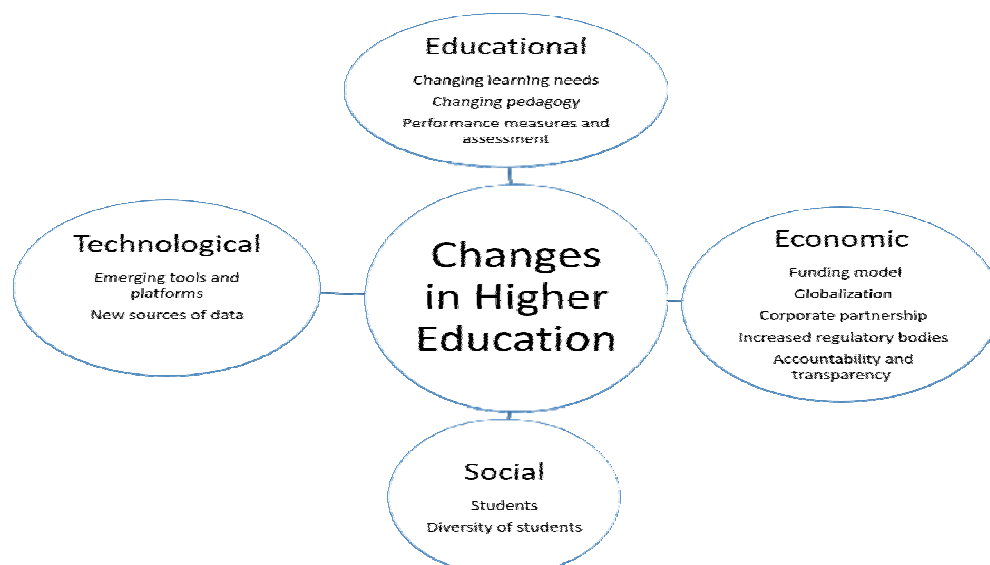


Fig.1. Big Data Analytics in Education Sector

Generally, the term “Business Intelligence” is used when corporates are trying to get some fruitful hidden information form insights of data. When it applied to education sector then these analytics has two major categories[7]. Both these categories have their own merits, effectiveness, affecting parameters, involved departments, etc.

A. Learning Analytics

These analytics mainly focuses on the effectiveness of leaning approach of the learners. It concerns with the data about learners which is to be collected analyzed and reported. These analytics uses data about learners for understanding and optimizing learning environment [7].

B. Academic Analytics

These analytics focuses on overall improvement in academics of the organization. It deals with the organizational processes, workflow and resources allocations. It uses data about learners, academic and institutes[7].

Generally, in the field of education, challenges are opportunities with Big Data Analytics are narrated as below.

The two approaches identified in the field of education are strictly concept but they do not have any idea about tool set which can actually evaluate the learner’s performance and comparison between the different set of learners. To overcome such weakness, an integrated and extensible toolset may be required which can assist the organization to evaluation performance of learner, academic process, success ration of the organization, etc. Such integrated and extensible toolset may produce the results in terms of statistical or mathematical analysis to empower the overall process of the education system. The tool must produce accurate assessment, results are broad enough to cover many aspects of the study, based on multi source of data and more importantly integrated.

Apart from such robust tool which may produce high degree of results, following challenges have also been identified when analytics is going to apply over educational data[6]. Ensuring data flow, training to practitioners, synchronization of data and sources of data, data smoothing, converting vision to venture and practices. Available data of different purpose and nature are available through educational institute and they may be utilized with a single and simple objective. These variety of data that may evolve through educational institute are shown in following figure.

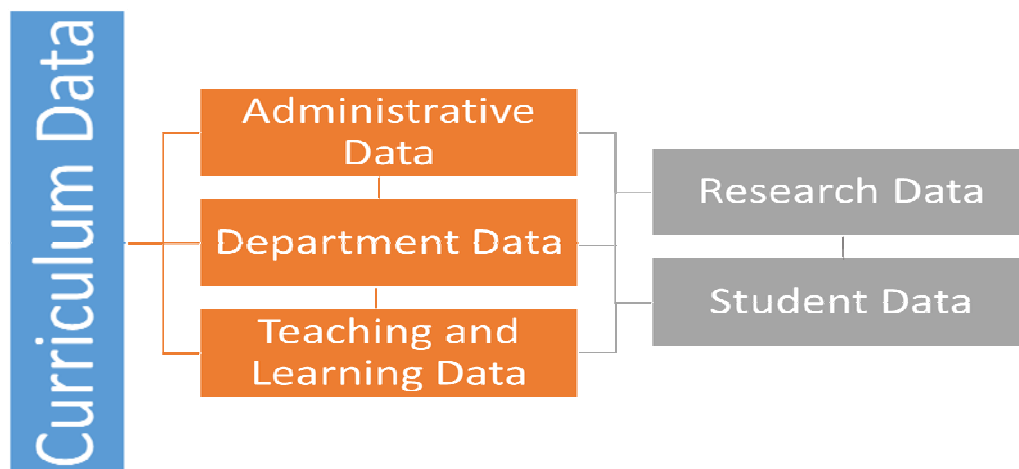


Fig.2. Various Data Sources in Education Sector

Opportunities are also there when we are dealing with the Big Data analytics in the education sector. Students, faculties, administrators, management, educationalist may get avail with such opportunities. Identified opportunities are as follows:

- a. Risk involved in learning and academics can be reduced at high level for the betterment in overall working.
- b. High success ration can be achieved by the institutions by maintaining smooth process framework and knowledge flow.
- c. Collaborative approach can be adopted to empower the relationship with industries and institutes.
- d. Self-measurement techniques can be developed and established the effective learning environment.
- e. Financial Performance can be improved by proper utilization of available resources.

IV. CONCEPTUAL FRAMEWORK OF BIG DATA IN HIGHER EDUCATION

Big Data analytics results guarantees that some meaningful decision can be taken out of the efforts deployed on huge amount of data. Actionable information can be gathered or expected as a result of this analytics. In recent times, big data analytics influence higher education practice at a very high extent to achieve effective and evidence-based strategic decision making [8]. Big data makes this decision making process cost effective by use of a précised proposed framework model[8]. Components of big data is shown in the following figure[8].

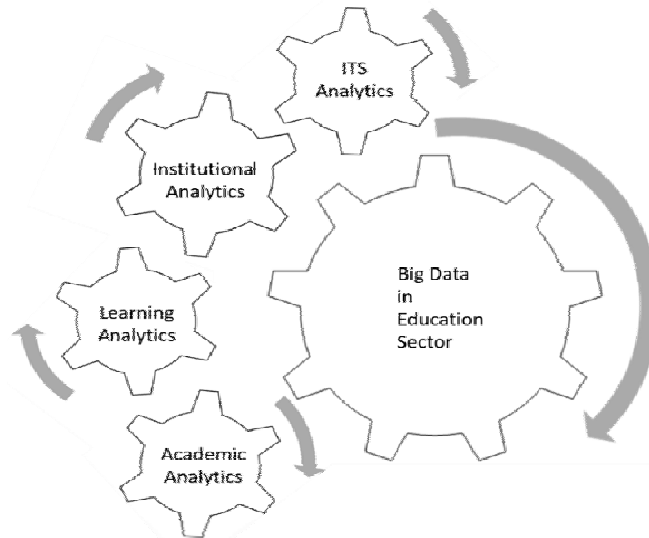


Figure – 3: Components of Big Data Analytics in Education Sector

A. Institutional analytics

This analysis is mainly used for decision making improvement at institute level. It uses variety of operational data for policy assessment analytics, instructional analysis and structural analytics. It makes use of reports to make timely data-driven decision across the entire institute and divisions.

B. IT analytics

It integrates data from different sources to monitor technological impact over the performance of the abstracted institutional process. It analyzes the data and identifies the need of technology at different level of institute process, required technological tools for betterment in process flow, required and suggested mechanism for information flow with the aid of Information Technology, etc.

C. Academic analytics

The essential part of data analytics in the education sector is academic analytics. Academic analytics covers different activities of higher education. Such activities may affect resource utilization and allocation, resource management, administrative task, research etc [8].

Mainly, academic analytics combines large datasets to improve decision making process and also established methods to compare with other institutions.

D. Learning analytics

In learning analytics, data about learners and its context are collected to understand and optimize learning environment. Learning analytics software are commonly used for improving processes and workflows, measuring academic and institutional data to improve organizational effectiveness [8].

V. LITERATURE REVIEW

Cristobal Romero and Sebastian Ventura elaborated various Education Data Mining workshops and conferences organized at different locations. He also listed some applications and tools with their intended goals[9].

Communication related to Big Data is merely superior in IT firms as compared to that with other organization like Education and Business Intelligence was stated by Samson Oluwaseun Fadiya and Emeka Joshua Chukwuemeka[1].

Faisal Kalota stated that different sources generate varied form of electronic data and their privacy issues are also important. These kind of data works much better with Big Data and Analytics [10].

Phelim Murnion and Markus Helfert described three level of complexity on a tentative model that extends the current existing academic analytics work[11].

Now-a-days students are forced more on process of learning but David Nicol in his article states the importance of formative assessment and feedback which can address a wide continuum like motivational, communicative and intellectual aspects of self-regulation[2].

Maximum used tools for new learning environments is expanded with a case study by Vidal Alonso and Olga Arranz in their paper which helps to set new strategies of teaching-learning[12].

Impact of multilingualism for a user to access a portal is well expanded with a survey on search query fired to reach a portal. It shows that use of English language is dominant as compared to native languages, which is a barrier for identifying user's behavior who use only their native language. Such users are prohibited from accessing the portal in their own native language thus, they could not help out themselves with educational materials of e-learning[13].

Judy Kay, Irena Koprinska and Kalina Yacef expounded group work among the students for developing any software. Using data mining technics, they categorized groups as strong and week groups[14].

Jin Soung Yoo and Moon Heum Cho used data mining techniques to mine concept maps to understand student learning patterns[15].

Educational data is hierarchical. Techniques for hierarchical data mining and longitudinal data modeling is the main expansions in mining educational data[16].

There are many benefits of Big Data and open data in educational sector. It helps parents and students to find the best educational program, transparent education financing and matching student and employment. Athanasios Drigas and Panagiotis Leliopoulos also briefed the usage of Data Analytics with Big Data in Industry [17].

VI. CONCLUSION

In this paper, we tried to showcase the application of big data analytics in the education sector. We have shown the different aspects of analytics that may be present in the education sector. Sources of data, nature of data, kind of process can be applied over the data, etc, have been identified and narrated in the article. A general idea to develop a framework to analyze educational data is also proposed in the article. Hence, we conclude that with the aids of such modern technology like big data analytics, the education system will be enriched with the new learning ways, making more efficient and targeted.

REFERENCES

- [1] D. S. S. E. J. C. Samson Oluwaseun Fadiya, "Big Data in Education; Future Technology Integration," THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE, pp. 65-69, 2014.
- [2] D. M.-D. David Nicol, "Formative Assessment and Self-Regulated Learning: A Model and Seven Principles of Good Feedback Practice," vol. 31, pp. 199-218, 2006.
- [3] J. Theroux, "Real-Time Case Method: Analysis of a Second Implementation," Journal of Education for Business, pp. 367-373, 2009.
- [4] S. L. a. I. Wilensky, "Mining Students' Inquiry Actions for Understanding of Complex Systems," Computers & Education, vol. 56, pp. 556-573, 2011.
- [5] J. S. B. M. a. K. K. Vincent Aleven, "Rapid Authoring of Intelligent Tutors for," in Human-Computer Interaction Institute, Pittsburgh, Pennsylvania: Carnegie Mellon, 2006.
- [6] S. Anirban, "Big Data Analytics in the Education Sector: Needs, Opportunities and Challenges," International Journal of Research in Computer and Communication Technology, vol. 3, no. 11, pp. 1425-1428, 2014.
- [7] P. B. D. D. G. O. John P. Campbell, "Open Learning Analytics: an integrated & modularized platform," EDUCAUSE Review, vol. 42, no. 4, p. 53-54, 2007.
- [8] B. Daniel, "Big Data and analytics in higher education: Opportunities and challenges," British Journal of Educational Technology, pp. 1-17, 2014.
- [9] S. V. Cristobal Romero, "Data Mining in Education," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, pp. 13-27, January 2013.
- [10] F. Kalota, "Applications of Big Data in Education," International Scholarly and Scientific Research & Innovation, vol. 9, no. 5, 2015.
- [11] M. H. Phelim Murnion, "Academic Analytics in quality assurance using organisational analytical capabilities," 2013.
- [12] O. A. Vidal Alonso, "Big Data & eLearning: A Binomial to the Future of the Knowledge Society," International Journal of Interactive Multimedia and Artificial Intelligence., vol. 3, 2016.
- [13] G. S. K. K. S. S.-A. Vassilis Protonotarios, "USING MULTILINGUAL ANALYTICS TO EXPLORE THE USAGE OF A LEARNING PORTAL IN DEVELOPING COUNTRIES," Journal of Asynchronous Learning Networks, vol. 17, no. 2, pp. 101-118.
- [14] I. K. K. Y. Judy Kay, "Educational Data Mining to Support Group Work in Software Development Projects," in University of Sydney, NSW 2006, , Australia, 2006.
- [15] M. Jin Soung Yoo, "Mining Concept Maps to Understand University Students' Learning".
- [16] M. F. M. F. Marie Bienkowski, Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief, Washington, D.C., 2012.
- [17] P. L. Athanasios S. Drigas, "The Use of Big Data in Education," IJCSI International Journal of Computer Science Issues, vol. 11, no. 5, pp. 58-63, 2014.

AUTHOR PROFILE

Dr. Jaimin N Undavia possess 14 years of teaching experience and awarded doctorate in the filed of Data Mining by Charotar University of Science & Technology.

Sheenal Patel She receivedher Bachelors and Masters degree in Computer Applications from Dharmsinh Desai University, Nadiad,Gujarat, India in 2012 and 2014 respectively. She ispresently working as an Assistant Professor at Smt.Chandaben Mohanbhai Patel Institute of ComputerApplications, CHARUSAT University, Changa, Gujarat,India since 2014. Her research area includeknowledge processing, Data mining, Big Data.

Dr. Atul Patel possess 20+ years of teaching experience at higher level and awarded doctorate in the field of networking. Currently he is serving CHARUSAT as a Dean of Faculty of Computer Science & Applications.