

INTELLECTUAL ANALYSIS OF EDUCATION DATA

<https://doi.org/10.5281/zenodo.10144027>

A.A.Khakimov

*Fergana branch of TUIT named after Muhammad al-Kharizmi,
assistant of "Information Technology"*

Introduction

The relevance of the study is due to the fact that there are now more questions than specific answers on the topic in the context of educational data mining: how it is done, for what and how we can use it, what metrics to include in the sample and how to make predictions. Undoubtedly, in the coming years there will be a transition from discussions to the practical introduction of educational analytics into educational processes. The study analyzes these categories, and fundamentally differences educational analytics from pedagogical diagnostics and other methods of data collection. One of the objectives of the study is to build a model for individual educational solutions, provided that the data is well-assembled, so the study considers the types of educational analytics. The practical significance of the study is that all this can lead to a change in editorial educational policy, the language in which communication with students takes place, which is already a strategic task of learning.

The relevance of the study is due to the fact that there are now more questions than specific answers on the topic in the context of educational data mining: how it is done, for what and how we can use it, what metrics to include in the sample and how to make predictions. Undoubtedly, in the coming years there will be a transition from discussions to the practical introduction of educational analytics into educational processes. The study analyzes these categories, cites the fundamental difference between educational analytics and pedagogical diagnostics and other ways of collecting data. An analytical review of educational analytics studies [1-4,10] showed, that for the first time in 1995, Corbett and Anderson conducted a study of 'Bayesian Knowledge Tracing', a key early algorithm of educational analytics that has not lost its popularity to this day. In 2000, the first Workshop associated with EDM is created. In 2001 - Theoretical study of Dina Zayan on EDM methods. In 2005, it was the first workshop to use the term 'educational data mining'. In 2006, the first published book about EDM: 'Data mining in e-learning, Romero and Ventura'. In 2008, it was the first international

conference on the analysis of educational data (EDM). In 2009, the first issue of the Journal of EDM was published, and to date it has been quoted 189 times (15.75 quotes per article annually). In 2010, the first directory "Educational data mining, EDM" was published by the authors: Romero, Ventura, Pechenizkiy, and Baker. In 2011, the first conference "Analysis of Learning and Knowledge" was held. The International Symposium of Electronic Devices and Materials (IEDMS) was founded. In 2012, SoLAR, a community of educational analysts, was founded. In 2013, the first summer institute of learning analysis was established. In 2017, data analysis and extraction are considered necessary for a full-fledged educational study, and it is expected that by 2022 it will become mandatory.

So, first, in normal diagnostics, data is collected based on the responses of students through questionnaires, surveys, while in educational analytics they are recorded in information systems, transmitted via a computer. In pedagogical diagnostics, the work of the teacher is analyzed, and in educational analytics-first of all, the students, in order to predict the success or failure of training. But the main difference is that data collection in the first case takes place in accordance with a special procedure at a certain point in time (usually at the end of training), while in training analytics it is carried out constantly, continuously.

Accordingly, in traditional diagnostics, the results of surveys and questionnaires are a secondary, auxiliary tool, and decisions are made not only on the basis of data, but taking into account subjective judgments and so on. But everything is changing, and above all, thanks to the phenomenon of big data - today computing power is not comparable to those of 30-50 years ago. There is confidence that soon there will be the concept of data-driven company - an organization whose management is based on data, and analytics is integrated into all operational processes.

Intellectual analysis of educational data allows you to get away from this, especially in such a question setting always the answer, socially expected, while real information is needed. Intellectual analysis of educational data uses analytics to improve content in order to achieve the goal of learning. There is no single mathematical model for assessing the impact of learning on the effectiveness of educational results, because each educational decision has its own goals and influence on the development of competencies and so on.

One of the objectives of the study is to build a model for individual educational solutions, provided that the data is well-assembled, so the study considers the types of educational analytics.

The practical significance of the study is that all this can lead to a change in editorial educational policy, the language in which communication with students takes place, which is already a strategic task of learning.

The empirical basis of the study is analytical research Gartner Says Advanced Analytics Is a Top Business, IBM SPSS Modeler CRISP-DM Guide Priorit, Mayer-Schenberger V., Kukier K., and the studies of M: Mann, Ivanov, and Ferber.

Learning analytics, or learning analytics, is to measure, collect, analyze, and present student and educational environment data to understand and optimize learning and the conditions in which it occurs. Whereas intellectual educational data mining is the process of finding patterns in large data sets, including methods at the intersection of artificial intellect, machine learning, statistics and database systems.

Timely and sufficient intellectual analysis of educational data allows you to optimize educational solutions.

The user of the online course creates information during the training process, which is accumulated in the database of the training management system (LMS).

How to analyze this data? Gartner Consulting Company (Gartner, Inc.) offers four types of analytics [9,13,15]. Of these, the first two types are an analysis of what has already happened, and the next two are forecasts and prescriptions for future actions. Thanks to this approach, the decision-making process is increasingly automated when moving from identifying a problem to eliminating it. Learning analytics determines the strategy at the stages of implementation and evaluation of the effectiveness of the educational solution. Descriptive analytics answers the question "What happened?" and allows you to identify the problem based on simple data: the time of passing a particular stage of the course, the quality of responses, and so on. Diagnostic analytics identify the cause of the problem. For example, why did 70% of users give incorrect answers? The analysis will show what led to this result: the content, how it is presented, the conditions for completing the course, or the low basic level of users. The third type is predictive, or predictive, analytics. By answering the question "What will happen next?", you can predict how motivated users are to complete a module or course. And finally, prescriptive analytics. On its basis, decisions are made about actions after completing the course: how to change a course or a separate module, what new courses to offer the user.

There is a model in which data is raw. Taking it as a basis, a person forms information, further compares information with knowledge about the world and makes certain decisions. Consider the example of a bored programmer over a

cup of coffee: armed with a thermometer, he measures the temperature of the drink, translates the numbers into specific degrees, compares the indicators with the ideal temperature and decides to send the office coffee machine in the trash. In the world of big data, the decision for it would be made by artificial intellect, and much faster. Powerful computing technologies capable of analysing vast amounts of data emerged between 1996 and 2015. They have been changing the market ever since. Previously, the future development of the company was universally determined by subjective judgments. The reporting included small data from internal sources and played a secondary role. Now the so-called data driven organizations are developing, where analytics is integrated into all operational processes.

This allows you to analyze specific points in the operating process, identify the causes of problems, find out how certain trends will affect production volumes, product movements, and optimize performance. Statistical analysis, forecasting, triggers, predictive modeling and optimization speak to the maturity of analysts in the organization. CRISP-DM Decisions are made faster, but behind them is painstaking preparatory work. First, the data must be created: a bank employee, for example, must enter the number and series of the client's passport. These, as well as data from external sources, need to be collected in one place. Next, select the ones you need for a specific task; Transform into an analysis-friendly format, clean up bugs and other noises. assess the accuracy, relevance, completeness and consistency of the information. Only then can artificial intellect form a report, identify patterns and make recommendations. According to the book "Big Data" by V. Mayer-Chenberger, the first five stages account for 80% of the work.

All education is based on the principle of feedback. The process of learning itself will not happen if a person does not ask himself the question: "What have I learned?" How do we usually remove feedback? Often, any type of survey is conducted: what did you like?

Intellectual analysis of educational data allows you to get away from this, especially with such a statement of the question, the answer is always socially expected, while real information is needed. Educational data mining uses analytics to improve the content in order to achieve the learning goal.

There is no single mathematical model for assessing the impact of learning on the effectiveness of educational outcomes, because each educational decision has its own goals and its impact on the development of competencies, and so on. Of course, one of the tasks is a similar model for individual educational solutions,

provided that the data is collected qualitatively. Of course, we need to talk more about predictive analytics, but this can only be done when data is available and when some models have already been tested.

In most cases, business problems are solved using data by CRISP-DM (Cross-Industry Standard Process for Data Mining) [8,20]. It involves six successive steps: business analysis, data analysis, data preparation, modeling, evaluation of results and their subsequent implementation. In the first stage, the objectives of the analytical project are set, risks are assessed, the amount of resources spent is determined, metrics for evaluation of results are determined, quality criteria are prepared and a work plan is prepared. The second is to determine what data is available and how it can solve the problem. On the third- the analyst receives the final set of data. Next, the modeling technique is chosen, the model is built and evaluated. The final stage is planned to monitor and support the implementation of the model and prepare a final report on the project. You can skip some of the steps depending on the problem you're up to.

It is important to remember that the data processing process presents many problems: the lack of a data search tool and clear procedures for providing access to data, the need to coordinate the purchase of data that may be heterogeneous, contradictory, uncertified and of poor quality.

REFERENCES:

1. Qadamova, Z., Khakimov, A., & Sotvoldieva, D. (2023). APPLICATION OF LIST METHODS IN PRACTICE AND ITS ADVANTAGES. *Лучшие интеллектуальные исследования*, 7(2), 43-47.
2. D.B. Sotvoldiyeva, A.A.Hakimov, & Z.E.Qadamova. (2023). PYTHONNING NUMPY MODULI YORDAMI BILAN CSV FAYLLARNI O'QISH. *Лучшие интеллектуальные исследования*, 7(2), 39-42.
3. Isroilovich, H. A., & Abdimahamatovich, H. A. (2023). KIBERJINOYAT JAMIYAT UCHUN YANGI TAHDID SIFATIDA. *World scientific research journal*, 15(1), 249-252.
4. Обухов, В. А., & Хакимов, А. А. (2022). ОСНОВЫ ИСПОЛЬЗОВАНИЯ РЕКУРСИВНЫХ ФУНКЦИЙ В СТРУКТУРАХ ДАННЫХ. *Journal of new century innovations*, 11(1), 92-99.
5. Sotvoldieva, D. B. (2023). DISKRET KONVOLYUTSIYANING MATLAB DASTURIDAGI TAHLILI. *Educational Research in Universal Sciences*, 2(10), 245-249.

6. Хусанова, М. К., & Сотволдиева, Д. Б. (2020). ИСПОЛЬЗОВАНИЕ ДЕЦИМАЦИИ И ИНТЕРПОЛЯЦИИ ПРИ ОБРАБОТКЕ СИГНАЛОВ В ПРОГРАММЕ МАТЛАВ. In *ЦИФРОВОЙ РЕГИОН: ОПЫТ, КОМПЕТЕНЦИИ, ПРОЕКТЫ* (pp. 970-975).