# An Overview on Role of Big Data Analytics in Higher Education

[1]K.Keerthana, [2]P.Yamini Devi

[1,2]Vignana Bharathi Institute of Technology,  Vignana  Bharathi Institute of Technology
Email: [1]Kaumudi55@gmail.com, [2]yamini1220@gmail.com

**ABSTRACT: A New technology which draws the insight of researchers to focus both in science and in industry is big data and analytics. This paper discusses how big data helps in transforming education system and discuss the role of big data analytics in present higher education system. Specifically this study will presents a survey report on role of big data and learning analytics in higher education system. In current educational system, Information is becoming more readily accessible across a wide range of interconnected systems. Even though the data driven decision making is popularized in 19th century, is transforming in to vastly more sophisticated concept known as big data and analytics. This paper proposes some rules for framing     an effective Adaptive learning environment design.**

**Keywords: Adaptive learning system(ALS), big data and analytics,**

## I. INTRODUCTION:

Day by Day many higher educational institutions are evolving and operating in an increasingly complex and competitive environment, every educational institution is adopting different learning environments but still suffering from several limitations in student evaluation scheme .Many institutions fail to make efficient use of huge amount of data available.   According to Ben Danien they are many contemporary challenges higher education institutions are facing worldwide and how the potential of big data and analytics helps in addressing these challenges [1].

### 1.1 What is Bigdata analytics?

Big data analytics is a technique to operate on datasets really about two things big data and data analytics. Analytics helps to discover what has changed and how one should react .Most of the organizations are confused about selection of right form of analytics for their Big Data analysis. Though we have related experience in data warehousing, reporting and online analytic processing (OLAP), we find that business and technical requirements are different for advanced form of analytics.

Big data focus on the size of the data in storage that is data volume along with its other attributes that are   data variety, data velocity and data veracity, Thus Big Data View is driven by these four variables.

**Data volume** is the primary attribute of big data that define big data in terabytes or sometimes peta bytes to show how much data is involved.

**Data Variety** defines different types of data such as structured, unstructured and semi structured data that is used by various applications such as credit files, utility records, or even face book posts[9].

**Data Velocity** implies fast and frequent processing of data by applications as one can't waste time for massaging or cleansing data that is needed for time sensitive, real time processes.

**Data Veracity** focuses on accuracy of data that is how far the data is trust worthy for usage by business applications.

### 1.2 Role of Big Data Analytics in higher education

In his paper he identified various factors affecting current education system. Finally he outlined a number of opportunities and challenges associated with the implementation of big data in context of higher education , and concluded  by presenting future directions relating to the built out and implementation of an institutional project on big data. In his paper he described predictive models has to be created to examine students at risk by performing retrospective analysis of student data and provide appropriate intervention, hence enabling instructors to adapt this teaching or initiate tutoring ,tailored assignments and continuous assessment. He also examined the role of big data and analytics in resolving the challenges facing by current higher education system. He outlined the key global factors [fig1] affecting the higher educational institutions and how the potential of big data and analytics explored in addressing these changing trends [1].
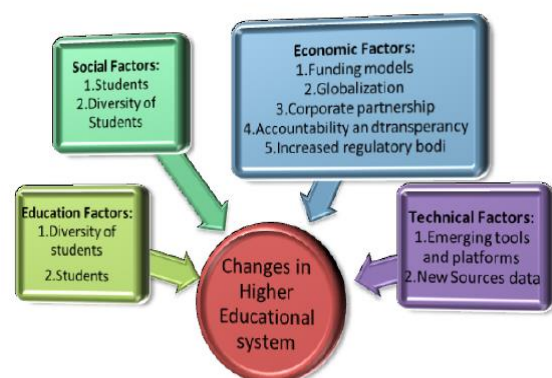


**Fig.1**

### 1.3 What kinds of data the institutions have long traditionally collected and stored?

Tracking grades, attendance textbook purchase, test scores, etc .But this huge data is not atomized to enhance students learning and to predict student's success in course completion.

According to education theorist George siemens[2] if universities collect huge data about the learners in their institutions drawn from university applications i.e. ; location ,previous learning activities, health concerns (physical and emotional, mental),attendance ,grades parental income, parental status and so on . Analyzing this data could help in capturing the profile of potential at -risk students [2].

Every institution has right to capture how learners interact with content and the discourse they have around learning materials as well as the social networks they form in the learning process to analyze whether the student move towards or away from course completion .Many tools are available such as SNapp ,uses the significant data such has time spent resource, frequency of posting, number of logins etc. captured by LMS like module or desire 2 learn to analyze social networks, degrees of connectivity, and peripheral learners and evaluate the qualitative attributes of discourse and discussions and rate each learners contributions by depth and substance in relation to selected course .If learners have access to the same dash board seen by educators and institutions for analytics, analytics can be a powerful tool in learner motivation [2].The foundational tool for informed change in education is evolved as learning analytics. Once we better understand the learning process -the inputs, the outputs, the factors that contribute to learner's success we can start to make informed decisions that are supported by evidence. Unfortunately, we don't really know which academic practices need to be curbing which need to be encouraged. Use of learning analytics can provide valuable insight into the factors that influence learners success (time on task , attendance, frequency of logins, position within a social network ,frequency of contact with faculty members or teachers access to library resources and learning help services.

According to simen[2] analytics should also focus on softer elements of learning, such as motivating encouragement from a teacher and the value of informal social iterations so that it can provide dramatic, structural change in education.

## II. RELATED WORK

The current teaching methodology in educational include creating and providing learning content in advance to learners who are taking the course in the form of curriculum like textbooks. This process is terribly inefficient since each learner has differing levels of knowledge when they start a course and one course may not fulfill the need for 60 learners hence each learner should have his/her own course based on their

life experiences, learning pace, familiarity in the course content by learner the topic should be as adaptive, flexible and continually updated. In the evaluation process of learners assessment should be done in each phase, not at the conclusion of a course in the form of an exam or a test, in order to make it possible we need to develop semantically defined learning materials and ways to compare learner –produced artifacts(in disc using, papers, texts)to the knowledge structure of a field automatically the outcome of our knowledge profile will reflect how we compare to the knowledge architecture of a domain such as let us consider an example" you are 64% on your way to being a radiologist" or "you are 38% on your way to being a diabetologist ".Evaluation should be done basically on a complete profile of an individual, not only on how narrowly he destined subject area but also on non-school-related learning (prior learning assessment) should be considered under learner program of study, Since a student who volunteers with a local charity or play sports outside of institution is acquiring skills and knowledge that is currently ignored by the current education system. In order to move beyond the nano –focus of exams" whole person analytics "is required, so that for students who return to university mid –career to gain addition qualifications can be analyzed to provide early indications which students are at risk for dropping out and these students can be offered early interventions by institutions to reduce dropouts dramatically.

Finally learning analytics if adopted can serve as a foundation for informed change in education, altering how universities and institutions create curriculum, deliver it, assess student learning, provide learning support , and even allocate resources.

### 2.1 What technologies are much focused for performing learning analytics?

In both the innovations in technology space and research in university research labs these is an exciting cross –over point between them ,both technical and social innovations in learning offers huge potential for a better , more effective learning model. Language recognition, neural networks, artificial intelligence, machine learning & selected concept are being combined with the growth of social network services, collaborative learning, and participatory pedagogy. George Siemens, Stephen downers and Dave cornier have experimented with "massive open online courses "and has resulted in software developed to encourage distributed learning .Tools like open study take an approach of decentralized learning and centralized analytics[2].

The latest analytical tools like R, R shiny[5] are used for performing student performance prediction, preventing risk retention and Graphical visualization of collected data and its analysis are also used for Unlike traditional learning system where syllabus is fixed for a course before they take up learning analytics helps to create personalized adaptive learning platform i.e

_____
ISSN(Online): 2349-9338, ISSN(Print): 2349-932X Volume -4, Issue -2-3, 2017

15

grockit and knowton. Many learning management system providers (such as desire 2 learn and blackboard) are actively building analytics options in this offerings.

An integrated learning and knowledge model is required where the learning content is adaptive, prior learning is included in assessment, & learning resources are provided in various contexts in order to have a broad impact in education by learning analytics. The non pre-planned content for all learners, profile of learner needs to be considered to drive curriculum & learning opportunities.

## III. METHODOLOGY

Educational Adaptive Learning technique involves the $W^3H$(who, where, why) concept for adaptation[3] which covers placement of adaptive system, goals adaptation, factors, methods and techniques of adaptation.

They are 2 major approaches micro and macro adaptation to reach e-learning systems adaptability.

Micro adaption leads to personalized learning where teaching approaches vary according to student's subject knowledge and interest.

Macro adaption enables instructors to derive type of learning environment and instruction that best suits to individual learners.

### 3.1 What is Adaptive Learning System (ALS)

According to sarbani mukherjee [4] an ALS is an ADAPTIVE LEARNING SYSTEM that is quite simple and easy t et adapted to the needs of individual learners. Unlike conventional teaching methodology where course curriculum is set prior to take up course by individual irrespective of his/her subject knowledge and interest, use to train the students not focusing much on the factors of their attrition and retention in course completion, ALS focuses much on student needs/styles, co focuses much on student needs/styles, competence levels, subject prerequisite to setup course curriculum and provides easy path for learning by assessing student progress parallel throughout the course training program.

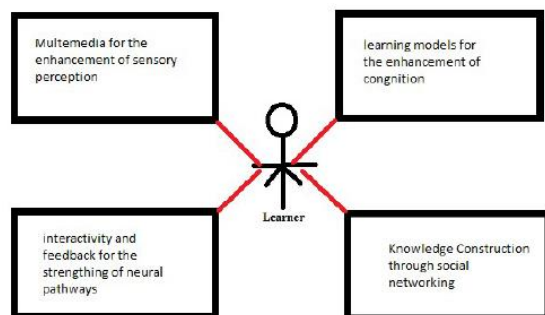The parameters considered in constructing a learning path under ALS are



**Fig:2**

### The 10 golden rules for framing an effective ALS environment design are:

**"**Same teaching style does not fit to all learners of a course. Providing information alone is not education .Teaching requires a cognitive process to be adapted for learners. Teaching styles need to be varying with each student and content: A student may be deductive learner in science, but have different style in social. Avoid forcing rigid timeline slots on learners for online course completion. Instructor teaching style should be individualized for each student to maintain student retention in course completion. Need to consider multimedia content and cognitive strategies for adapting ALS.Tutoring style should improve learner's competence for reaching effective Assessment: This is possible by replacing the traditional "Student punishing system" with "Student rewarding system". Students should follow appropriate strategy to reach current learning competency. Learners should be motivated with content but not with score". The reference table [10] comprises the **lists of top 99 universities in United Countries**. This Well developed countries has millions of students all over the world who aspire to study in good universities but the demand outnumbers the supply (especially of quality universities) [8].Hence the top 99 universities mentioned for the year 2012 are given national rank based on their score calculated according to following attributes :

**"Teaching ", Learning and Resources," "Research and Professional Practices," "Graduation Outcomes," "Outreach and Inclusivity," and "Perception".** Data visualization plays a key role in making human brain to grasp and understand the image of analytics results. Hence scatter plot fig.3 and box plot fig 4 depicted using R language shows the affect of each parameter in ranking the universities. We have taken a dataset Named as "College Rank.csv" from kaggle.com [10] for analysis purpose to show the impact of secured score of colleges on their National-rank. The college data set consist of 99 records which include "RANK", "COLLEGE", "COUNTRY", "NATIONAL_RANK", "PATENTS", "SCORE", "YEAR" as Table Attributes. They are 99 colleges from various united countries such as USA, UK, Japan, Switzerland, Israel, Canada, France, Sweden, South coria, Italy etc. The status of each college is graphical represented against other colleges located in United Countries based on their respective score and national rank in the year 2012.

**Scatter plot represents status of colleges based on the National-Rank assigned according to their respective Score**

In The Graph [fig.3] X-Axis depicts Score secured by colleges and Y-Axis depicts National-Rank of each college among listed 99colleges from various countries .The generated scatter plot from R depicts that many University colleges from various united countries whose students have taken average patents are assigned 24-39 rank based on their score in 2012 colleges .
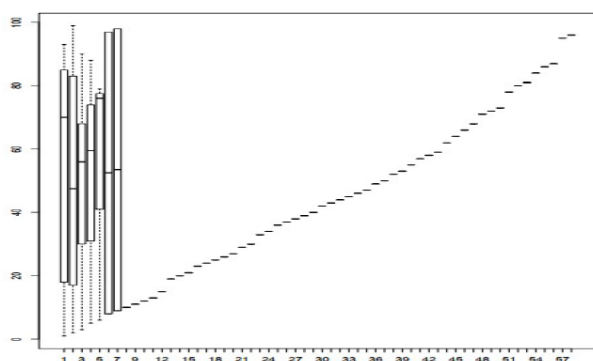
_____
ISSN(Online): 2349-9338, ISSN(Print): 2349-932X Volume -4, Issue -2-3, 2017

16

_____



**Fig.3**

The Scatter Plot [fig.4] shows that they are very less colleges whose rank is from 1 to 10 in the year 2012.
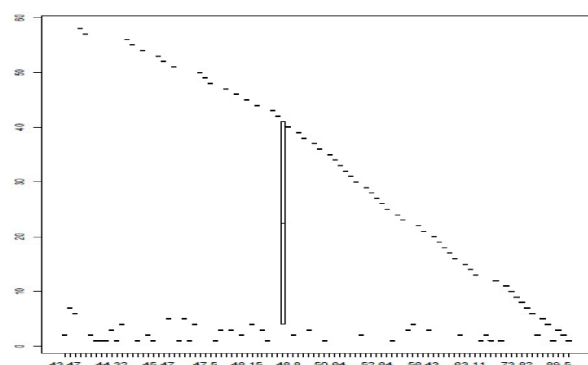


**Fig.4**

**Box plot representation of college status based on the National-Rank assigned according to their respective Score.**

The Box Plot [fig.5] shows the distribution of ranks across various universities based on their student's patents. There are nearly 50 colleges whose score is between 40 and 55.There is only one college Harvard University from USA whose score is 100 secured National-Rank
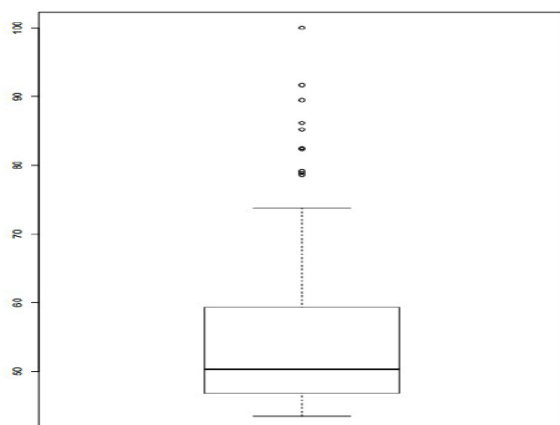


**Fig.5**

## IV. LITERATURE REVIEW

### 4.1 Major obstacles faced by education data and analytics?

The critical issue to handle in using learning analytics is privacy for learners and teachers .On other hand analytics serves as a means to improve learner success but it is used to evaluate & criticize the teachers performance. Data access and ownership are symmetrically important issues like who should be able to see the analysis that institutions perform on learners? And other concerns relate to error-correction in analytics. If educators relay heavily on analytics effort should be devoted to evaluate the analytics models and understanding in which contexts those analytics are not valid. With respect to the adoption of learning analytics, this is an exceptionally practical time to explore analytics, the complex challenges that universities & private institutions are facing currently can at least be illuminated partially through analytics applications.

Some of the PERFORMANCE PREDICTION techniques using big data analytics mentioned by authors are

| AUTHOR | TITLE | PERFORMANCE PREDICTION TECHNIQUES IMPLEMANT |
|---|---|---|
| (1) Darrell M. was author of new brookings institution book | Digital schools how technology can transform education[7]. | Tracking performance through Dash boards and visual Displays |
| (2) Vincek ellen , cutter consortium adam reck ten ward and Stephen burr | "Applying big data in higher education: a case study " cutter consortium,VOL1 3 ,2013[8] | Usage of bigdata analysis tool:"SAP's HANA" provides K-score for each student depicts the involvement of students in learning activites. |
| (3) Michel cottam, associate dean for instructional design at Rio salado | The evolution of big data and learning analytics in American higher education | PACE progress and course engagement system is an analytics application used for automated tracking of student progress |
| (4) Sarbani mukharjee principal learning designes | Adaptive learning systems: a tool to personalize learning TATA Interactive systems | Paradigm a drift from teacher centric model to student centric model |
| (5) Ben K.Daniel, | Big Data and analytics in higher education opportunities and challenges" Article in British Journal of Educational Technology · September 2015. | retrospective analysis for performance prediction |

| | | |
|---|---|---|
| (6) Ali Daud,Naif Radi Aljohani, Rabeeh Ayaz Abbasi, Farhat Abbas | "Predicting Student Performance Using Advanced Learning Snalytics"[6] | A Predictive model called "CHAID" is developed to predict students performance |

## V. SUMMARY

In this paper Introduction focuses on how bigdata analytics is used for addressing contemporary challenges faced by higher educational institutions worldwide. Related work focuses on various technologies used for performing learning analytics such as R,R shiny. It also focuses on methodology such as adaptive learning system and rules for framing an effective ALS design. We have taken a dataset from kaggle.com to show how R is used for Data analytics in comparing higher educational institutions based on their teaching methodology. Finally we covered various PERFORMANCE PREDICTION techniques used in BigData learning analytics mentioned by various authors in Literature Review. We are planning to design a new framework for OFFLINE Course performance prediction of students based on their skill set and personal history so that we can prevent drop outs and promote successful course completion of students by cultivating learning interest in them.

## REFERENCES

[1] Ben K.Daniel,"Big Data and analytics in higher education opportunities and challenges", Article in British Journal of Educational Technology · September 2015.

[2] Audrey Watters, "George Siemens on the applications and challenges of education data". July 25, 2011, https://www.oreilly.com/ideas/education-data-analytics-learning.

[3] J.Liebowitz , "Thoughts on recent trends and future research perspective in big data and analytics in higher education", Springer International Publishing Switzerland 2016.

[4] Sarbani Mukharjee, "Adaptive Learning Systems: A Tool to Personalize Learning", Tata Interactive Systems 2013

[5] https://www.r-project.org/about.html

[6] Ali Daud,Naif Radi Aljohani, Rabeeh Ayaz Abbasi, Farhat Abbas," Predicting Student Performance Using Advanced Learning Snalytics", International World wide web conference vommittee (IW3C2),

[7] Darrell M.,"Big Data for Education: Data mining, data analytics, and web dash boards", Executive Summary.

[8] Vincekellen , cutter consortium adam reck ten ward and Stephen burr, "Applying big data in higher education: a case study " cutter consortium ,VOL13 ,2013.

[9] Michel cottam, associate dean for instructional design at Rio salado, "The evolution of big data and learning analytics in American higher education", Journal of Asynchronous Learning Networks, Volume 16:issue 3

[10] www.kaggle.com.

❖ ❖ ❖