International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-3, February 2020

Vectorization and Optimization of User Behavior Data in E-Learning Systems

Krupa Tatiana

***Abstract*: At the first stage, an applied scientific research developed a procedure for collecting data on the parameters of user interaction with the user interface. This input procedure receives many heterogeneous messages about the actions of a particular user in the interface, while the output represents a vector that describes the user in aggregated form. The set of vectors for different users, in turn, was then used as input for the k-means clustering algorithm, the result of which is the user's attitude to one of the k clusters that distinguish the user by the type of behavior.**

User interface interaction data is available to 67.8% of GlobalLab platform users. There is no such data for the Diary.ru electronic diary. Considering that not all users of the GlobalLab platform took measures to create a project, ideas, work with questionnaires and educational materials, the total number of students for whom the value of all 4 variables differs from the neutral one was 9.7 thousand.

***Keywords*: education, learning trajectory, student, educational process, mathematical model.**

I. INTRODUCTION

Adaptive Learning Technology uses a "non-linear approach to learning and debugging, adapting to student engagement and demonstrated level of performance, and then predicting what types of content and resources students need at some point to succeed". Then a kind of automated teacher. The exciting thing is how quickly major textbook publishers have embraced this technology:

- Pearson teamed up with Knewton, which allowed Pearson to offer its MyLab and Mastering adaptation tools for a wide range of subjects, mainly in the sciences.
- McGraw-Hill introduced ALEKS and Smartbooks, the newest adaptive LearnSmart technology.
- New Ventures Macmillan has partnered with Knewton and has access to PrepU technology.
- Wiley has partnered with Snapwiz to create a new product offering called WileyPlus [1, 5, 9, 14].

There are other companies and even universities here:

- Companies include Smart Sparrow, CCKF and ScootPad.
- Brightspace from D2L has acquired the launch of Knowillage and adaptive LeaP learning technology.

Revised Manuscript Received on February 05, 2020.
* Correspondence Author
Krupa Tatiana*, GlobalLab, LLC, Moscow, Russia. Email: krupa.tatiana@mail.ru

In Europe, about a dozen institutions have formed INTUITEL to "improve the content of the e-Learning and Learning Management System (LMS) with features that so far have only been provided by human educators".

The University of Phoenix has invested heavily in its adaptive learning technology.

Ohio State University professors have created an adaptive learning module (called MOOCulus) that they have embedded into the Coursera platform for their MOOC calculation.

Adaptive technology has established the beach in higher education. Significant early projects are the use of MyLab Pearson by Arizona State University and the use of ELEXF ALEKS. The reports on these projects are ambiguous, as might be expected when young technologies are just beginning to develop, but the set of significant interests and investments promise to make this key technology a viable future.

II. METHODOLOGY

Education Data Education (EDM) is an interdisciplinary research area that is concerned with developing research methods for data that emerge in a learning context. EDM uses computational approaches to analyze educational data to explore educational issues [21].

The method of electronic educational systems to obtain additional information on the user's behavior has been developed. This approach is part of a pattern of behavior, typology, algorithm, the session was characterized by a set of variables in the course of pupils' behavior Sx in Section 3 of the introduction of plants and was cleared by the fact that the session is stored in the warehouse.

The main objective of the algorithm in any level of physical activity is a set of measures (storage PhisEvent entity) session is part of a document (ESO session, VI, and CM) length of a Session of variables, including the value of Sx [7, 13]. Watching the events on stage at the 1st session of command and used as a type of input. Nestled in a hierarchical structure, and the session of the Session, that person can be. In the algorithm of the session variables that reflect the user's behavior pattern Sx, is a Tuple. for all users, for example, Tuples sources on the internal variable is used to determine the UbehPrj, k-means clustering algorithm is used to learning UbehData, UbehStd, Ustyle [6, 7, 9]. Thus, the algorithm for the introduction of a certain degree of time-ordered set of events during one session. If a line in the direction of (tuple) is a representative example of the code of conduct and style of learning.

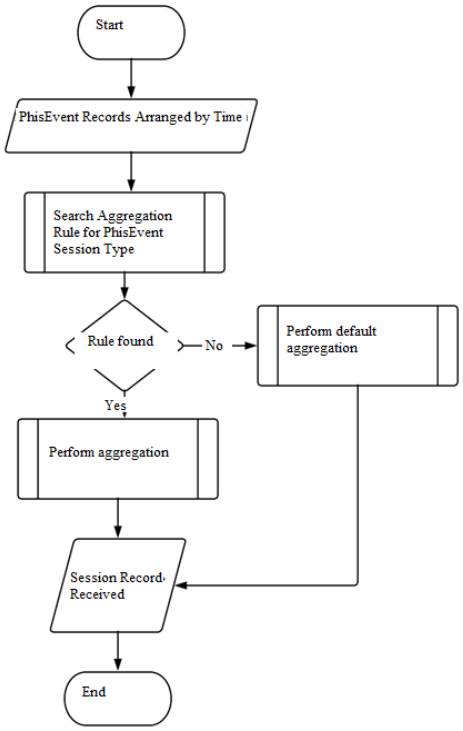

Fig. 1. The algorithm of the session variables that reflect the user's behavior [1]

III. RESULT AND DISCUSSION

To illustrate how the algorithm works, we give the following example of processing PhisEvent physical events that occurred during a document-level session.

After opening the webpage page, the user has performed the (PhisEvent) actions corresponding to those listed in Table 4 of Section 9 of the PNI Report for the first stage [1]:

All events were handled by a private user model algorithm that found the appropriate rule for this session type.

As a result of applying this rule, the following snippet of session record will be obtained:

{
$S_{scrollVector}$: 0.15,
$S_{scrollLength}$: 0.24,
$S_{scrollCoverage}$: 0.87
}

The example above is much simplified: it is assumed that a user only performed scroll actions during a session. In real-life, sessions are mouse clicks, text selection and on-screen dragging [8, 11, 18]. For each of these events, the rules for converting and calculating the corresponding values of the Sx variables are searched.

As already shown in Section 1, the clustering of session type records is performed by the k-means algorithm.

Updating the Housing for User Behavior in Electronic Educational Systems

In step 1, a body of user behavior data is generated that includes all the information that can serve as input and predictions during a xMANN recurrent architecture training network [17, 19].

The case included:

- data already available in the GlobalLab database;
- data on user interaction with the user interface;
- data on academic success.

In addition, the corps includes student achievement data obtained from Diary.ru's electronic diary, which was integrated within the GlobalLab platform.

In the second phase of the PNI, the body of collected user behavior data in eLearning systems was updated and modified as follows:

1. Data already available in GlobalLab has been reduced to a format compatible with Model 1 of the input repository for an algorithm for predicting the optimal learning trajectory [1, 16].
2. Training performance data has been reduced to a format compatible with Model 2 of the Input Repository.
3. The order of integration with Diary.ru's electronic diary has been revised to provide additional information about students, educational institutions and grades.
4. Software implementation of procedures for converting the output data from sections 1 and 2 of the repositories into xAPI format (section 3 for storing the input data) has been developed.
5. Software has been developed to convert user-specific information during a session (a PhisEvent entity) into a session record [1, 2].
6. Software implementation of vectorization of event components describing the stage of the educational trajectory and methods of reducing the dimension of the vectors obtained (see Sections 4, 5 of this PNI report).

The method described above, based on empirical rules, on the one hand, is flexible enough to create a large number of simulated user models, and on the other hand, ensures the realism of the obtained models by observing general patterns in user behavior. in GlobalLab EOS. It seems obvious that, due to the low knowledge of the laws of user behavior in EOS, a more reliable method for obtaining simulated data is not yet possible.

In step 2, the PNI test method generated trial generation of a set of simulated data. For generation purposes, the method developed in Stage 1 was programmatically implemented as a utility, which in Stage 3 will be included in the experimental sample [3, 4, 10, 12]

Ensuring Deployment, Configuration, and System Administration of Cloud Infrastructure that Updates the Housing of User Behavior in Electronic Educational Systems

In Phase 1, PNI, using the resources of an industrial partner, deployed, configured, and systematically administered the cloud infrastructure to generate an array of user behavior data in eLearning systems (see Infrastructure Description in Section 12 of the PNI Report for Phase 1, Phase 2 [1], in the deployment, configuration, and administration of the system has been continued due to an update of the database of user behavior in e-learning environments.

The total number of services providing work to update the user behavior data set in e-learning systems in Phase 2 has been expanded to 160 [15, 20].

The services were implemented programmatically, which ensured closer integration with Diary.ru's electronic diary through calls to Diary.ru API methods. These services implement steps 1.1 - 1.9 of the input algorithm for the optimal learning path prediction algorithm [1, 20].

In Step 2, we transitioned from using the Google Geocoding API service to using the OpenStreetMaps Nominatim open source system, which required the deployment of this service on a separate leased server. The move to the use of the open source OpenStreetMap Nominatim system deployed on your own industrial server servers also allowed for a sophisticated version of the UgeoAdmLvl, Ulat, Uurb variable calculation method. Implementation of the methods was made possible by the fact that the number of requests for direct and reverse geocoding is now unlimited.

In the future, the use of the OpenStreetMap Nominatim service opens up additional opportunities for modeling the students' living environment as it allows:

- determine the number and distance of different objects of educational and cultural infrastructure from the student;
- population in the settlement where the student resides;
- natural features of the student's place of residence (altitude, type of landscape, average annual air temperature, etc.).

Deployment, configuration, and system administration of the cloud infrastructure was performed on dedicated dedicated FastVPS EXV-10 and EX-4 servers.

In the framework of the patent research, a search for a patent, an analysis of the current patent situation, development trends and development prospects in the field of computer programs for the formation, analysis and vectorization of the student's educational path, etc. were carried out. including information about his behavior as part of machine learning to automatically determine the various properties of a student's learning trajectory.

The inventions, utility models and samples adjacent to the business object were taken for the study.

As a result of the scientific and technical literature search, 14 abstracts of computer program registration certificates and 5 patents were selected [1].

An analysis of patent documents and computer registration certificates shows that there are no inventions to date that use specific machine learning methods to model the properties of learning pathways. Often, the object of protection is not a specific model, but a method of applying it to a more general utility model. Methods based on the use of specific inputs in models of learning trajectories and students are also protected.

IV. CONCLUSION

The method of extracting additional user behavior data developed in Step 1, developed an algorithm for generating a user behavior data array that is intended to automatically transform the User's specific physical activity data into one aggregate EOS session his type of behavior and teaching style.

Applied scientific research completed all necessary deployment, configuration, and system administration work for the cloud infrastructure to update the corpus of user behavior data in electronic education systems.

The body of the user behavior data was updated, both related to updating the database with new data and converting the data to a format compatible with the model for storing the input of the optimal learning trajectory algorithm.

The simulated data about the users used to train the RNS was generated using the software implementation of the method developed in step 1 of the applied scientific research.

ACKNOWLEDGMENT

Applied research described in this paper is carried out with financial support of the state represented by the Ministry of Science and Higher Education of the Russian Federation under the Agreement #14.576.21.0100 of 26 September 2017 (unique identifier of applied research – RFMEFI57617X0100).

REFERENCES

1. T. V. Krupa, *Theoretical studies of the performed for the Stages of the ASR tasks*. Moscow: GlobalLab, LLC, 2018.
2. L. Furukawa, "Trajectory of Learning Experience based on the Performance of Canada's Youth in Mathematics", *International Journal of Innovation in Science and Mathematics Education, 26*(6), 62–75, 2018.
3. G. Lipatnikova, A. S. Polyanina, "The formation of the target component of educational activities of students using decision-making techniques in the framework of the reflexive approach", *Problems and Methods of Teaching Natural Sciences and Mathematics: Materials of the III All-Russian Scientific and Practical Conference (Ekaterinburg, December 2007)* (pp. 194-197). Ekaterinburg: Publishing House of the Ural Institute of Economics, Management and Law, 2007.
4. G. Lipatnikova, "The creation on an individual educational thakectory as one of ways to reach students in decision making", *Fundamental researches*, 5: 108-110, 2009.
5. O. A. Granichina, "Mathematical models of the quality control of the educational process in the university with active optimization", *Stokhaisticheskaya optimizatsiya v informatike*, 2: 77-108, 2006. Retrieved from https://www.math.spbu.ru/user/gran/sb2/granolga.pdf
6. H. Choi, K. Cho, Y. Bengio, "Context-dependent word representation for neural machine translation", Computer Speech & Language, 45: 149-160, 2017.
7. M. Taguma, F. Gabriel, M. H. Lim, "Future of Education and Skills 2030: Curriculum anaktsis: A synthesis of research on learning trajectories/Progressions in mathematics", *OECD. EDU/EDPC(2018)44/ANN3*, 2019. Retrieved from https://www.oecd.org/education/2030-project/about/documents/A_Synthesis_of_Research_on_Learning_Trajectories_Progressions_in_Mathematics.pdf
8. "Transforming Our World, the 2030 Agenda for Sustainable Development", *United Nations General Assembly Resolution A/RES/70/1*, 2015. Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E
9. M. McCartney, *The Indian Economy*. Newcastle upon Tyne: Agenda Publishing, 2019. DOI: 10.2307/j.ctvnjbfk1
10. M. Ndlovu, "Modeling with Sketchpad to enrich students' concept image of the derivative in introductory calculus: developing domain specific understanding", 2008. Retrieved from https://www.researchgate.net/publication/323120414_Modeling_with_Sketchpad_to_enrich_students'_concept_image_of_the_derivative_in_introductory_calculus_developing_domain_specific_understanding

11. I. Azam, M. A. Hasan, T. Abbasi, S. Ahmed, M. Abbasi, "Education Technology Based Models of Teaching and Learning", *Proceedings of the 4th National Conference; INDIACom-2010 Computing For Nation Development, February 25 – 26, 2010 Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi*, 2010.
12. Duke University. "Machine learning predicts behavior of biological circuits: Neural networks cut modeling times of complex biological circuits to enable new insights into their inner workings." ScienceDaily. ScienceDaily, 2 October 2019. Retrieved from www.sciencedaily.com/releases/2019/10/191002165235.htm.
13. P. Sztajn, J. Confrey, P. H. Wilson, C. Edgington, "Learning Trajectory Based Instruction: Toward a Theory of Teaching", *Educational Researcher, 41* (5), 2012.
14. I. Gatopolos, "Vectorization: How to speed up your Machine Learning algorithm by x78", *Towards Data Science*, 2019. Retrieved from https://towardsdatascience.com/vectorisation-how-to-speed-up-your-machine-learning-algorithm-by-x78-times-faster-e330df8c9b27
15. E. Merzon, E. Galimullina, E. Ljunimova, "A smart trajectory model for teacher training", In *Cases on Smart Learning Environments* (pp. 164-187), 2019. DOI: 10.4018/978-1-5225-6136-1.ch010
16. M. N. Ryzhkova, "The mathematical model of the process of educayion management", *Bestnik Cherepovetskogo gosudarstvennogo universiseta, 6*: 41-47, 2015. Retrieved from https://cyberleninka.ru/article/n/matematicheskaya-model-protsessa-upravleniya-obucheniem
17. A. Freiberg-Hoffman, J. B. Stover, N. Donis, "Influence of Learning Strategies on Learning Styles: Their Impact on Academic Achievement of College from Buenos Aires", *Problems of education in the 21st century, 75* (1), 2017. Available: http://www.scientiasocialis.lt/pec/node/files/pdf/vol75/6-18.Freiberg-Hoffmann_Vol.75-1_PEC.pdf
18. H. Jafari, A. Aghaei, A. Khatony, "Relationship between study habits and academic achievement in students of medical sciences in Kermanshah-Iran", *Advanced in Medical Education and Practice, 10,* 2019. DOI: 10.2147/AMEP.S208874
19. P. V. Ovchinnikov, Mathematical Models and Instruments for Designing Adaptive Educational Trajectories for Preparation of Competitive Specialists in Universities: Abstract thesis. Rostov-on-Don, 2014. Retrieved from http://economy-lib.com/matematicheskie-modeli-i-instrumentariy-proektirovaniya-adaptivnyh-obrazovatelnyh-traektoriy-dlya-podgotovki-konkurentosp
20. D. Sharma, "Artificial Intelligence and Big Data in Higher Education: Promising or Perilous?", *SmartDataCollective.* Retrieved from https://www.smartdatacollective.com/artificial-intelligence-and-big-data-in-higher-education-promising-or-perilous/
21. C. Romero, S. Ventura, "Educational Data Mining: A Review of the State of the Art. Systems, Man, and Cybernetics, Part C: Applications and Reviews", *IEEE Transactions, 40:* 601-618, 2010. DOI: 10.1109/TSMCC.2010.2053532