Predictive online analytics tools in the educational environment

A I Vinogradova¹, A N Gorodishcheva¹, A V Gorodishchev², Yu V Paskhalskaya¹ and E S Yaskov¹

¹Department of Advertising and Cultural Studies, Reshetnev Siberian State University of Science and Technology, 31, pr. Krasnoyarskiy Rabochiy, 660037 Krasnoyarsk, Russia
²Krasnogvardeysky District Centralized Library System, 8, Sredneokhtinskiy ave., St. Petersburg 195027, Russia

E-mail: vinogradova50@gmail.com

Abstract. The article on the higher education topic identifies four trends associated with the introduction of digital technologies and tools into the educational process: the formation of a blended learning model; transition to online learning; creation of a virtual (digital) educational environment; changing the approach to the management of educational organizations. The perspective directions of using artificial intelligence in the higher education field are considered and analyzed. Artificial intelligence is an auxiliary, but valuable tool that can perform and improve a large number of different operations carried out in a university, help in organizing an effective educational process and building the necessary communications.

1. Introduction

Digital technologies are changing the person way of life, his social space with such an intensity that humanity has not yet encountered in its history. In the political and scientific debate, it becomes more and more obvious that the answers to these changes must lie in the field of education [1]. The promotion of digital capital is called a “wicked challenge” by the authors of the analytical report EDUCAUSE, which is difficult even to define [2]. The increasing digitization of teaching and learning and the related changes require new approaches to organizing learning processes. Today we have no need in preparation and future development of the student’s qualification for the digital environment. We have to learn how we can change teaching strategies at universities and all educational process with considering of digital design. As a result, new methods of educational data (Educational Data Mining, EDM) appeared with added value of studying exclusive data types. Also, there is a process of adaption the education environment because of the accessible data expandable space, understanding student’s life goals. All this is necessary in order to predict intellectual results in academies, improve the routine of the educational process, and make organizational decisions.

The point of this paper is to analyse the educational data analytics levels. An approach to predicting the learning success based on a cognitive model is important for understanding the efficiency of mastering learning materials by students in an information-rich environment.

Learning and understanding of so-called digital competencies by students is increasingly forming the basis of the student achievement system. According to the analysis many of the competencies formed
to promote “digital skills” with a qualification approach. As a result, professionals should develop their qualification for deep understanding of the digital environment, which provides intuitive adaptation to new contexts and content creation, from the level of teaching skills with using isolated technologies [3].

In the technology development context, the highest requirements are imposed on the intellectual and cognitive-cognitive potential of future professional personnel [4]. These include: the ability to generate and process complex information, think critically, make quick decisions, use various forms of evidence, pose the necessary questions, be adaptable and flexible in relation to new information, identify and solve real problems [4]. Skills of critical thinking, reasoning, operating with categories, understanding complex concepts, connections and context spaces are required. Most of the listed skills are related to cognitive processes of a higher order, it is necessary to pay special attention to them in the educational process. There is a problem of obtaining data that is accumulated in the learning process and allows one to draw conclusions not only about the academic performance of students, but also about the development of the cognitive sphere. Not only the learning goals are changing, but also the strategies for designing the learning process using different methods and forms [5, 6]: group, individual, team, independent - using mixed technologies, where digital and physical objects can coexist together.

2. Learning platforms and dashboards
Learning platform providers and business analysts are adopting dashboards that combine data into aggregated views that drive innovation. The data space today offers institutions new opportunities for assessing, measuring and documenting learning [1]. The volumes of data are generated at an increased rate and the process of diversification of the data sources become stronger.

As a result of the development of technologies, previously disparate data can now be collected in educational data sets Dataset and represent a new direction of research in the field of educational data analytics, Learning Analytics [7, 8].

Attraction specialists of cognitive neuroscience, developmental psychology, psycholinguistics, computational linguistics, ergonomics, artificial intelligence technologies, and knowledge engineering become a new requirement for designing the educational process. There are two main research communities in EDM area- Educational Data Mining (EDM) and Learning Analytics and Knowledge (LAK). They unite a large group of researchers dealing with the development of methods and tools for data analysis in interdisciplinary area of EDM. Since 2008, they have held annual seminars and conferences for researches collaboration [9]. The main areas related to data mining of the educational process / student analytics is shown on figure 1.

![Image of Education Data Mining Learning Analytics Venn Diagram](image_url)

**Figure 1.** Main areas related to Educational Data Mining/Learning Analytics.

New educational system must include different exercises which have been linked and resolved through Data Mining process. As an example of Baker’s learning environment we have five approaches/methods: prediction, clustering, relationship mining, distillation of data for human judgment and discovery with models, which is used in four key areas of application for EDM: improving student models, improving domain models, studying the pedagogical support provided by learning software,
learning and learners scientific research. On the other hand we have Castro model (Table 1), that consist of applications dealing with the student’s learning performance assessment, applications that provide course adaptation and learning recommendations on the student’s learning behavior based, approaches dealing with the learning material and educational web based courses evaluation, applications that involve feedback to teachers and students in e-learning courses, and developments for atypical students’ learning behaviors detection. [10].

Table 1. EDM approaches key applications.

|   |   |
|---|---|
| 1 | Prediction |
| 2 | Scientific research into learning and learners |
| 3 | Clustering |
| 4 | Relationship mining |
| 5 | Data distillation for human judgment |

In general, Learning Analytics using different methods to measuring, collecting, analyzing and presenting data on the students’ progress and activity in the learning environment, identifying the reasons for success / failure, ways to help students at academic failure risk [11]. The introduction of analytical research methods displays the identification of problems in student learning already in the second week of the semester, which was not possible before, or data analytics is used to predict student progress in the next academic year, to choose strategies when learning is difficult, to improve understanding of key concepts, search and organization of information in individual training, etc. [1].

3. Levels of educational data analytics

Educational processes are increasingly supported by information technologies, thanks to which streams of heterogeneous data are created. Aggregating data from various sources (online sources, performance data, survey results, student portfolios, administrative data, etc.), their analysis and meaningful assessment can have an impact on the learning process of an individual student.

Data analytics results can be used to identify success / failure factors in learning or to tailor curricula accordingly and optimize teaching / learning.

Collection and aggregation of data accumulated in the learning process, to which the methods of machine processing and data analysis can be applied, is the goal of the Learning Analytics approach [12]. One of the main challenges is to recognize the needs of individual learners in order to support and guide the individual learning process through evidence-based adaptation. Learning Analytics covers a wide range of analysis represented by three levels: macro, meso and micro levels [13].

Macro level includes access to data at the national, regional, national levels. At the macrolevel, it is possible to change academic models and pedagogical approaches.

Going to the meso-level obliges us to work at the level of individual educational institutions, sectors, including the analysis of unstructured data, forecasting and optimization of work processes to develop the decision-making and resource allocation at the organizational level of the process.

At the micro level we start to track and explain the individual processes, like data about students (or groups), user activity and productivity, student interests, interpersonal data. Obtaining the feedback can be the main way for information gain. Micro – level Data can be obtained in a short period of time and
give a great level of detailing. We use the information on this level to find out students at risk, to improve their learning and cognitive performance as soon as it possible.

As a result of Data mining and accumulation we receive a pool of different relevant resources for teachers, managers, methodologists, administrators, data analysts. Also, we can use analytic categories such as predictive analytics, adaptive analytics, social network analytics, discourse analytics [13] in education in various scenarios.

4. Conclusion
Designing the educational processes with using the analytical categories and methods provides representative models and structures educational forms that are invariant to the students characteristics.

The experimental method is reduced to the limit of the parameters, which is used in the individual model associated with specific student. Also, for data mining, we can use limited ways of collecting necessary information, such as the processing of data from active experiments (testing, diagnostics) or passive observation, which refers to the identification of control objects [15]

We believe that the most informative and predictive for the education success monitoring are information-graphics models based on integral-generalized indices and the structure of the dynamics of learning success indicators, extrapolation and phase models that allow to adequately, clearly and easily build an operational, medium-term and long-term forecast the success of teaching quality management in academic quarters, semesters, a year and long-term dynamics; to highlight and analyze interdisciplinary relationships, to ensure the effectiveness of the training programs implementation, motivation to learn, the effectiveness of the situational management functions implementation.

The effective use of AI technologies in the field of higher education will make it possible to select the most optimal learning strategy, adapted to the student’s individual abilities and labor market needs.

References
[1] Kotova E E 2019 Predicting the success of learning in an integrated educational environment using online analytics tools Computer Tools in Education 4 55-80
[2] Minina V N 2020 Digitalization of higher education and its social results. Bulletin of St. Petersburg University Sociology 13(1) 84-101 https://doi.org/10.21638/spbu12.2020.106]
[3] 2018 European Skills Index Technical report. URL: https://skillspanorama.cedefop.europa.eu/en/useful_resources/2018-european-skills-index-technical-report (appeal date: 15.09.2020)
[4] Organisation for Economic Cooperation and Development. Skills for a digital world. Background report 2016 Ministerial Meeting on the digital economy. 2016. № 250. 56 p. doi: 10.1787/5jlwz83z3wnw-en
[5] Graesser A C et al. 2018 Challenges of assessing collaborative problem solving Assessment and teaching of 21st century skills. Springer 75-91 doi: 10.1007/978-3-319-65368-6_5
[6] De Wever B et al. 2006 Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review Computers & education 46(1) 6-28
[7] Shum S B and Crick R D 2012 Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics Proceedings of the 2nd international conference on learning analytics and knowledge (ACM) pp 92-101 doi: 10.1145/2330601.2330629
[8] De Liddo A et al. 2011 Discourse-centric learning analytics Proceedings of the 1st international conference on learning analytics and knowledge. ACM 23-33 doi: 10.1145/2090116.2090120
[9] Siemens G and Baker R S J 2012 Learning analytics and educational data mining: towards communication and collaboration Proceedings of the 2nd international conference on learning analytics and knowledge. ACM. 252-4 doi: 10.1145/2330601.2330661
[10] Romero C and Ventura S 2010 Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 40(6) 601-18
[11] Sclater N, Peasgood A and Mullan J 2017 Learning analytics in higher education (London: Jisc) Retrieved from: https://goo.gl/g0roCB (appeal date: 15.09.2020)

[12] Learning Analytics Retrieved from https://www.e-teaching.org/didaktik/qualitaet/learning_analytics (appeal date: 15.09.2020).

[13] Shum S B 2012 Learning analytics policy brief. UNESCO Institute for Information Technology in Education 12

[14] Dmitrievskaya N A 2011 Integrated intellectual environment for continuous development of competencies Open education 3 4-8