Big Data Architecture Framework and Application Tools for Education Institutes

Paduri Veerabhadram*

*vpaduri@gmail.com,   *vpaduri@nust.na,

Faculty of Computing and Informatics

Department of Informatics

Namibia University of Science and Technology

Dr. Nalina Suresh**

nalina.kss@gmail.com*, *nsuresh@unam.na,

Faculty of Science, School of Computing

Department of Information Technology

University of Namibia

Abstract— This paper is a study of the utilization and analyses of huge information on education. However, the massive information and open information technology will actually involve to education sector also. Moreover huge amounts of unused information will profit and improve education system. Academic Institutions of higher education are operating in an increasingly complex and competitive environment. This paper identifies many challenges faced in institutions of higher education worldwide and explores the potential of big data in addressing these challenges.This paper explore the variety of opportunities and challenges related to the analysis and implementation of huge information within the context of upper and higher education.

Index Terms— Data Mining, Big Data, Education, Open Data, Cloud Computing, BI, SPSS, smart university, sharing knowledge

———————————— ◆ ————————————

INTRODUCTION

BIG data is a term that describes the large volume of data both structured and unstructured that inundates a business on a day-to-day basis. A big data designed to integrate data from various source to a single platform and structure. Ubiquitous social networking, and pervasive mobile devices and the rise of big data has created an inflection point making real-time data collection and analysis mission critical for businesses today. As an IT solution, big data mirror the growth in content and data source, as well as the pervasiveness of technology in our everyday lives. As more and more of what we do is both connected to and often empowered by a network and the devices that we connect to be themselves powered by an array of sensors. It is expected that the ongoing stream of data will grow. Within data centres, every node (servers, storage, and applications) generates a tremendous number of log files and isolated data streams that also can be collected, collated, and analysed. With storage costs dropping, the cost associated with saving and leveraging even the most mundane data becomes a nonissue. In this study we are going to see the employment of huge skill and open data in education. Conjointly, however huge quantity of information may be used and extracted to one thing helpful, serving to the industry to lift their profits. Thus a definition of huge data and open skill, explain, however that these two technologies are a unit classified. After that it's vital to say the goals and purposes of huge skill in education, giving a transparent image of the value and effects of huge skill in education. Big Data has also been defined by the four(4) "V"s: Volume, Velocity, Variety, and Value. These become a reasonable test to determine whether you should add Big Data to current information architecture. A simple Big Data model of 4 V's is show in figure 1 and described as below.
Volume: Comes in one size: large. Organizations are awash

with data, easily amassing hundreds of terabytes and petabytes of information.
Value: is the most important aspect of big data it refers to the process of discovering huge hidden values from large data sets with various types and rapid generation.
Variety: Extends beyond structured data and includes semi-structure or unstructured data of all varieties, such as text, audio, video, click streams, log files, and more.
Velocity: Sometimes must be analyzed in real time as it is streamed to an organization to maximize the data's business value. It means how fast data is being produced and how fast the data must be processed to meet requirements or demands.
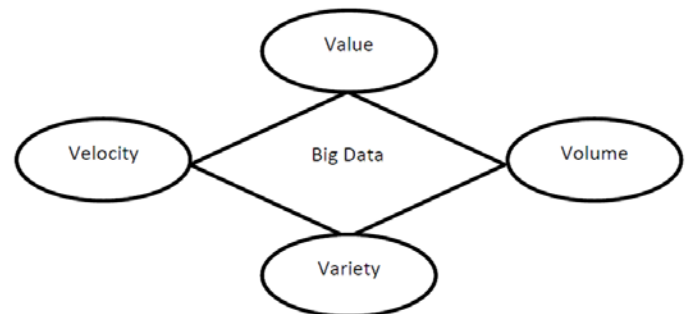

Fig 1: Big Data Model [1] [2]

Big data is a term utilized to refer to the increase in the volume of data that are difficult to store, process, and analyze ex-pand. The benefit of gathering large amounts of data includes the creation of hidden information and patterns through data analysis. Laurila et al. [3] provided a unique collection of longitudinal data from smart mobile devices and made this col-

lection available to the research community.

Objectives:
Big Data are becoming a new technology focus both in science and in industry and motivate a technology shift to data centric architecture and operational models. There is a vital need to define the basic information/semantic models, architecture, components and operational models that together comprise a so-called Big Data Ecosystem.

African Universities are still left behind in terms of University world ranking due to several short comings including the lack of technology implementation which includes big data and offering big data courses. Namibia University of Science and Technology (NUST) in Namibia take the initiatives to implement big data program and becoming a smart university.

2 BACKGROUND

We are living in the era of data. Everywhere we can get large amount of data. Data is generated almost from every sector. It can be Aviation sector, sports, social media and also in the education sector. The primary focus of this paper is to analyze how the big data is generated in the education sector and also how those data can be managed and how to apply predictive modeling to best use of the dataset. Big data-simply means large unstructured data. Education sector now a days started to become technology oriented. There are lot of Massive open online courses (MOOC) which are generating huge amount of data.

According to Mobile Cloud Application Design Process Model for Education believe Cloud computing will surely improve the current system of education and improve quality at an affordable cost [4]. Mobile Cloud computing application helps to increase the storage space in the Cloud service. There are more applications for teaching and learning as opposed to a single platform, independent tools and scalable data storage and pay-as-you-use.

Hence, Technologies such as Data mining and Data analytics can provide a fast feedback to students and teachers about their academic performance. These methods can provide a deep analysis of some education patterns and extract valuable Skill from them[5] .In this way, collective and big scale data can predict who student needs more help from the education system, avoiding the danger of failure or drop out [6] .Big Data can provide more opportunities for new learning experience for children and young adults. Hence, students can share information with educational institutions in this way they can expand their skill. Furthermore, Educational institutes and Universities are able to help and prepare their future students[7] .

Table 1: Benefits of Big Data and Open Data in Education [8]

Improved instruction	Can improve students'performance and learning abilities, making the lessons more personal. The courses can be adjusted from the teachers with the help of analytics.
Matching students to programs	Open Data are able to help parents and students to find the best school or educational program.
Matching students to employment	Companies and candidate staff can discover alternative and more effective tools to use open data to qualify their skills with the needed skills. Also students can find and make applications for jobs which can match with their abilities, more efficient than before
Transparent education financing	This leaves to students to participate in education activities, which previously they don't have the ability. Furthermore are able to choose anything about higher education and to discover the most proper education programs for them.
Efficient system administration	School education systems are able to develop a skillful school supply which can help administrators to allow more affective education resources. In that way this secures a high performance and afford to a versatile and smart plan for future education interests

3 PROPOSED BIG DATA FLOW ARCHITECTURE FOR EDUCATION INSTITUTES.

The data flow is shown in figure 2 through a box and arrows diagram with a content box on the top with an arrow to a student and two engines underneath showing as boxes: an adaptation engine and an intervention engine, with arrows for each up to the content box. Another arrow connects a predictive model box to the adaptation engine. The predictive model is connected to two databases with incoming arrows. On the right is the student learning database and on the left is the student information system. Below the predictive model and connected with an incoming arrow is a dashboard that is shown connected with arrows to faculty and educators and administrators

In addition to these six internal components, an adaptive learning system often uses the student information system (SIS) that is maintained by a university, college and school district, or institution as an external data source. Student profiles from the SIS are usually downloaded in batch mode, as they do not change often, and then are linked with performance data in the student learning database using student identifiers in compliance with applicable law. Student profiles contain background information on students that can be used to group them into specific categories or to provide more variables that might suggest a particular student is at risk.

The Future proposed paper work intends to provide a consolidated view of the "big data "phenomena and related challenges to modern technologies and initiate wider discussion. By becoming a smart university, not only the university's ranking will rise, but also other aspects related to the university, such as increased profit, student enrollment, researching abilities and publishing and the university would also be able to collaborate and share and receive skill from other Universities[10].

In addition to these six internal components, an adaptive learning system often uses the student information system (SIS)

that is maintained by a university, college and school district, or institution as an external data source. Student profiles from the SIS are usually downloaded in batch mode, as they do not change often, and then are linked with performance data in the student learning database using student identifiers in compliance with applicable law. Student profiles contain background information on students that can be used to group them into specific categories or to provide more variables that might suggest a particular student is at risk.

The Future proposed paper work intends to provide a consolidated view of the "big data "phenomena and related challenges to modern technologies and initiate wider discussion.
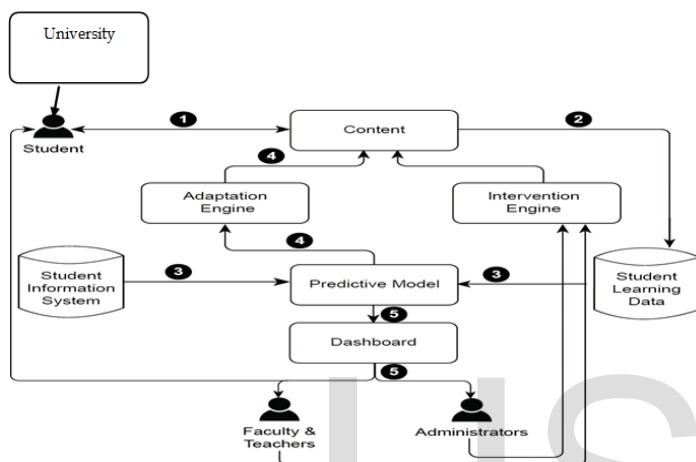

Figure 2. The Components and Data Flow through a Typical Adaptive Learning System [9]

4 RESEARCH METHODOLOGY

This is paper informative reviewed research paper, supported literature review, connected existing work and exploring potential use cases, targeted on the big data of a theoretical framework for style theories in Cloud computing for service, application, data storage, analytic, reduce map and data mining platform or infrastructure access. The authors embody the conclusion of the standard theoretical context framework for style theories in big data and Cloud computing literature theory. The value of big data and Cloud style model and application solutions aren't nonetheless specific, however wants more attention. Analysis ought to specialize in the relation between a context framework for style theories in big data and Cloud computing theory. From a user perspective, the disposition to share pooled resources needs more attention.

A quantitative analysis approach is typically utilized in studies meant to live lecturer and student perspective towards big data, specifically supported a theoretical framework for style theories in big data architecture. This study examines the factors that influence the adoption of a theoretical framework for style theories in big data usage in Cloud technology models for instructional functions among the scholars or researchers of the NUST in Namibia.

5 DISCUSSION AND CONCLUSION

However, we should be developing this infrastructure with an eye toward the horizon. Adopting technologies and underlying infrastructure, including networking that can provide the scale, performance, and headroom for tomorrow's technologies is critical in offering the highest levels of investment protection, business agility, and time to market.

As we are able to see from the previous research, huge skill will really improve the education. In addition, will afford to form a modern and dynamic education system, which each individual student will have the most enjoy that. Moreover, lecturers will have valuable tools, where they do not have before, which may build their selections on a lot of specific and area unit ready to opt for an enormous sort of new learning strategies. Hence the massive skill is literally concerned to vary the way of industries as well as the education. Within the new era of Data the normal difficulties are now does not exists, keeping the nice strategies. The education system is enriched with new learning ways in which, creating a lot of economical and targeted. But the approach of this new era, has simply begun and there are many difficulties like the shortage of tough personnel in the science of massive skill and Skilled analytics. What is more important is, the lecturers and teachers should really train and involved on them and at last the scholars should settle for and use these new tools.

MOOC becomes popular day by day and the amount of data generation increases day by day. So handling this huge amount of data will be a challenge now. For the other big data management researchers started using various new databases like NOSQL. For the education data also, lot of analysis and researches are going on to handle the data with the proper techniques and lot of researches are going on to invent new tools to generate real time solution, prediction in this particular sector.

As a conclusion, the integration of big data into University for a smarter University is definitely something to be looked into as it would provide a better future for both students and the Universities' staff. Besides increasing profit and the ranking for the University, it would also increase the stakeholders' satisfaction towards the services of the University. This project will ensure continuous innovation to the University, which may open ways to other innovations for a better University since this integration of big data and data warehouse will provide this ecosystem.

Challenges

Another important tool that is becoming popular that is called Virtual classroom. All the distance programs now days are based on it. In these tools the instructor can share his/her desktop, so all the slides he wants to share in the webinar class and all the demonstration he wants to show can easily be shown. There is always an option to raise the question from the students' side which creates the notification to the instructor. All these tools generate lot of data, data that can be stored, analyzed and on the basis of that forecasting can be done; but one thing should be kept in mind that learning analytics data are big, but not google big. The size can be like this. For a open source free online program almost 1,00,000 students have registered. In the

PSLC data shop, 2,50,000 hours of students are using educational softwares, more than 30 million students actively participate to access data. So all those data are really large

REFERENCES

[1]   Michalik, P., Stofa, J., & Zolotova, I. (2014, January). Concept definition for Big Data architecture in the education system. In Applied Machine Intelligence and Informatics (SAMI), 2014 IEEE 12th International Symposium on (pp. 331-334). IEEE.

[2]   Sagiroglu, S., & Sinanc, D. (2013, May). Big data: A review. In Collaboration Technologies and Systems (CTS), 2013 International Conference on (pp. 42-47). IEEE.

[3]   Laurila, J. K., Gatica-Perez, D., Aad, I., Blom, J., Bornet, O., Do, T. M. T.,. & Miettinen, M. (2013). From big smartphone data to worldwide research: The mobile data challenge. Pervasive and Mobile Computing, 9(6), 752-771.

[4]   Veerabhadram, P. (2016). Mobile Cloud Application Architecture in Education Institutes. International Journal of Advanced Studies in Computers, Science and Engineering, 5(3), 6. [5] West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. Governance Studies at Brookings, 4, 1-0.

[5]   West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. Governance Studies at Brookings, 4, 1-0.

[6]   Siemens, G., & Gasevic, D. (2012). Guest editorial-Learning and knowledge analytics. Educational Technology & Society, 15(3), 1-2.

[7]   United States. Executive Office of the President, & Podesta, J. (2014). Big data: Seizing opportunities, preserving values. White House, Executive Office of the President.

[8]   Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Doshi, E. A. (2013). Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute, 21.

[9]   Bridges, J. E., Bajzek, T. J., Hofer, K. E., Spencer, H. L., Smith, L. G., & Young, V. R. (1991). U.S. Patent No. 5,070,533. Washington, DC: U.S. Patent and Trademark Office.

[10]   Linoff, G. S., & Berry, M. J. (2011). Data mining techniques: for marketing, sales, and customer relationship management. John Wiley & Sons.