Learning Analytics – New Flavor and Benefits for Educational Environments

Aleksandra KLAŠNJA-MILIĆEVIĆ, Mirjana IVANOVIĆ

University of Novi Sad, Faculty of Sciences
Trg Dositeja Obradovića 3, Novi Sad, Serbia
e-mail: akm@dmi.uns.ac.rs, mira@dmi.uns.ac.rs

Received: Mach 2018

Abstract. Amount of educational data has been constantly increasing for years in all domains and kinds of education (formal or informal) and educational activities (teaching, learning, assessment, use of social media and collaboration and so on). Accordingly, Learning Analytics (LA) become a powerful mechanism for supporting learners, instructors, teachers, learning system designers and developers to better understand educational processes and predict learners’ needs and performances. In this paper, we analyze the important dimensions and objectives of LA, application possibilities and some challenges to the beneficial exploitation of educational data. The required skills and capabilities that make meaningful use of LA techniques and technologies in this domain are considered and identified. Presented findings can act as a valuable guide for setting up LA services in support of educational practice. Also, they can be used as learner guidance, in quality assurance, curriculum development, and in improving learning process effectiveness and efficiency. Finally, this paper proposes the unavoidable constraints that affect LA technologies in education.

Keywords: Learning Analytics (LA); education; application; learning environments; effective learning system.

Introduction

During the last years, developments of new learning technologies, devices and environments, such as digital learner records; different types of learners’ cards; mobile devices; use of eye-tracker, virtual reality, augmented reality devices; sensors; flexible classroom design; and Massive Open Online Course (MOOC) are completely transforming the approach of learning and teaching (Sharple et al., 2014). Higher education institutions are collecting more data than ever before using different kinds of Learning Management Systems (LMS), Content Management Systems (CMS), Tutoring systems and educational platforms through combination of traditional and blended learning style. Management of this vast amount of collected data, named “big data”, should offer valuable comprehension about the learning process, insights about risk of learner’s dropping out,
and support for increasing learners’ success. In order to comprehend the patterns of value that exist and should be recognized within the large amount of data, new, effective and innovative approaches are required. Lot of exploration and researches aim to handle the data with the proper techniques, methods and new tools to produce real time solutions and predictions in this certain area. The utilities of proper techniques and new tools could be: operative self-learning, useful peer groups, available class time for creativity, and possibilities for critical thinking in problem solving (Jevsikova et al., 2017; Pamanitsiou, Economides, 2014).

Learning Analytics (LA) encompasses numerous analysis techniques that successfully extract and recognize important information from “big data” such as discovering patterns and applying them to the education process (Ruipérez-Valiente et al., 2017). In its initial steps of evolving, there has been numerous definitions used for Learning Analytics. Siemens defines it as “the use of intelligent data, learner product data and analysis models to discover information and social connections, and to predict and advise on learning” (Siemens, 2012). Elias described it as “an emerging field, in which sophisticated analytic tools are used to improve learning and education” (Elias, 2011). Learners and teachers leave many traces behind them in which LA can convert them to be beneficial for the education sector (Duval, 2011). The most cited definition of LA came from an open online course on learning and knowledge analytics (LAK11) and was adopted by the associated First International Conference on LA and Knowledge in 2011 (Long, 2011): “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”.

Many researchers, research groups and developers of educational systems study the use of LA in different fields related to the data availability, the competence, applicability, the cost, the privacy, the relevance and the ownership (Adams Becker, 2017; Klašnja-Milićević et al., 2017). In this paper we will concentrate on contemporary study that examines the benefits, applicability and positive educational effects of LA in higher education institutions.

The rest of the paper is organized as follow. The second section gives an overview of key trends in education environments, dataset and analytics. The third section is devoted to educational LA applications and their role in contemporary education. The challenges of implementation of LA in education environments are pointed out in the section 4. The fifth section of the paper provides the concluding remarks and possible future research directions.

This paper is an extended version of a paper presented at Constructionism 2018 and published in conference proceedings.

What Kinds of Benefits Educational Data Can Bring to Learning Environments?

Learning Analytics can obtain advantages of available educational datasets from different Learning Management System (LMS), Content Management System (CMS), Tutor-
Learning Analytics – New Flavor and Benefits for Educational Environments

Educational Institutions have been constantly producing and storing large amounts of different kinds of students’ data and use these data for different purposes like: preparing different kinds of reports for wide range of stakeholders, assessing teachers’ and students’ activities, make range of statistical processing. Administering students’ progress and reporting in order to apply for and to receive funding from the public authorities are the most commonly known purpose of using available institutional educational data. Linking such available datasets would facilitate the development of rather complex mash-up applications that can lead to more learner-oriented services and therefore improved personalization in learning and teaching activities. LA strongly relies on data about learners and one of the major challenges that LA researchers are facing is the accessibility of publicly available datasets to evaluate their LA methods. Most of the data produced in institutions is protected, and the protection of student data and created learning artefacts is a high priority for Information Technologies (IT) services departments. Nevertheless, analogous to Open Access publishing and related movements, calls for more openness of educational datasets have already been brought forward (Drachsler et al., 2010). Additionally, researchers need “good and useful” datasets i.e. such collected data that gives good basses for discovering useful learning patterns.

Greller and Drachsler (2012) developed a framework that includes six critical dimensions related to an LA initiative. Each of the dimensions must be addressed to institutionalize an LA initiative successfully. Fig. 1 illustrates critical dimensions according to educational environments. These essential dimensions include key educational Stakeholders, Educational Data and Environments, necessary educational Competences, inevitable Objectives, and adequate contemporary Technologies and Techniques. But we have not to neglect importance of different Constraints, especially Security, Privacy and Ethics.

Some researchers specified that higher education has to include the analytics tool into the system in order to improve productivity (Daniel 2016; Hrabowski et al. 2011;
First it is essential that data analytics software and adequate methods can provide feedback to learners and teachers about educational performances:

1. **Feedback**: Learner often might fail at particular subject but do not know what the reason was. It becomes valuable, even necessary in fair educational activities, when the learner can look not just at himself and his actions, but at other colleagues who have had the same or similar experience. (S)he can get an insight either that would describe it so (s)he is not frustrated or that (s)he could use to correct it so that (s)he could be successful again. The improvement of electronic learning modules supports evaluation of learners in logical, real-time ways. In order to predict learner outcomes such as dropping out, needing extra help, or being capable of more demanding assignments, this approach can analyze underlying patterns. Pedagogic approaches that seem most effective with particular learners could be identified and after that applied adequately.

2. **Tracking**: In order to understand the real learning patterns and activities of learners more effectively tracker devices can be used for teachers, by allowing them to track a learner’s experience in an e-learning course. In observing the digital paths learners leave overdue. Teachers can track learners’ passage during the whole learning experience within particular topic, some bigger portion of teaching material or even during the whole course.

3. **Efficiency**: LA can save many hours of time and effort, when applying strategies and try to achieve various goals. This means students can receive better support from lecturers, and take greater control of their own learning progress.

4. **Understanding the learning process**: By using LA in e-learning, teachers can see which parts of a course were too easy and which parts were too difficult that the learner has failed to solve. Other parts of the learner’s path teachers can analyze after that and consider pages re-entered often, preferred learning styles, sections recommended to peers, and the time of day, when learning operates at its best.

5. **Collaboration**: Experts from many different fields (Education, Pedagogy, Psychology, Cognitive Sciences, Engineering, Computer Science especially from Artificial Intelligence and Data Mining fields, Instructional Design and so on) have to come together to retain a Learning Management System function at its best. Recently the most important part of educational systems and environments is Human-Computer interaction. It should offer to learner friendly, highly functional and pleasant interface to communicate with the system and easily use teaching material. This encourages high-level and quality cooperation, powerful teamwork, and interdisciplinary and multidisciplinary approach through all educational stages and processes.

6. **Personalization**: LA can also be effective in guiding teaching material designers and designers of e-learning systems by allowing them to personalize courses to adjust their learners’ individual needs. This will allow developers of e-learning systems and educational environments to promote the standard for effective and exceptional e-learning courses.

Regardless of the educational performances, many educational institutions and especially faculties, and universities have confirmed that analytics can support signifi-
cant improvement of their activities and performances, including resource distribution, learner achievement, even administration, and financial sector. Some important features that LA offers to institutions are listed below (Klašnja-Milićević et al., 2017; Daniel, 2015; Siemens, 2011).

- Assisting in creating common sense of complex educational topics and content through the combination of social, technical, and information networks. Algorithms can recognize and provide insight into data and at-risk challenges.
- Innovating and transforming the college and university system and functioning, in addition to already existing educational models and pedagogical approaches.
- Improving and make faster and more efficient administrative decision-making processes and effectiveness of organizational resource provision.
- Helping leading institutions, educational policy makers, and educational authorities transition to holistic decision-making through analyses of “what-if” scenarios.
- Exploring how different components within a complex discipline (e.g., remembering learners, decreasing costs) connects and discovers the influence of varying essential components.
- Increasing administrative efficiency, effectiveness, and productivity by providing up-to-date information and allowing fast and as much as possible proper reaction to challenges.
- Constant testing and evaluation of curricula and prompt reactions if some parts have to be changed and/or improved.
- Helping official leaders to control the hard (e.g., research, patent) and easy (e.g., quality of teaching, reputation, profile,) value created by faculty and university activities.
- Evaluating typical and widely spread grading techniques, instruments, and practices (i.e., departmental and licensing exams).

**Adaptive Learning System**

Online learning environments (learning software, learning platforms, LMS – learning management systems), Content Management Systems (CMS), Tutoring Systems and so on should have the ability to recognize fine-grained learner behaviours feedback to improve learning, teaching, and decision making (Olajuwon and Adebayo, 2017; Klašnja-Milićević et al., 2009; Camilleri, 2017). Based on the conclusions we reached in the study of adaptive learning system in the education field, this section describes a principal and general architectural framework suitable for support higher educational research (Klašnja-Milićević et al., 2017). The functioning of such adaptive system, using the predictive models incorporates following important components, elements and interactions (Fig. 2):

- Students interact with a content component in order to deliver personalized subject content and evaluation for student learning support.
- Database devoted to students’ learning stores time-stamped student data, learning behaviors and learning sequences recorded as students’ activities within the system.
- A predictive model uses learning behavior data from the database of students’ learning and to track students’ progress in order to make predictions about their future activities, preferences, and course withdrawals.

Fig. 2. Adaptive learning system with Predictive component.
A reporting server uses the output of the predictive model to produce dashboards that provide visible feedback for various users.

Background information on students can be used to group them into specific categories. Student profiles are collected in Student Information System (SIS).

According to the output of the predictive model, an adaption engine refines the content of delivery component to prepare material for students’ preferences and knowledge level.

Administrators, system developers or teachers can use an intervention engine which allows modification for better serving student’s learning processes.

Adaptive learning system enables various learning sequencing, according to individual learner rhythm, considers knowledge level, provides a collaborative interaction when solving tasks, personalizes materials prepared for learners, and gives in-depth analysis of learners’ performances. The application of adaptive learning systems is extremely important for higher education institutions and all stakeholders of educational processes. The ability of an adaptive learning platform to harmonize and coordinate more efficient students’ learning will expand the way in which students will learn in the future.

Application Possibilities of Learning Analytics in Educational Environments

Nowadays, researchers in LA and educational data mining fields tries to answer increasingly important and complex questions: what a student knowledge is and whether a student is actively engaged in learning process. Scientists and researchers from different disciplines connected to educational processes have actively been included, analyzed and experimented with new techniques and new kinds of learning system data that have shown promise for predicting students’ learning achievements and outcomes (Jivet et al., 2018, Becker et al., 2017). In this section we pointed out a range of potential application areas for LA and educational data mining. These areas represent the broad categories in which LA can be applied regarding students’ online activities in e-learning environments and systems. Fig. 3 presents a summary of selected areas and appropriate data types needed for analysis.

In fact, Learning Analytics systems apply variety of models and techniques to answer such questions as:

- When are students ready to move on to the next topic during use of teaching material presented in e-learning system?
- When are students falling behind in a course and discover what topics were bottleneck?
- When is a student at risk for not completing a course and what topics and activities have been discouraging?
- What grade is a student likely to get without any additional external intervention?
- What is the best next course for a given student according to achievements (s)he obtained in previous learning paths?
Should a student be referred to a teacher for help, in which situation and part of teaching material i.e. course?

Responding to these questions involves the collection of students’ entry and correctness of the answer, student activities on learning systems over time, when and to which group is specific for learning strategy, and performance of students on pre / post-tests.

Research on the development of new and improvement of existing various components of the learning system can significantly contribute to the design of better learning...
systems as it has strong implications for student learning. This area is in progress and represents a key point for educational data processing and LA (Baker, 2010). Haixiang et al. (2017) proposed specific approach known as learning decomposition which is in fact an alternative and effective method. Learning decomposition method involves exponential learning curves to performance data and relating student success to the amount of each type of pedagogical support. The weights indicate how effective each type of pedagogical support is for improving learning.

Scientists from Carnegie Mellon University went a step ahead and have built cognitive models of mathematics, which have become the basis for high school programs that include intelligent tutorials (Ritter et al., 2007). In these systems, complex tasks are divided into individual components and portions of knowledge, and the model is used to improve the learning process and the problem-solving strategy. Each student activity is associated with one or more skills. In this way, researchers can use data from the tutorial system to dynamically assess the teaching efficiency. Evaluations and improvements to this model have been carried out over the last 15 years with increasing quality effects.

**Learning Analytics Dashboard Applications**

Learning analytic dashboard can be defined as a form of “personal informatics” applications (Li, Dey, Forlizzi, 2010). These applications track students’ log-files and collect personal data about different characteristics of behavior patterns, interests, habits, feelings, which encourage user self-knowledge by making available tools for the review and analysis of their personal history (O’Donoghue, Rabin, 2003). We can distinguish four stages of a learning analytics application: Awareness, Reflection, Sensemaking and Impact. In the field of higher education, dashboard applications can support learners and teachers in online and blended learning or traditional face-to-face teaching. Such dashboards (Few, 2006) provide graphical representations of the current and historical state of a learner or a course to enable flexible decision making.

Verbert et al. (2013) reviewed fifteen learning analytics dashboard applications in order to discover and precise their characteristics and functionalities: GLASS (Gradient’s Learning Analytics SyStem), CourseVis, TeacherADVisor, Course Signals, Moodle, Social Networks Adapting Pedagogical Practice, TUT Circle, CALMsystem, Student Activity Meter, Student Inspector, LOCO-analyst, StepUP!, Classroom View, OLI dashboard and Tell Me More.

In Table 1, short description of 10 more influential developed learning analytics dashboards applications are presented. They are organized into three types according to foreseen target users: LA dashboards applications for a) teachers, b) learners and c) both teachers and learners. Ten systems have been evaluated with teachers and/or learners. Usability, usefulness and effectiveness have been evaluated for four systems. Evaluation results indicate that providing learners with a dashboard has an impact on their knowledge level, satisfaction and exam results. Significant difference is observed in retention rates of learners using the dashboard in at least one course and learners which not using the dashboard.
In these LA dashboards applications, several distinctive data are tracked: target users, social interaction, time spent on the system, use of documents and tools, artifacts produced, and exercise/quiz results. Evaluation of linguistic data provide useful knowledge about learners’ preferences, needs, worries, preoccupation and language development. Learning analytics are universal and can be used in learning course without focus to language learning in particular. Learning analytics dashboard applications have the opportunities for encourage student’s awareness, reflection, sense making, and impact that such dashboards provide. As well, these applications have the potential to improve learning, to get better at getting better.

### What Challenges Learning Analytics Can Offer to Higher Education Stakeholders

Regardless of the fast development of methods, techniques, and applications that support implementation of processing and powerful use of big data techniques still there are more places and challenges for improvements and advances. This trend is extremely important and promising in higher education systems and environments there exist big
amount of different kinds of educational data (umak et al., 2016). Moreover, there are also several apprehensions that affect educational data mining and Learning Analytics technologies. The main issues that can be observed in the application of educational data mining and LA in education are related to data profiling, privacy, and the rights of learners with respect to their individual behavior recording (Boyd, 2010). For example, some important issues must be taken into consideration (Picciano, 2012):

- Should learners be informed that their activity is being followed? It is one of very important issues that has to be considered in e-learning systems. If students know that their activities are recorded their learning behavior can be strongly influenced by this fact. So, collected data might not be realistic and might lead us as teachers even into wrong conclusions.

- How much information is necessary for faculty, students, parents, scholarships, and other issuers? The question which kind and amount of data are optimal to be collected during the educational processes is also essential. If e-learning systems and environments are implemented in manner to collect and record too many different kinds of data, then processing could be cumbersome. Even more different educational stakeholders after processing collected data can get rather useless information or based on them to get wrong decisions.

- In what manner should faculty affiliates act in response? Faculty affiliates are not University employees but are associated with an academic department or organization of the University for the purposes of teaching or academic oversight. Those who hold the courtesy appointments typically contribute to the instructional, research, creative activity, or service functions of the University. Colleges, Schools, Departments, and Programs may recommend these appointments for approval, according to the Faculty Affiliate policy.

- Do learners have a requirement to look for support? This aspect in e-learning systems recently appears as crucial, not only for learners, but also for teachers. Students in such systems have to have possibility for additional support. Contemporary technologies significantly influence development of high-quality HCI in e-learning environments that facilitate adequate support for students during learning activities. Additionally, starts to be very important to give learners opportunity to see effects of their activities and somehow based on them to determine their further learning paths through topics.

Protection should be undertaken to confirm that the well-known collections of personal data of learner transactions are not hurt individuals. One encouraging and very usual approach to resolving these issues could be masking or anonymized the data at its source (Barlow, 2013). Masking is one type of creative approaches that will make large-scale applications of data possible while still protecting the confidence of students’ and teachers’ information. New performances and competences for software applications permit sensitive data to be masked at the database level, when brought into a data warehouse.

With the adoption of LA in the educational fields, educational institutions and key stakeholders are required to adjust their policies with legislative framework. Many in-
itutional policies failed to fully reflect the ethical and privacy issues and implications of LA (Prinsloo, Slade, 2013). In the rest of the section, we specify several possible principles that an ethical LA policy should describe that can be useful for educational institutions:

1. Precisely specify principles of collection of sensitive personal information: sex, date of birth, address, ethnicity, occupational status, qualifications and study records.
2. The precise specification of use of personal information, if it is for the benefit of the students, such as predicting students’ learning behavior and advices based on conclusions obtained by LA, or if it is for research reasons to achieve LA objectives.
3. Precise comprehensively all aspects of methodology of data collection either by the student’s input him/herself or by other services, such as browser cookies.
4. Security principles for keeping the data protected are of essential importance in wide range of e-systems and applications including e-learning environments.
5. A description of the time period of keeping learners’ data and rules for deletion process. (Singer, 2014)

All LA tools should follow convenient security principles to keep the analysis results and the students’ records safe from any threat (Anciaux et al., 2006).

- The confidentiality property promises that the data can never be accessed by an unauthorized stakeholder/parties.
- The integrity property promises that the originally collected data cannot be watched, changed, or transformed. Or if there is a need for such activities (for example to repeat processing of data, to additionally experiment with slightly modified data, and so on) that adequate agreed procedures and steps will be followed.
- The availability property guarantees that the data should be available for authorized parties to access when needed.

A key component of protecting learners’ information is encrypting their data in order to achieve the confidentiality concept. Encryption guarantees that only authorized parties and stakeholders can use the data. Moreover, assuring confidentiality can include: invoking file permissions and granting a secure operating environment, while cryptographic hashing of datasets can assure the integrity property of students’ records (Chen & Wang, 2008).

As LA is an emerging research topic in the field of Technology Enhanced Learning and a forthcoming unavoidable and promising trend (Ebner, Schön, 2013), accuracy and validity of information is highly questionable. Faults related to selection a wrong dataset, or not recognizing the component and attributes relevant to data will negatively affect the accuracy of the outcome (Krasnow Waterman, Bruening, 2014). Therefore, a wrong selection of educational dataset will lead to inaccurate results and can highly negatively influence learning processes. Results and conclusions obtained by processing wrong educational datasets, if they are applied in educational systems for example
for changes of teaching material, can completely violate effects and gaining of adequate knowledge and skills.

The essential questions we can ask here are: What if LA results were wrong? And what if the predications or the interventions went wrong? Accordingly, LA would aim to provide guarantees that it’s analyzing, and picking the data, fit quality criteria and produce an agreed level of accuracy.

Nevertheless, all participants (but predominantly teachers, learners and LA experts) in developing and using e-learning system and processing collected educational data have to be aware of abovementioned problems and issues and try to seriously take care about them in all stages and life circles of such systems.

Data protection and copyright laws are legal restrictions that limit the beneficial use of LA. Such legal restrictions are: limitations of keeping the data for longer than a specific period, which are regulated differently in each country; the data should be kept secure and safe from internal and external threats; data should be used for specific purposes and the results of any process should be as accurate as possible. The restrictions could be stronger when it relates to personal information. Applying social network analysis as a method of LA causes the adoption of personal information.

There are two main perspectives that are very important issue i.e. about who own the data: students and/or institutions. Jones et al. (2014) concluded that single ownership is not good model and neither the students nor the institutions should win the ownership of the data. They suggested a hybrid module that merges both perspectives. Institutions can invest the students’ data in analytics, develop new personalized learning platforms and benchmark their learning management system. While students want to enhance their learning and maintain their performance, they would like to ensure that their information is kept confidential. An uprising question we like to address here is: What if LA methods have to modify the students’ data for prediction purpose and additional more quality findings.

Conclusions

Learning Analytics is a definitively promising research field, which provides new and innovative methods, tools and platforms that influence researchers in Technology Enhanced Learning. Higher education institutions are applying LA to improve the facilities they provide for students but also other educational stakeholders and to improve observable and measurable learning outcomes (e.g. grades, success and awards, as well setback and obstruction) (Cooper, 2013). Nonetheless, this emerging new field lacks an approach that explains a complete overview of its processes. This research study revised the definitions of LA and the aspects that encourage its extension. This provides an entire overview consisting of: learning environment, dataset, analytics, constraints and the interventions which are interpreted to achieve the main goals of LA. Based on this approach, we identified the key educational stakeholders, introduced particular examples and possible areas of LA, presented methodologies and discussed the objectives.
Finally, we covered the way of determining the challenges that surround LA and shed the light on the extremely important issues in e-systems like privacy, security and ethical issues. Furthermore, we arise and anticipated various questions that need serious efforts for further research in near future.

In the future, educational institutions should try to balance the institutions own philosophy of learner development on the one side and various federal privacy laws on the other side. It is significant that organizations comprehend the dynamic nature of educational success and retaining, offer surroundings for open dialogue, and enhance practices and approaches to solve these issues.

We hope that research achievements and clarification of some important aspects of LA application presented in this paper and the implications derived from them will advance the discussion about building more effective learning systems for learners, instructors, course designers, and similar activities in educational institutions.

References


A. Klašnja-Miličević is an assistant professor at Faculty of Sciences, University of Novi Sad, Serbia. She joined the graduate program in Computer Sciences at Faculty of Sciences, Department of Mathematics and Informatics, University of Novi Sad in 2003, where she received her M.Sc. degree (2007) and PhD degree (2013). Her research interests include e-learning and personalization, Intelligent Tutoring Systems, information retrieval, Internet technologies and recommender systems. She actively participates in several international projects. She has also served as Program Committee member of several international conferences. She co-authored one university textbook and one monograph. She has published over 30 scientific papers in proceedings of international conferences and journals.

M. Ivanović holds the position of Full Professor at Faculty of Sciences, University of Novi Sad. She is a member of the University Council for Informatics. She is author or co-author of 14 textbooks, several monographs and more than 350 research papers on multi-agent systems, e-learning, and intelligent techniques, most of which are published in international journals and conferences. She is/was a member of Program Committees of more than 230 international conferences, participant of numerous international research projects and principal investigator of more than 15 projects. Mirjana Ivanovic delivered several keynote speeches at international conferences and visited numerous academic institutions all over the world as visiting researcher and teacher. Currently she is Editor-in-Chief of the Computer Science and Information Systems journal.