



Big Data in Education and Learning Analytics

Swati Bajpai¹ and S. Mani²

¹Department of Education, University of Lucknow, UP, India

²Department of Educational Planning and Administration, Tamil Nadu Teachers Education University, Chennai, India

*Corresponding author: kldangwal@gmail.com

ABSTRACT

Due to increasing role of computers, the internet and digital data across many forms of human endeavor is the widespread uptake of the notion of "Big data." Researchers and educators seem to explore the potential of technology to grab, analyze and utilize huge amount of data in increasingly powerful ways. analysis of educational data is not something new in the field of education but increasing interest towards the use of learning management system, collaborative learning electronic portfolio, moodle, intelligent tutoring system, gaming, sharable content object reference model (SCORM), e-learning, m-learning, e-assessment, virtual worlds and augmented reality has fueled research in the field of Big data and learning analytics. Learning analytics is a field related to educational data mining. It describes attempts to use digital data about student's background and learning behaviors in online contexts to monitor and predict student performance. This paper explores the concept of learning analytics, its dimensions, approaches and applications in detail it also describes the relationship between educational data mining, academic analytics and learning analytics.

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

Due to increasing role of computers, the internet and digital data across many forms of human endeavor is the widespread uptake of the notion of "Big data." The term big data has gained prominence not only in academic and research activities but in public domain as well as in political and policy rationales. It is a field of growing interest. It's emergence in the field of education has provided the learners with lots of scope to improve teaching learning process both formally and informally through new extensive educational media, combined with advanced computer technology. Researchers and educators seem to explore the potential of technology to grab, analyze and utilize huge amount of data in increasingly powerful ways. Big data that is present everywhere is too big to handle for standard database software to

process. In the field of education, it is beginning to grow. Although analysis of educational data is not something new in the field of education but increasing interest towards the use of learning management system, collaborative learning electronic portfolio, moodle, intelligent tutoring system, gaming, sharable content object reference model (SCORM), e-learning, m-learning, e-assessment, virtual worlds and augmented reality has fueled research in the field of Big data and learning analytics. 'Learning analytics' is the name for a big data-inspired innovation (Baker and Inventado, 2014) in many developed countries it is being implemented in educational systems. It describes attempts to use digital data about student's background and learning behaviors in online contexts to monitor and predict student performance. Learning analytics is a field related to educational data mining. There are conflicts among the researchers about the definition and aims of learning analytics. An earlier definition suggested by the community emphasized that *"Learning analytics is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections for predicting and advising people's learning."*

However, this definition has been criticized on following grounds:

1. *According to George Siemens, 2010 - "I somewhat disagree with this definition - it serves well as an introductory concept if we use analytics as a support structure for existing education models. I think learning analytics - at an advanced and integrated implementation - can do away with pre-fab curriculum models".*
2. *"In the descriptions of learning analytics we talk about using data to "predict success". I've struggled with that as I pore over our databases. I've come to realize there are different views/ levels of success." Mike Sharkey 2010.*

Another definition of learning analytics given by Erik Duval is:

"Learning analytics is about collecting traces that learners leave behind and using those traces to improve learning."

According to Erik Duval, learners during the process of learning leave some indications or signs related to their learning behavior, which are identified through learning analytics to help them to improve their performance.

Although there is no universally agreed definition of learning analytics, a recent and more often used definition of Learning Analytics comes from the First International Conference on Learning Analytics and Knowledge in 2011:

"Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs." LAK11

This definition draws the attention of many researchers for several reasons. First, it points out towards the practices used in the process of learning analytics such as measurement, collection, analysis and reporting of data. Second, it makes clear that this data is all about learners and in their context. Third, it highlights the purpose of learning analytics that is optimizing learning and learning environment to its maximum level. It also provides us with a tool to improve teaching learning process and the environment in which it takes place through observation, reflection and feedback.

Further, Greller and Drachsler (2012) provide a more comprehensive definition by using general morphological analysis (GMA) to divide the framework of learning analytics into six “critical dimensions”.

Later on, a systematic overview on learning analytics is provided by Chatti *et al.* (2012) and Chatti *et al.* (2014) through a reference model for learning analytics based on six dimensions:

1. Data
2. Environment
3. Context
4. Stakeholders
5. Objectives
6. Methods

It has been observed that awareness of analytics has a wider application in the field of educational institutions/setting. The way ‘learning analytics’ is explained and implemented may vary in different context like:

- ❖ For learners, it reflects their achievements and patterns of behavior in relation to others.
- ❖ For students, it helps them in SWOT analysis.
- ❖ For educators to guide their teaching learning programme and improvement activities.
- ❖ For educational planners to modify current course content and to develop new improved curriculum programmes.
- ❖ For educational administrators to take decisions about human resource management, infrastructure facilities or efficiency and effectiveness measures.

HISTORY

The roots of learning analytics are not from educational background but from business need to understand internal organizational potential and external customer demand. Technological advancement has provided solution to their problems by gathering data from various sources.

Mitchell and Costello first referred to the term Learning Analytics in 2000 in their analysis of the of international market opportunities for online learning products (Mitchell & Costello, 2000).

According to Buckingham Shum & Ferguson, 2012, the emergence of learning management systems such as Blackboard, Moodle, and Desire2Learn that afford the amassment and visual representation of a large amount of student information has enabled the active development of the field of Learning Analytics in the last five years. The development of learning analytics took place through following path or we can say following are reasons for emergence of LA:

- ❖ The increasing interest to explore 'big data' for various purposes.
- ❖ The inclination towards the digital education and online courses using Virtual Learning Environments (VLEs), Content Management Systems (CMSs) and Management Information Systems (MIS) to exploit digital data.
- ❖ To optimize technology to improve teaching learning process and education system.
- ❖ Increasing focus on evidencing progress and professional standards for accountability systems
- ❖ This focus led to teacher stakeholders in the analytics - given that they are associated with accountability systems.
- ❖ Thus an increasing emphasis was placed on the pedagogic affordances of learning analytics
- ❖ This pressure is increased by the economic desire to improve engagement in online education for the deliverance of high quality - affordable – education.

DIMENSIONS OF LEARNING ANALYTICS

There are six dimensions of learning analytics: Objectives, data, instruments, stakeholders, external limitations and internal limitations.

- ❖ **Objectives:** There are two main objectives of learning analytics: Reflection and Prediction. Refection is done through the monitoring and analysis of student's interaction on a particular platform, identification of different types of network connections among them and identification of isolated students to bring them back into mainstream of interaction. Prediction is made on the basis of intervention programs provided, tutoring/mentoring skills of instructor, feedback based on assessment of students. On basis of above activities mentioned, mentor makes the predictions about student's success and their upcoming future.
- ❖ **Data:** Two types of data are available for learning analytics: Open data and protected data. Before using any kind of data first, we have to ensure that whether it is ethical to use that data. Open data is available to everyone in form of big data. Anyone can

access that data as per his or her requirement. While protected data is the data, which is available because of student's interaction and posts in a discussion forum of LMS. It is not available to everyone.

- ❖ **Instruments:** Instruments applicable to learning analytics includes technology, algorithm, pedagogic theories and presentation. The technologies, which are used, are social network analysis (SNS), statistic, visualization, data mining and web surfing. Pedagogic theories such as socio-constructivist theories are used with the assumption that active learners show better results (learning outcomes) in discussion. Network diagram visualization and stats table are used for presentation.
- ❖ **Stakeholders:** In learning analytics, there are two types of stakeholders. First one is Data Subjects that consist of group of learners and the second one is Data Clients, which includes tutors, students, policy makers, academic administrators and discussion moderators.

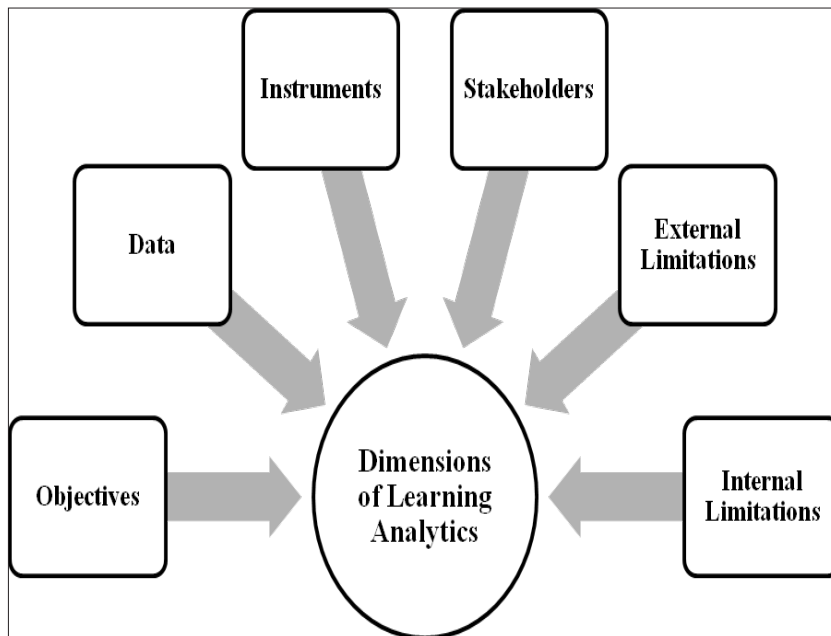


Fig. 1: Dimensions of Learning Analytics

- ❖ **External Limitations:** Convention and norms are two aspects of external limitation. Convention includes ethics, personal privacy, and similar socially motivated limitations. While norms are restriction by laws or specific mandated policies or standard. Another big challenges for learning analytics are ethics and privacy.

- ❖ **Internal Limitations:** Competence and acceptance are two major internal limitations of learning analytics. Competence is user's capability to optimize the benefits. A prolific application of learning analytics entails new higher-order competences to take out the best of learning and teaching. Acceptance factors can further influence the application or decision making that follows an analytics process.

APPROACHES OF LEARNING ANALYTIC

Ferguson and Buckingham Shum (2012) "*Social Learning Analytics: Five Approaches*" defines five approaches of social learning for which one could create instruments and these are:

- ❖ **Content Analysis:** Content analysis involves the analysis of content, which students create, or systematic analysis of oral, written or visual communication of students.
- ❖ **Context Analytics:** Context analytics analyze the environment in which learning takes place. Mobile computing is transforming access to both people and content.
- ❖ **Discourse Analytics:** Discourse analytics is applied to analyze written, vocal, or sign language use, or any significant semiotic event using number of approaches to identify meaningful data on student interaction to explore the quality of language used.
- ❖ **Social Learning Analytics:** The aim of social learning analytics is to explore the role of social interaction in the process of learning, the significance of learning networks and discourse used etc.
- ❖ **Disposition Analytics:** Disposition Analytics play an important role in gathering data regarding student's dispositions to their own learning and finding the relationship of these to their learning.

APPLICATIONS OF LEARNING ANALYTICS

Learning Analytics have a wider application in the field of education such as:

- ❖ **Prediction Purposes:** Learning analytics is used as a tool for prediction of student's performance, achievement and success. Educators use it to monitor their growth in terms of their learning. It is also used to identify students 'at risk' in terms of drop out or course failure.
- ❖ **Intervention Purposes:** Once students at risk are identified then they are provided with intervention programme by their respective educators to support them to upgrade their level of performance.
- ❖ **Personalization & Adaptation Purpose:** Learning analytics provides the students with tailored learning pathways or assessment materials as per their own pace, strengths, needs and interests.

- ❖ **Information Visualization Purpose:** Information Visualization that is the study of (interactive) visual representations of abstract data to reinforce human cognition is used to overview learning data through data visualization tools in the form of so-called learning dashboards. The abstract data may be in numerical or non-numerical form or both such as text and geographic *information*.

SOFTWARE USED FOR LEARNING ANALYTICS

Many of the software tools that are currently used for learning analytics are also applied in the field of web analytics but learning analytics software is used in learner's context rather than general. Some of learning analytics software tools are given below:

- ❖ **Student Success System:** An intervention tool that is used for predictive analysis, visual diagnosis and spot potential problems based upon engagement and performance predictions to guide the students before it is too late.
- ❖ **SNAPP:** Social Networks Adapting Pedagogical Practice (SNAPP) is used as a diagnostic instrument, which allows educators to evaluate student behavioral patterns against learning activity design objectives and intervene as required in a timely manner.

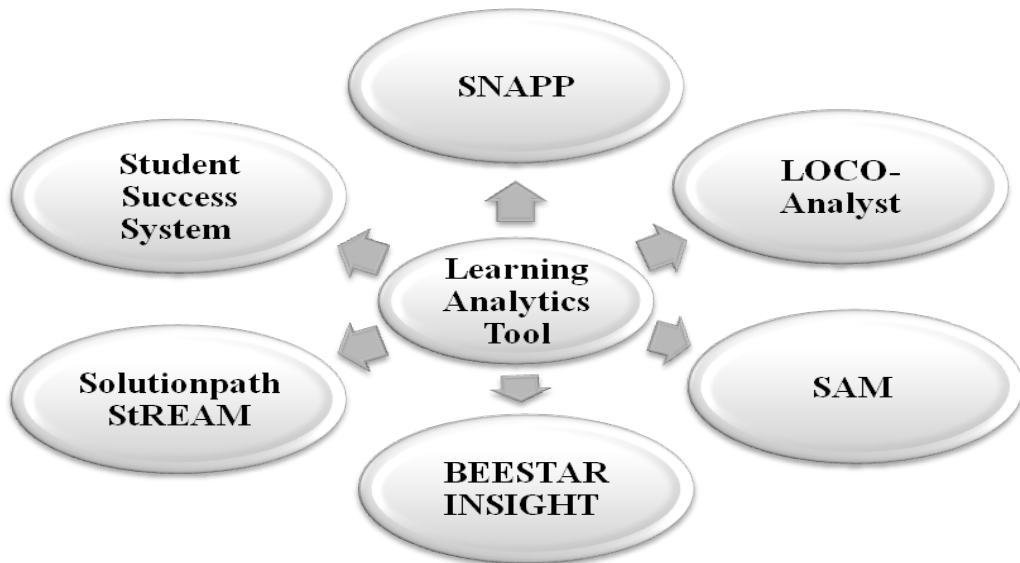


Fig. 2: Learning Analytics Tools

- ❖ **LOCO-Analyst:** LOCO-Analyst is basically designed for general purpose analysis and now it is being applied as educational tool that intended to provide feedback to the teachers about the learning process which is taking place in a web-based learning environment,

which in turn helps them to improve the content and the structure of their web-based courses.

- ❖ **SAM:** Software asset management (SAM) monitor student activity for 'Personal Learning Environments'. It is used for analysis and visualization of learner's data. It also provides feedback for supportive instructor and recommendations for students.
- ❖ **BEESTAR INSIGHT:** It is a tool designed for general-purpose analysis, data extraction and data visualization. It is a real-time system that collects student attendance and their engagement to various assignments automatically & provides analytics tools and dashboards for students, teachers & management.
- ❖ **Solutionpath StREAM:** A leading UK based real-time system StREAM by Solutionpath is an award winning toolset that supports decision making, enables higher education institutions to maximize their resources. It helps in identification of risks and opportunities using reliable data analysis before they happen.

LEARNING AND ACADEMIC ANALYTICS

When we talk about analytics in the field of education, three types of terminology come into existence:

1. Educational data mining
 2. Academic analysis
 3. Learning analysis
- ❖ **Educational Data Mining (EDM):** Techniques, tools and research designs which are meant for extracting meaning from large repositories of data either generated by or related to learning activities of learners in educational settings is referred as educational data mining. It is concerned with the application of data mining, machine learning and statistics to information generated from educational settings such as universities, colleges and tutorial systems. It is a well-developed community running its own journal. In other words, EDM is an emerging discipline that is concerned with developing methods for exploring the unique types of data that come from educational background, and using those methods that are crucial to better understand students and the environment in which they learn. It is necessary to develop and improve the methods used for exploring data at higher level of education, which results in multiple levels of meaningful hierarchy to discover new insights about how people learn in the context of particular settings. Moreover, theories of learning propagated by researchers in educational psychology and the learning sciences are the contributions of EDM in the field of education. This field is closely related to learning analytics and the two have been compared and contrasted on several basis. The

ultimate goal of Educational data mining is predicting student's future learning behavior, analysis of effects of educational support provided, discovering new models of teaching or improving old one and advancement of scientific knowledge among the learners.

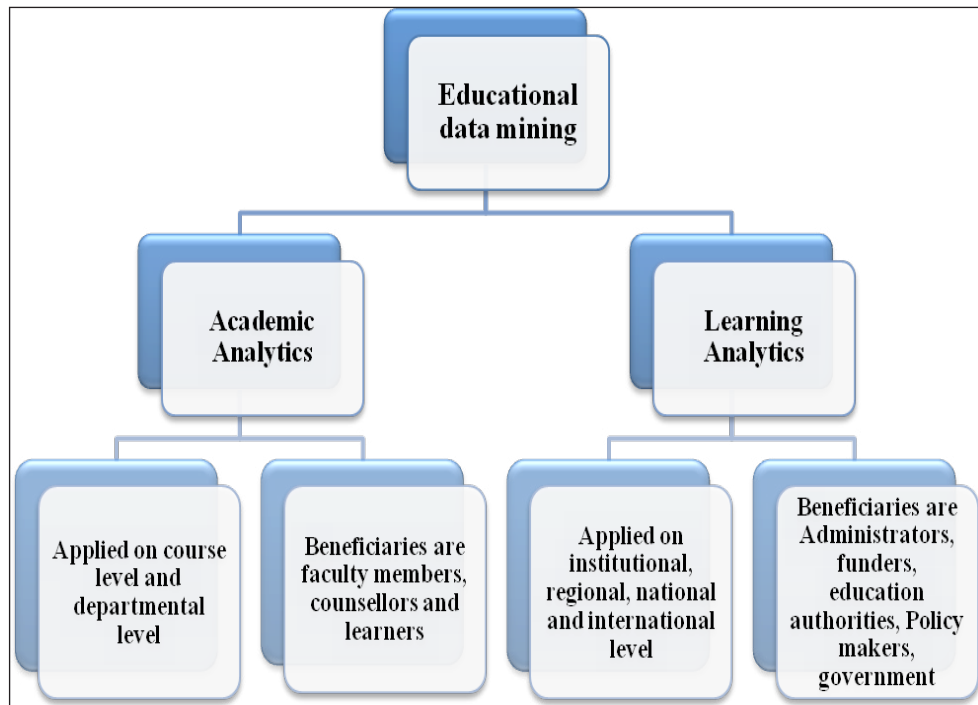


Fig. 3: Relationship between Educational Data Mining, Academic Analytics and Learning Analytics

- ❖ **Learning Analytics:** This community is still developing, but various government and educational reform settings have a great interest in learning analytics and its application. Learning analytics may be defined as data mining plus interpretation and action. Learning Analytics (LA) is closely related to EDM, although the LA lays extra emphasis on a simultaneous investigation of automatically collected data along with human observation in the context of teaching and learning. In contrast to educational data mining, learning analytics is not generally related to the development of new methods of computation for data analysis rather it is concerned with the application of known methods and models to answer the important questions that affect student learning and organizational learning systems. Learning analytics has a relatively greater focus on human interpretation of data and visualization (though there is a tradition of this in EDM as well -- cf. Kay *et al.*, 2006; Martinez *et al.*, 2011) while EDM has a relatively greater focus on automated methods.

- ❖ **Academic Analytics:** The term Academic Analytics has been around for about a decade, based on early work by Diana Oblinger and John Campbell. As initially presented, the concept addresses a mix of administrative and learning analytics. For clarity sake, this concept is now closest to what is called business intelligence in corporate settings. Academic Analytics is “the improvement of organizational processes, workflows, resource allocation, and institutional measurement through the use of learner, academic, and institutional data. Academic analytics, akin to business analytics, are concerned with improving organizational effectiveness.” (Siemens *et al.*, 2011).

These three terminologies are interrelated to each other. Educational data mining plays a vital role for both learning analytics and academic analytics. Although learning analytics and academic analytics are two distinct communities but both are contributing in transforming large educational data into meaningful educational practices. Thus also contributing in the field of educational research by enriching users with new tools and techniques to elevate the level of teaching learning process.

CONCLUSION

Learning analytics is a field of growing interest, there are lots of opportunities to explore and discover its potential to enrich and enhance teaching learning process. Every year, new researches in the field make us to better understand the potential application of these technologies to optimize our teaching learning resources. Better, we know the technologies better, we apply them in various situations to obtain fruitful results. Innovation and discovery of new techniques is not something our goal but the precise application of them is crucial. This is an area of big concern on which we all need to work. As a field, learning analytics is still under the process of development and requires a lot of attention and research work to develop it to its maximum potential.

REFERENCES

- https://en.wikipedia.org/wiki/Learning_analytics retrieved on 13/6/2017 at 06:19pm
- http://etec.cctl.ubc.ca/510wiki/Learning_Analytics:_An_Introduction_and_Critical_Analysis retrieved on 16/6/2017 at 11:20am
- http://edutechwiki.unige.ch/en/Learning_analytics retrieved on 16/6/2017 at 11:35am
- <http://www.learninganalytics.net/?p=131> retrieved on 16/6/2017 at 06:29 pm
- <http://www.laceproject.eu/faqs/learning-analytics/> retrieved on 16/6/2017 at 06:39pm
- <https://tech.ed.gov/learning-analytics/> retrieved on 16/6/2017 at 06:41pm
- <http://digitallearning.northwestern.edu/learning-analytics> retrieved on 16/6/2017 at 06:58pm
- <http://www.solutionpath.co.uk/> retrieved on 17/6/2017 at 11:54am
- <http://edutechwiki.unige.ch/en/SAM> retrieved on 17/6/2017 at 12:18pm

http://edutechwiki.unige.ch/en/Beestar_insight retrieved on 18/6/2017 at 07:47pm
<http://edutechwiki.unige.ch/en/LOCO-Analyst> retrieved on 18/6/2017 at 08:19pm
<http://edutechwiki.unige.ch/en/SNAPP> retrieved on 18/6/2017 at 08:30pm
<http://firstmonday.org/ojs/index.php/fm/article/view/6872/6089> retrieved on 19/6/2017 at 11:05am
https://en.wikipedia.org/wiki/Educational_data_mining retrieved on 20/6/2017 at 12:03pm
<http://circlcenter.org/educational-data-mining-learning-analytics/> retrieved on 20/6/2017 at 12:13pm
<http://www.upenn.edu/learninganalytics/ryanbaker/Educational%20Data%20Mining%20and%20Learning%20Analytics%20-%20DRAFT.pdf> retrieved on 20/6/2017 at 12:58pm
http://sites.arteveldehogeschool.be/efye/sites/default/files/efye_2016_0.pdf retrieved on 22/6/2017 at 12:10pm

