Learning Analytics and Big Data in Higher Education

Tulasi. B,
Assistant Professor,
Department of Computer Science,
Christ University,
Bangalore
India

Abstract

Learning analytics analyses both static and dynamic data extracted from teaching - learning environment to allow timely interventions by educator. It paves way for optimizing the learning and thus helps both educator and learner to achieve success. Learning analytics allows personalization of learning process .It lets the educator to cater to needs of individual learner to a greater extent. The voluminous data (termed as Big Data) available with educational institutions can help to enhance the effectiveness of teaching learning in many ways. Data driven decision making is the need of the hour .With the rapid increase in generation of digital data, the need to extract potential information effectively also arises. Big Data can prove to be an effective tool. This paper explores the area of learning analytics and big Data in higher education.

1. Introduction

In the current scenario of technical advancements, educators have potential to exploit technology for greater impact on learner achievement. Over past few years technology has been envisioned to pioneer transformation in the field of higher education .Though the effect of it is not profoundly seen, technology has been embedded in all streams of education. For over a decade universities have been using various systems like ERP, LMS and have generated massive amounts of information [1].With Internet and mobile technologies along with social change gaining newer heights, amount of data being captured is increasing exponentially. This abundant data which is available in different structural format is being termed as "Big Data". To extract useful insights from the voluminous data stored, it requires more than the compliance reporting. Institutional data can be effectively utilized to improve educational results. Analytics is a possible approach which can help higher education change the way it operates. Analytics when applied on large volumes of data available along with advanced technologies can facilitate learner's academic progress and add value to the institution.

Analytics provide statistical evaluation of rich data which can help educational institutions to make data or evidence based decisions. Colleges and universities can potentially use this feature to develop models which can identify learners who are under the risk of not succeeding academically. As in business sector, Big Data and technology advancement are motivating usage of predictive analytics in higher education. During the process of teaching learning learners leave trail of data which is digital in nature. This data can to be mined to optimize the teaching learning process. Though analytics is not the only solution, it is one of the key components which can lead to effective actionable data. Actionable data provides insights to both educator and learner to make the learning process more effective.

Learning is the result of interaction and discussions done by the learner with the educator, other learner's and course material. A learner then interprets these interactions and act based on the understanding of them. Educators involve themselves in generation of more learning opportunities to make the teaching learning process effective. During this process the questions on understandability of a learner, effectiveness of the course and many more similar arise. The educators usually get reports on these questions at the end of the course which is usually incomplete and delayed. Most of the content and resources are now moved online. This data provides avenue to analyze, design, and deliver content which would make the learning process more efficient. It would also be tailor fit for the learners.

Before learning can be improved, it has to be determined whether the learner is responding to the instruction. It is usually measured by qualitative approach which is educator's instincts. This approach often leads to incorrect conclusion. It has been observed that more quantitative and actionable data about classroom progress is beneficial to both the learner and the educator. Data available to educators is usually learner's average scores which needs to be further simplified to make it actionable. Also the results are mostly obtained after a topic is completed, hence would not be very effective. Digitized learning modules provide real time and systematic data which enables assessment of learner and feedback process immediate.

In this competitive world there is demand for skilled and resourceful citizens. This increases the expectation and accountability of educational institutions. On the other side educational institutions are finding it difficult to graduate learners who can meet the requirements. To identify the approach to be adopted, educators and administrators need deeper insight. Big Data and analytics can provide tools to measure learner performance and further help in acquiring the skills to succeed. Analytics in higher education can help in promoting academic success and measuring learning effectiveness

2. Analytics in Education

Analytics combines large data sets, statistical methods and predictive modelling to produce actionable data in educational institutions. [10] It focuses on building models which can predict appropriate activities to optimize learning.

Analytics in education has diverse and inter disciplinary face. An analytical project would have computer professionals, statisticians, graphic professional and psychologists as team members. It is relatively newer and also not associated with any particular domain.

2.1 Forms of analytics in education

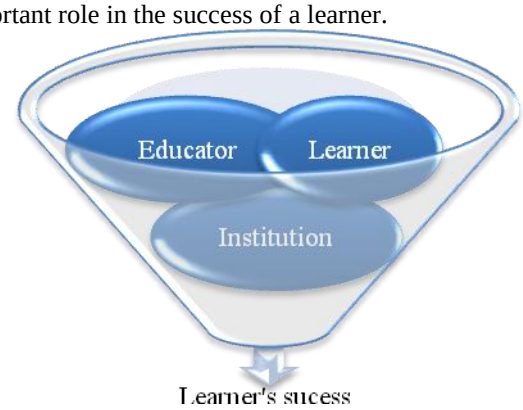
Educational institutions are utilizing two forms of analytics: academic analytics and learning analytics. [10]. Academic analytics is application of business intelligence tools and strategies for decision making in educational institutions. It reflects role of data analysis at an institutional level. Learning analytics is analysis and reporting of learner data for facilitating teaching learning process. It also emphasizes on use of data intelligently. Distinction between academic and learning analytics as given by Siemens and Long states that learning analytics is more specific than academic analytics. The former focuses on learning process which includes analyzing the relationship between learner, content, institution and educator. Academic analytics involves learners and institution while learning analytics is beneficial to all the stake holders: learner, educator and the institution. The following table provides a comparison of the two types of analytics

Table 2.1 Comparison of Academic and Learning Analytics

Academic Analytics	Learning Analytics
Academic analytics is application of business intelligence in higher education	Learning analytics is associated with the collection, measurement and analysis of data about learner.
It can be applied at three different levels: Institutional, Regional, National or International.	It can coupled at course and department level
Beneficiaries are administrators, funders educational authorities	Beneficiaries include learner and instructor
It would facilitate in comparison of different educational systems.	It would help the learner and instructor to understand the pattern of success or failure
It would assist administrators to ensure optimum utilization of resources.	It would facilitate in identification of an intelligent curriculum.

2.2 Learning analytics –A micro process

Learning analytics is a micro process which follows bottom up approach. It looks at short term goals like improving learner's learning process.[13] It does not uphold any one of the pedagogy rather is versatile in nature. The key components in learning analytics are educator, learner, content of the course and the institution. The educator needs to have a real time feedback regarding the teaching learning process to ensure that the process is helping the learner in his/her academic endeavors. All the components of learning analytics play an important role in the success of a learner.



Learning analytics is a powerful instrument which has the potential to make effective change in higher education. It can act as a decision foundation to provide insight into the data which influences learner's success. Analytics encompasses the learner information like attendance, frequency of logins, position in social network, time required for a task completion. Tangential data analysis could also include use of physical services and accessing of library resource. Though analytics tries to encompass all the factors it fails to take in things like mentoring done by educators, the influence of social interactions which might play an important role in learner's success.

Learning analytics has the potential to change the structure of higher education. It would lead to development of curriculum which would be dynamic to cater of the requirements of the varied set of learners. Each learner is unique and so is requirement of each learner. To provide an effective learning environment the course content should be adaptive, flexible and rationalized. This necessitates "intelligent curriculum" to be developed and utilized.

2.3 Stages involved in Learning Analytics

Learning analytics can be viewed as incremental and iterative approach which leads to maturity in implementation. The iterative approach helps to evolve a rationalized model.

The following are the stages involved: [9]

- **Collection and incorporation of data**

Data from different resources like social media, learning management system, library logs, and class interaction logs have latent information which can provide insights to both the educator and learner. Educational institutions collect diverse data and identifying the potential data is a vital task. This varied data could be incorporated in analytics matrix.

- **Develop model**

Models for individual learners can be developed .This approach focuses on personalization of models rather than applying a generic one for all learners. The model would pave way for further plan of action to be adopted by the educator and learner.

- **Application of analytics – action and adaptive**

Predictive analytics would be able to identify learners who could be categorized as "risk learners". This information would assist in designing activities to enhance the learner capability thus avoiding a failure. It would also identify the advanced learners who could be motivated to outperform .Effective activities could be designed to maintain the same level of learning interest. The automatic and adaptive analytics would take in the real time learner data which would help in making operative decisions .The model thus would be updated automatically on the availability of data .This prevents in making decisions made through outdated model.

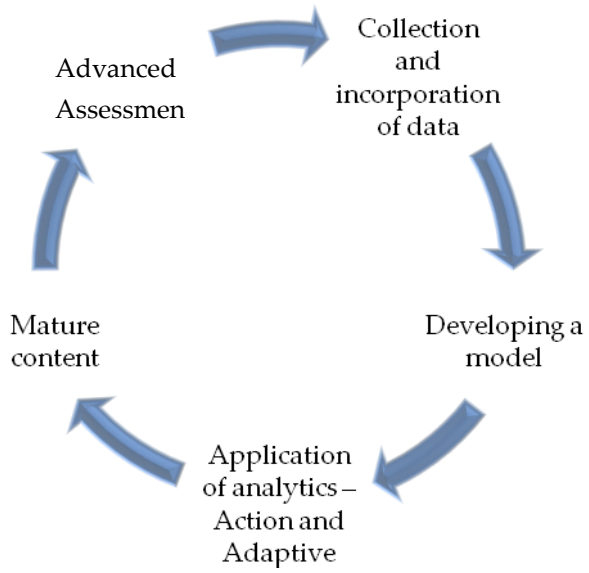
- **Mature content**

Defining the learning content semantically leads to intelligent curriculum. The benefits of intelligent curriculum are mammoth. It leads to personalized, tailored fit content satisfy the specific requirements of each learner. The pertinent and current data is available for the learners, which helps in in-depth learning. Intelligent curriculum would be self adaptable to the skill sets of learners. This would define smoother path towards the learning objective.

- **Advanced Assessment**

Grading learners taking into consideration their profile and the knowledge domain associated would be an effective methodology of assessment. These assessments would be dynamic in nature which would be reviewed and adapted periodically to make the teaching learning process more effective. Assessments both qualitative and

quantitative should clearly relate to the learning objective.



2.4 Benefit at different levels

Learning analytics can be applied at different levels: Individual learner or educator, organizational and at regional/national/international level. Due to the rich diversity and varied operational level of learning analytics, the impact of it on higher education is profound. At the course level learning analytics can be applied to:

- Understand teaching learning process.
- Identifying learners who are risk learners
- Obtain information of learning habits of learners and give recommendations for improvement
- Measure learner learning against the learning objective

The department level learning analytics would help decision makers in [8]

- Identifying the best instructional practice.
- Improve existing instructional practices.
- Identifying key attributes of a good teacher.
- To set standards for teachers performance.
- Measure and monitor teaching progress.
- Teachers can also be evaluated with help of attributes identified.
- Evolve curriculum to fit right of a particular group of learners.

The impact of learning analytics on the mentioned two levels could lead to transformation at the institutional level .This could evolve new academic models and pedagogical approaches further impacting the global scenario.

3. Big Data

Data has always played a pivotal role in decision making. With the advent of technology, data is being created in an exponential manner. There is a surge of data seen in every area globally.[6] Digital data is available in every segments of an organization, educational institutions being no exception. Due to availability and sharing of data over the digital network, "Exhaust Digital data" is being churned in tremendous volumes. [3].The data is termed as exhaust as it is a by-product of other activities .Social networking, smart phones are few means through which this voluminous data is being created .This voluminous data is termed as "Big Data" as these datasets are beyond the ability of a normal database tool to capture, store, manage and analyze. The definition of Big Data can vary from sector to sector depending on the tools used and the average size of data sets associated with the sector. The current growth rate of data collected is overwhelming. It is a major challenge to handle and analyze the data effectively.

3.1 Characteristics of Big Data

Big data is defined as three dimensional volume, variety and velocity. It has been defined as "New generation of technologies and architecture designed to extract value for large datasets of wide variety by high velocity capture, discovery and analysis. [2]

Volume

Data is being generated in terms of zeta bytes. The volume of the data is the data generated or available to an organization.

Velocity

The rate of creation of data is termed as velocity of data. It is seen that data is being generated at exponential rate. The digital being generated is enormous and is overwhelming.

Variety

This is a measure of varied data representations like text, audio, video, images .These varied data are structured, semi structured and unstructured forms of data.

Apart from the above mentioned three dimensions there have been other dimensions like of "value", "variability" being considered by different data scientists.

Variability

Data apart from being varied, voluminous is also highly inconsistent in flow. Data flow can be event triggered, seasonal among other reasons. These variable data loads are challenging .Unstructured and variable load of data makes the task of analyzing more complex.

Value

This" fifth V "is the key element which understands the behavior or pattern generated by the data. This paves way for a predictive model generation.

3.2 Big Data – Value creation

Big Data creates value within and across disciplines. This comes from the actionable data obtained after the analysis of the data. The following points indicate five ways in which value creation is done by Big data. [3]

i. Transparency

Transparency of data to all stakeholders in an educational institution i.e. learners, management and educators at the appropriate time would lead to effective value creation. When relevant data is available across the dissection of departments, it facilitates in improving the quality of the processes involved in educational institutions.

ii. Discovering varied needs of learner to improve performance

With the availability of data in digital form data collection would be more accurate or real. This data further facilitates the decision makers, educators to discover different needs of learners and industry. The varied needs of learners when catered, lead to improvement of performance.

iii. Customized action

Big Data allows designing and implementing processes for specific requirement of learners. This provides a tailor fit solution to individual learners which in turn paves way to performance enhancement. This approach leads to refinement and evolving of teaching learning process.

iv. Decisions with support of data

Decision making when backed by data proves to create greater value. With Big Data the decision makers analyze the real data which leads to better decisions. The risks involved during the procession of decision making and implementing are also minimized.

v. Innovating new models

Big Data enables to create new methodologies in teaching learning process. Educators would have accesses to data which would allow them to refine their teaching paradigms .Identification of better approaches and refinement of them could lead to new paradigms and pedagogy.

4. Implementation Examples/Breadth of Learning Analytics

The field of learning analytics is quite extensive. Various applications have emerged encompassing the wide spectrum of learning analytics. The following examples try to highlight the expansive nature of learning analytics.

Analytical dashboards

The impact of BI products on learning domain has given rise to the analytics dashboards. This kind of analytics is now seen profoundly in online learning platforms. Visualization of data through graphs, tables which is easily interpretable is accessible to educators, learners and administrators. Simplification of instructional content planning and delivery, real time assessment are the key take away for educators Quiz scores, forum discussions provide learners a understanding about their relative performance. Examples for this are LMS/CMS systems like moodle and "Insight 360º".

Predictive analytics

Data associated with students can be identified as different types. It can be static (like previous academic performance) and dynamic (like discussions on forums, frequency of login). [1]. This data about learners is voluminous, varied, unstructured and the speed about which it is being created is quite challenging. Studying this data can lead to classification of learners under different categories like advanced learners, slow learners. Identification of variables on which the prediction is done is a complicated procedure. It requires extensive statistical analysis. These variables would be validated using previous history to ensure that they are strong predictors. Though the most common used parameter is the test score from the beginning of a course, there is a need to evolve more complex predictive model.

Course signals a project of Purdue University is one of the applications of predictive analysis in teaching learning. It uses student's grades, past academic performance, student's efforts with help of the Blackboard Vista (LMS of Purdue University).The premise is to analyze the ample data available with educational institutions. The algorithm which predicts the learner's success considers four parameters performance, learner's effort, prior academic performance and learner's characteristics like age. Each parameter is assigned a weight and is computed to obtain the final result. [5]

These predictive models usually are context based like for a particular course, institution or culture. The need to generalize these models is singular. To obtain a generic model, common parameters need to be identified among the learner's data available across different educational institutions.

Adaptive Learning analytics

Fine grained feedback always provides good insight of learner's understandability over a generic feedback. Adaptive analytics depends on understanding the content and not on superficial characteristics. This requires the data to be tagged for analytics. Though complex by nature, it is quite useful to both learner and educator. Development of model for learner cognition is quite intense and the results more personalized.

Examples are the intelligent tutoring systems. Significant research and progress has been done on Intelligent tutoring systems. Tutoring systems are said to have two loops. Outer loop deals with a task which is usually consisting of few complex steps. The inner loop would assess the learner's evolving competence and update the student model which would suggest next task to be given by the outer loop.SQL tutor, Algebra cognitive tutor are few intelligent tutor systems. These systems are used as a element in a course. [4]

Social Network Analysis

Social network analytics provides details of interpersonal relationship of a learner. It provides imperative information about the connections of a learner. This might lead to deeper understanding of the effective learning of the learner. This process is time consuming and exhaustive.

A wide range of tools have been developed to mine the essence of social network. SNAPP allows rendering forum discussions, posts as a network to help in identifying disconnected students and visualizing educator effort in the network .SNAPP also allows educator to evaluate learner behavior against learning objective and devise early interventions.[11]. With the leap seen in use of social networking the amount of data created is also quite massive.

5. Issues and concerns

Learning Analytics also has its share of issues and concerns. The privacy of data is one the major concerns in learning analytics. Deciding the ownership of the data collected is a tricky concern. The analysis leading to negative psychological impact rather than motivating the learners towards better learning is a vital issue which needs to be dealt with care. The effectiveness of the decision making on the pertinent data is under cloud.[7] The result of the analysis to be provided to the learner is a question to be considered. Above all the capability of the educator to provide required interventions is considered as the vital requirement of the success of the entire process.

6. Conclusion

Technology is transforming higher education. Learning analytics together with Big Data can prove to be a good catalyst for the change. Education without analytics would prove to be inefficient as it would not be able to evolve itself to the needs of 21st century, the century whose backbone is Internet and social networking. The pedagogical changes required in higher education to cater to the current tech savvy learners might not materialize without help of analytics. Also teaching learning process is on threshold of evolution. With Big Data and automated analytics, the learner can be studied at a magnitude which was not possible earlier. This would empower educator to bring interventions which can lead to an effective teaching learning process. The impact of this process would also been seen as fruition of a new innovative pedagogy for teaching learning .Although this is a new area of research , range of tools and methods have evolved. The challenge is to evolve learning as a data science, working with wider data or data sets and other issues which would emerge.

References

[1]. Penetrating the fog: Analytics in learning and education, Phil Long and Gorge Siemens, *EDUCAUSE Review* vol 46,no 5 September/October 2011.

[2]. Gantz.J,and E.Reinsel,2011,"Extracting value form chaos",IDC's digital universe study,sponsored by EMC.

[3]. Big Data:The next frontier for innovation, competition and productivity.James Maniyka, Executive summary ,McKinsey Global Institute ,May 2011, http://www.mckinsey.com/mgi/publication/big.data/MGI_big_data_exec_summary.pdf>.

[4]. Kurt VanLehn, LRDC, University of Pittsburgh, Pittsburgh, PA, USA,The Behavior of Tutoring Systems, *International journal of Artificial Intelligence in Education*,volume 16 Issue 3,August-2006.

[5]. Course Signals at Purdue: Using Learning Analytics to Increase Student Success Kimberly E. Arnold Purdue University ,Matthew D. Pistilli Purdue University.

[6]. Abele, Eberhard, Tobias Meyer, Ulrich Näher, Gernot Strube, and Richard Sykes, eds., Global production: A handbook for strategy and implementation (Berlin: Springer, 2008).

[7]. Ferguson, Rebecca (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6) pp. 304–317.

[8]. IBM white paper Analytics for Achievement Understand success and boost performance in primary and secondary education, 2011

[9]. https://learn.canvas.net/courses/33/wiki/week-5-smarter-curriculum

[10]. Oblinger, D. G. and Campbell, J. P. (2007). Academic Analytics, EDUCAUSE White Paper. http://www.educause.edu/ir/library/pdf/PUB6101.pdf

[11]. Dawson, S., Bakharia, A., & Heathcote, E. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. Networked Learning – Seventh International Conference, Aalborg, Denmark

.

[12]. Buckingham Shum, Simon. "Learning Analytics Policy Brief." UNESCO, 2012. http://iite.unesco.org/pics/publications/en/files/3214711.pdf

[13]. Dawson, S., McWIlliam, E. and Tan, J.P.-L. (2008) 'Teaching smarter: how mining ICT data can inform and improve learning and teaching practice', paper presented at *The Ascilite 2008*, Melbourne, Australia (30 November – 3 December).

[14]. Boyd, Dana and Crawford, Kate. "Six Provocations for Big Data." Working Paper - *Oxford Internet Institute*. 21 Sept. 2011 <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431>

[15]. Schwartz, Paul M. Data Protection Law and the Ethical Use of Analytics, Privacy and Security Law Report. Berkeley Center for Law and Technology. 2011< http://www.law.berkeley.edu/10159.htm>