European Scholar Journal (ESJ)
Available Online at: https://www.scholarzest.com
Vol. 4 No.05, May 2023
ISSN: 2660-5562

EDUCATIONAL DATA MINING AND LEARNING ANALYTICS

Saydakhmedova Barno Batirovna
INTERNATIONAL ISLAMIC ACADEMY OF UZBEKISTAN,
teacher of the department of modern ICT,
batirovna.b88@gmail.com
tel: (94) 617-69-78

Article history:

Received: 20th March 2023
Accepted: 28th March 2023
Published: 28th May 2023

Abstract:

The relevance of the study is due to the fact that on the topic in the context of the intellectual analysis of educational data there are currently more questions than specific answers: how it is done, why and how we can use it, what metrics to include in the sample and how to make forecasts. Undoubtedly, in the coming years there will be a transition from discussions to the practical implementation of learning analytics in educational processes. The purpose of the study is to systematize the methods of intellectual analysis of educational data in the context of the difference between educational analytics and pedagogical diagnostics and other methods of data collection. The results of the study will help to build a learning strategy and combine the goals of the training program with the effectiveness of the educational process and the expected results from the students. In this regard, the author considers the types of educational analytics. The scientific novelty of the research lies in the systematization of areas of research interests related to data mining in education and educational analytics. It is substantiated that educational analytics in conjunction with the intellectual analysis of educational data makes it possible to develop accurate models that characterize the behavior of students, their properties, weak and strong points of content and interaction with it, team and group dynamics. The practical significance of the study lies in the fact that the considered methods will allow us to assess the current state of the training system or program, predict the desired results and draw up a roadmap for planned changes. For pedagogical designers and methodologists, the presented methods will become the foundation for optimizing the program.

Keywords: education, digitalization, educational analytics, educational trends, intellectual analysis, training, continuing education, skills, educational data, effectiveness

INTRODUCTION

The relevance of the study is due to the fact that on the topic in the context of the intellectual analysis of educational data there are currently more questions than specific answers: how it is done, why and how we can use it, what metrics to include in the sample and how to make forecasts. Undoubtedly, in the coming years there will be a transition from discussions to the practical implementation of learning analytics in educational processes.

The subject of the research is the educational process. The purpose of the study is to systematize the methods of intellectual analysis of educational data in the context of the difference between educational analytics and pedagogical diagnostics and other methods of data collection. The results of the study will help to build a learning strategy and combine the goals of the training program with the effectiveness of the educational process and the expected results from the students. In this regard, the author considers the types of educational analytics.

The scientific novelty of the research lies in the systematization of areas of research interests related to data mining in education and educational analytics. It is substantiated that educational analytics in conjunction with the intellectual analysis of educational data makes it possible to develop accurate models that characterize the behavior of students, their properties, weak and strong points of content and interaction with it, team and group dynamics.

Two sets of methods are widely used in the analysis of educational data:

- Learning Analytics;
- Educational Data Mining (EDM).

Learning analytics, or educational analytics, is the measurement, collection, analysis and presentation of data about students and the educational environment in order to understand the characteristics of learning and optimize it as much as possible.

Educational Data Mining (EDM) is a range of methods to explore data and find hidden patterns in it for the purpose of making educational decisions. Educational Data Mining includes both methods typical of any Data Mining (data mining) (clustering, classification, regression, search for rules), as well as those specific to the field of education, for example, psychometric models.

To generate and test hypotheses, it is necessary to determine the appropriate metrics. In addition, it is necessary to decide what data will be needed, where it will come from, and for what analytical purposes it will be collected.

Learning analytics, together with EDM, helps to uncover relationships in complex data and hidden knowledge, which, when properly approached, can help improve the learning process. The main difference between these two sets of metrics is that in educational analytics insights and findings are generated primarily by a specialist analyst, while in Educational Data Mining many insights become "unexpected" results of using automated algorithms.

The results of the study will help to build a learning strategy and combine the goals of the training program with the effectiveness of the educational process and the expected results from the students.

The practical and theoretical significance of the study lies in the fact that the considered methods will allow us to assess the current state of the training system or program, predict the desired results and draw up a roadmap for planned changes. For pedagogical designers and methodologists, the presented methods will become the foundation for optimizing the program. Learners through the presented methods receive the most relevant, engaging and meaningful educational experience.

METHODS.

The revolutionary nature of Big Data can be explained by three classic characteristics that are attributed to the processes associated with big data: it is not only volume (Volume), but also speed (Velocity), as well as variety (Variety). By speed, we mean the ability to process incoming data as quickly as possible, almost in real time, and by diversity, we mean the sources of these data, different in quality and properties.

What did the advent of big data mean for education? Thanks to the advent of teaching information systems and Big Data technologies, for the first time in history, pedagogy got a chance to quickly, continuously and in full register a vast array of observations of the learning process, behavior and progress of students.

The gap between the identification of a problem and its solution has decreased many times over; the possibilities for interpreting and analyzing the information received have significantly expanded. Moreover, learning analytics systems have made it possible to automate many routine processes, identify problems at an early stage, and act proactively.

There are several categories of data that are typically used for learning analytics purposes. The list below is not exhaustive, since the list of data sources is adapted to the goals and needs of a particular organization:

1. Administrative data - information about the teacher, methodologist, availability and options for support, the experience of the author, the subject of the program or course, etc.

2. Preferred learning media or genres - retrospective measures of the learner's preferred media or genres where a choice was possible, e.g. video length, podcast listenability, longread readability, etc.

3. Interaction with educational resources - indicators of interaction during training, including the manner of navigation, answers to exercises and tests, the number of attempts, types of errors made, temporal characteristics associated with the student's activities during training events.

4. Past activity - retrospective indicators of the student's past activity, revealing the assimilation of ideas, skills or competencies at the current moment.

5. Temporal history - indicators of the immediate context, representing the temporal history of the student's actions, data about which is available on a particular day.

6. Social indicators - an indicator of the student's interaction with other students and the teacher in the learning process or with recorded speech (with all its various properties, for example, semantic content, prosody, etc.).

7. Demographic information - indicators of the peripheral context: region, age, gender, level of training, etc.

8. Social ties - indicators of the immediate environment: the number of ties, their strength and activity.

9. Type of thinking - data from the questionnaire or self-report on how the student establishes a connection between their strategic efforts during training and the development of competencies, as well as how the individual learning process occurs.

10. Emotional state - psychophysiological indicators related to learning, for example, emotional state, sleep quality, nutritional indicators.

The modern learning process involves the use of many educational technologies and platforms, most of which are equipped with their own tools for collecting statistics and analytics. We list in general terms some data sources:

– LMS (Learning Management System, learning management system), such as Moodle, Canvas, iSpring, open EdX, etc.;

– LXP (Learning Experience Platform, learning experience platform, or aggregators of educational projects), such as Degreed, EdCast;

– TMS (Training Management System, corporate training management system), such as TrainingSpace;

– video conferencing services, such as Zoom, Skype, MS Teams; - webinar platforms, such as Webinar.ru;

– survey services such as Kahoot!, Quizizz, Slido, StartExam; - courses created through constructors, for example, using Articulate 360, iSpring Suite, Adobe Captivate;

– VR and AR platforms like Uptale, Microsoft Mixed Reality.

Each of these sources has an internal analytics system, which complicates the centralized collection of data and their subsequent analysis. Partially this problem is solved by LRS - Learning Record Store, storage of educational records. Some LRS systems use the xAPI standard to transfer data from different sources and store it in a common format.

xAPI is an open specification that describes the format for passing statistics between a learning activity provider (for example, a course, website, mobile app) and the LRS database. It is important here that the learning activity provider return data in the format provided by xAPI.

There is the concept of Total Learning Architecture (TLA), which is being developed as part of the ADL Initiative program of the US Department of Defense (authors of the SCORM standard). TLA includes a set of technical specifications to build future learning ecosystems. One such specification is xAPI, a standard for transferring data about completed training. This standard is supported by many popular LMS systems (WebTutor, Moodle), course builders (Articulate Storyline, iSpring) and other educational services.

The basic unit of the xAPI standard is a statement (from the English statement). A statement is a record in JSON format about some action of the form "The user has learned lesson 1". The number of statements can reach several billion. To collect data in xAPI format, you need a special storage - the Learning Record Store (LRS). LRS validates the incoming data for compliance with the standard, processes, sorts and provides for the formation of visual custom dashboards in specialized educational analytics systems or universal BI tools (PowerBI, Tableau, Qlik).

Some LRSs have built-in learning analytics, come with add-on applications (for example, to launch eLearning courses), or are part of an LMS. Popular LRS systems are proven in the educational analytics market.

External learning analytics platforms and learning record repositories provide rich opportunities for data aggregation from various sources, their analysis and visualization. It should be borne in mind that when using such solutions, dedicated resources are required for installation, support, and most importantly, scaling.

RESULTS.

Typically, curriculum changes are made for three complementary purposes:

1. To improve the listener experience – learner's point of view.
2. To optimize the curriculum – the point of view of the methodologist / pedagogical designer.
3. To improve the efficiency of the organization the point of view of the customer training).

The selected goal and determines – the relevant data sources and indicators whose changes will be monitored and evaluated. Contexts in which different metrics are used can overlap in goals, so you need to clearly map metrics to goals.

A metric is just a numerical indicator, and apart from context and goals, it remains just a number. The metric turns into a flag that signals the need to change something, confirm the effectiveness or, conversely, refute the hypothesis, thanks to analytical work. Learning effectiveness metrics help to assess the quality of educational content and its individual parts, the level of training of teachers and presenters, and the progress of students.

Educational products for obtaining a new profession (vocational training) are provided by many market players: universities, organizations of secondary vocational and additional education, EduTech companies, online universities, etc. What metrics can be used to assess the success of retraining students who only yesterday worked as miners, confectioners or drove a tram? Simple solutions are not suitable: for example, you can't take accessibility as the main metric, because the goal of the student is not to reach the end of the course, but to get a new job in a new specialty.

For a student of the program with a request to change profession, the result is important - getting a job in a clear time with a clear income. To measure and track such a result, indicators such as:

– the speed of receiving an offer;
– number of proposals;
– starting salary.

When developing and assembling programs for each profession:

– analyzes the vacancy market (number and dynamics of opening/closing of positions) and data on starting salaries of employees of initial positions;
– a list of required competencies is compiled both for the analysis of open positions and based on the results of 30–50 interviews with HR and hiring managers;
– together with market experts, a program is compiled and validated.

This affects the content of the program: it introduced market-relevant business challenges to immerse newcomers in the industry into the market and the specifics, and added career counseling, mock interviews, resume preparation assistance, retrospectives with alumni after initial responses, and interviews. As a result, it is advisable to focus on three sets of metrics: internal product metrics, educational program effectiveness metrics, and market metrics for student employment and recall.

DISCUSSION.

Learning analytics in conjunction with the intellectual analysis of educational data makes it possible to develop accurate models that characterize the behavior of students, their properties, weak and strong points of content and interaction with it, team and group dynamics. Data collection for these purposes is carried out from a variety of sources, which requires additional resources for organizing the storage and access to this data, and their subsequent processing for analytical purposes. This problem is partially solved by the storage of educational records - LRS, however, in any case, one cannot do without specific knowledge in the field of engineering and data analysis.

The study presents the performance metrics of the educational product. These indicators help to assess the quality of educational content and its individual parts, the level of training of teachers and presenters, and the progress of students.

External educational analytics platforms and educational record repositories provide ample opportunities for data aggregation from various sources, their analysis and visualization. It should be borne in mind that when using such solutions, dedicated resources are required for installation, support, and most importantly, scaling.

In the learning strategy, it gives confidence that the decisions made will correspond to the strategic goals of the educational organization. For those responsible for learning, learning analytics, combined with educational data mining, provides a framework for prioritization, decision making, and efficient resource allocation.

For pedagogical designers and methodologists, the presented methods will become the foundation for program optimization. Learners through the presented methods receive the most relevant, engaging and meaningful educational experience.

REFERENCES:

1. Ширинкина, Е. В. (2022). Методы интеллектуального анализа данных и образовательной аналитики. Современное образование, (1), 51-67.
2. Saydakhmedova B. B. THE IMPORTANCE OF E-LEARNING COURSES IN THE EDUCATIONAL PROCESS //Актуальные научные исследования в современном мире. – 2019. – №. 10-5. – С. 94-99.
3. Saydakhmedova B., Makhkamova M., Rakhimova S. Content of innovation and innovative activity of the teacher //International Journal of Advanced Science and Technology. – 2020. – Т. 29. – №. 7. – С. 665-669.
4. XODJAYEVA M. Jumayev, T., Dadamuhamedov, A., & Saydakhmedova, B. CREATION OF MOBILE APPLICATIONS FOR THE SHRINES OF AL-HAKIM AL-TERMIZI //The Light of Islam. – 2020. – Т. 2020. – №. 2. – С. 176-182.
5. Saydakhmedova B. TA’LIMDA MA’LUMOTLARNING INTELLEKTUAL TAHLILI //Евразийский журнал математической теории и компьютерных наук. – 2023. – Т. 3. – №. 2. – С. 15-19.
6. А.О.Блинов, О.С.Рудакова, Е.Н.Благирева: Интерактивные методы в образовательном процессе. Учебное пособие. ИД Научная библиотека, 2014 г.