Improving the efficiency of developing meta-subject scenarios in the Moscow Electronic School by means of educational analytics

Ekaterina Lavrenova[1] and *Boris* Yarmakhov[1*]

[1]Moscow City University, Institute of Digital Education, Moscow, Russia

Abstract. One of the problems connected with the development of meta-subject skills in school students is the lack of publicly available educational materials with a focus on such skills. A possible solution is the formation of a collection of meta-subject materials stored in the library of learning scenarios of the Moscow Electronic School — now there are over 40,000 learning scenarios that have undergone moderation. This article presents the results of research that identified a cluster of teachers most inclined to create meta-subject scenarios and suggested recommendations for motivating teachers to create such scenarios. To achieve this purpose, a sample of authors of such scenarios published by the Moscow Electronic School were analyzed and clustered with the help of machine learning methods. As a result of this work, a gradient boosting algorithm was developed, which produced the best results. The clusters of users described as a result of the application of the algorithm followed five main behavior strategies in terms of the activity related to the creation of new scenarios. Teachers that are most likely to create meta-subject scenarios show interest in their colleagues' scenarios not only in their subject but also in other academic disciplines taught at school, willingness to copy and customize them. To develop teachers' readiness for the creation of meta-subject scenarios, it is recommended to conduct teacher training including their introduction to the best practices of developing such scenarios presented by the Moscow Electronic School. The research results are used in the development of a recommender system enabling easier search and navigation among the scenarios published by the Moscow Electronic School.

Keywords: recommender system, learning object repository, Moscow Electronic School.

1 Introduction

The formation of meta-subject connections in the educational process and meta-subject skills in students represents one of the key directions in the development of modern education. Meanwhile, diagnostics of the formation of such skills are hindered by the fact that the educational process at schools is still subject-oriented. At schools, there is a shortage of meta-subject tasks, usage of which in the educational process would encourage not only absorption

---

* Corresponding author: yarmakhovbb@mgpu.ru

of information on a certain subject but also the development of a wider range of skills. A possible solution to this problem could be more active usage of meta-subject materials created by teachers themselves and available via modern information systems.

One of such systems is the library of the Moscow Electronic School (MES). When creating a learning scenario for the library, the author indicates its category corresponding with a certain subject, which makes it easier to find it in the system for other teachers who are ready to use it. At the same time, the MES library of learning objects contains a certain number of meta-subject scenarios. What distinguishes them from other objects is that the authors mark them with two or more categories.

An increase in the number of such objects, their growing popularity among teachers, and their active usage in the educational process are the factors that encourage the formation of meta-subject skills in school students. However, the question as to what motivates teachers to develop such learning scenarios needs to be studied. In this article, we examine how meta-subject scenarios are created on the platform of MES and how such development is carried out by representatives of different types of authors producing such learning content. Identification of such interactions will allow us to determine potential authors of meta-subject content, increase the share of meta-subject materials on the MES library of learning scenarios, and create a conceptual framework for developing a recommender system for meta-subject scenarios in MES.

The activity of users of learning object repositories is an important factor in the analysis of web behavior, and education is no exception. The early 2000s became the turning point in the formation and development of learning object repositories when universities started publicly sharing their training courses on the web within the Open Courseware initiative [1]. It became clear that for educational organizations attracting new students to their resources has a much higher value than keeping their learning programs and materials secret. However, together with a growing number of learning objects available to students, a problem has arisen connected with navigation in a relatively unstructured space where these objects are located.

The number of learning objects stored in modern repositories is estimated at dozens or hundreds of thousands. For example, in the American repository MERLOT, there are over 20,000 objects and 70,000 users, and in OER COMMONS, there are over 18,000 resources. The Learning Resource Exchange from European Schoolnet contains over 43,000 resources from 25,000 providers [2-3]. In this context, the MES library, which contains about 50,000 learning scenarios used by over 100,000 users seems a significant project that is of interest both to users who implement these scenarios into practice and researchers. In repositories, the materials of which are used to construct the educational process, searching for a learning object that meets the requirements of the user becomes an independent research task, the solution of which calls for specialized instruments and services [4-5].

2 Methods

A specific feature of the MES library is that, as a rule, a repository contains a large number of objects with similar names and keywords in the text, and to offer the most suitable content to the user, the system needs to understand their deeper motives and preferences. This category of tools is called "recommender systems" [6].

Our work is aimed at the formulation of the main approaches to the development of recommender systems based on analysis of behavior patterns shown by the users of the MES learning object repository; in particular, it will allow such systems to recommend meta-subject learning objects to users.

When teachers are looking for the necessary lesson scenario, they are not always able to formulate clearly what they need, while the repository search system needs to know exactly

the name of the scenario, its author, and unique identification number (ID). In this case, there is a recommender system that analyzes the available information about the particular user and recommends the best option matching their implied ill-defined need [7]. For this to happen, the recommender system must predict which object is going to be the most popular with the user [8]. Different models of recommender systems are used depending on what data about the user are available and what the expected result should be [9].

The mechanism of a recommender system is based on the assumption that if a group of users had common preferences in the past and chose similar subjects, they will continue to make similar requests in the future [10]. Let us assume that teacher A and teacher B have chosen similar lesson scenarios in the MES library. If later teacher A chooses a scenario from the library that teacher B has not chosen yet, the user filtration system will offer it to the second teacher, since it is likely that this scenario will be suitable for this teacher too. The collaborative filtering method is considered to be a digital form of word-of-mouth communication based on the fact that people tend to choose things recommended to them by other people who they know and trust [11-12].

User-based neighborhood recommendation methods predict the rating R_{ui} of a user U for a new item I using the ratings given to I by users most similar to U. To determine which users' opinions should be used to predict ratings, the k-nearest neighbors (k-NN) method is applied [8].

3 Results

For the creation of a recommender system oriented at searching for meta-subject learning objects in the MES repository, clusterization of the system users is important, which would allow the system to recommend a learning object to a user depending on which cluster they belong to and the interests of which group they share [13-14].

In the framework of this research, to identify the key groups of users for the creation of a MES scenarios recommender system, clusterization of users has been conducted with the help of algorithms most commonly used for this end. For clusterization, a group of 7,431 users was selected, each of whom had created at least one scenario published in the MES scenarios library.

To identify user clusters, we analyzed their actions in the system most typical of the descriptions of the groups to which they belong:
- Creating a new object by the user;
- Viewing an object registered in the system by the user.

User clusterization was conducted based on specialized Python Scikit-Learn and XGBoost libraries with the help of four algorithms most commonly used for the solution of such tasks: the k-nearest neighbors method, support vector machines, gradient tree boosting, and a neural network. Accuracy, integrity, and F1-measurement parameters were applied to the data processing results and allowed us to assess how well each of the suggested algorithms copes with the task of clusterization.

Thus, with the help of a **neural network**, the obtained parameters amounted to the following values: the F1 — 0.84, the precision parameter — 0.82, and the integrity parameter — 0.88. These are the worst results out of those produced by the four applied methods.

The **support vector machines** algorithm produced somewhat better results: 0.96, 0.98, and 0.96 correspondingly.

The results shown by the **k-nearest neighbors model** were slightly higher: its weighted parameters amounted to 0.96, 0.95, and 0.96 correspondingly. As can be seen from the table, the model did not retrain to the same degree as the **neural network** and the **support vector machines** algorithm.

Finally, the best results were produced by the **gradient boosting algorithm**, which was subject to retraining by the dominant class to the lowest degree. According to our analysis, it was the gradient boosting algorithm that proved to be the best tool for creating a recommender system for the MES repository objects†.

4 Discussion

As a result of the clusterization of the MES library users, we identified five large clusters, representatives of which show different behavior patterns as authors of the MES scenarios library. Representatives of each cluster are described further with the indication of their share in the total sample.
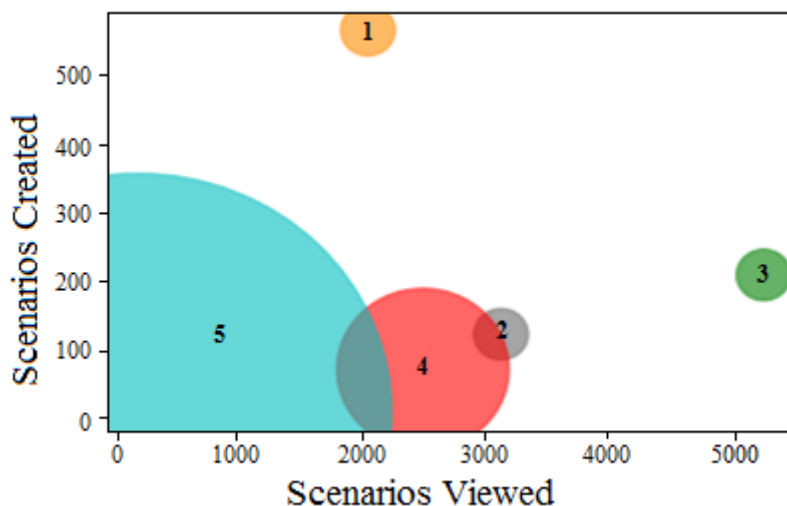
1 Authors — these users create content and lesson scenarios that are actively used by other teachers. Their share in the system is about 1%.

2 Methodologists — they bring the materials developed by the previous category of users (authors) to perfection and share them with other users (1%).

3 Mentors — they actively explore the MES materials, evaluate and use them in the educational process, and help young teachers to exploit the resources (1%).

4 Specialists — they use the MES materials in the educational process and create their own copies of learning materials (10%).

5 Practitioners — they use the MES materials to solve the tasks they face in the course of the educational process. They use the materials created by other teachers (87%).

Fig. 1. Clusters of MES authors.

At the next stage, we conducted a comparative analysis of the nature of scenarios created by representatives of different clusters.

---

†MES user clusterization was conducted with the participation of the Information Technologies Management of the Moscow City Pedagogical University.

Table 1. Authors of meta-subject scenarios.

	Mono-subject	**Meta-subject**
Practitioners	86%	14%
Specialists	72%	28%
Authors	70%	30%
Mentors	40%	60%
Methodologists	21%	79%

The majority of meta-subject scenarios are created by representatives of the "methodologists" cluster, i.e. the library users who do not create many scenarios of their own but view a lot of scenarios developed by other teachers. In general, this observation correlates with the patterns identified in the course of analysis of social networking sites, according to which the main activity in social interaction systems consists in viewing the content produced by other people rather than creating one's own content [15]. Conversely, low user activity in terms of viewing scenarios created by other teachers leads to a higher probability that such users will create a mono-subject scenario rather than a meta-subject one.

5 Conclusion

The task of developing meta-subject materials to be used in the educational process, including the ones designed for distance and blended learning, is a complex challenge. Its solution is possible by means of the application of modern data analysis methods, such as cluster analysis and machine learning methods. These techniques help to develop recommendations for teachers that will allow them to get necessary competencies for the development of meta-subject materials. The conducted research has shown that the most natural route to increasing the meta-subject nature of the materials created within the MES library consists in the development of professional ties between teachers, raising the interest of potential authors of MES scenarios in each other's works, and formation of a community of learning object users.

References

1. H. Abelson, Journal of Science Education and Technology, **17(2)**, 164–174 (2007)

2. H. Drachsler, K. Verbert, O.C. Santos, N. Manouselis, *Panorama of recommender systems to support learning,* in Recommender systems handbook (Springer, Boston, MA, 2015) http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-1-4899-7637-6_12

3. N. Manouselis, H. Drachsler, K. Verbert, E. Duval, *Survey and Analysis of TEL Recommender Systems,* in Recommender Systems for Learning, SpringerBriefs in Electrical and Computer Engineering. Springer, New York, 2013) https://doi.org/10.1007/978-1-4614-4361-2_3

4. E. Patarakin, V. Burov, B. Yarmakhov, *Computational Pedagogy: Thinking, Participation, Reflection,* in Digital Turn in Schools—Research, Policy, Practice, Lecture Notes in Educational Technology (Springer, Singapore, 2019) https://doi.org/10.1007/978-981-13-7361-9_9

5. S. Vachkova, E. Patarakin, E. Petryaeva, SHS Web of Conferences, **79**, 01017 (2020) https://doi.org/10.1051/shsconf/20207901017

6. S. Dwivedi, V.K. Roshni, *Recommender system for big data in education*, in 2017 5th National Conference on E-Learning & E-Learning Technologies (ELELTECH), IEEE, 3-4 August 2017, Hyderabad, India (2017). https://doi.org/10.1109/ELELTECH.2017.8074993

7. N. Hoic-Bozic, M. Holenko Dlab, V. Mornar, IEEE Transactions on Education, **59(1)**, 39–44 (2016). https://doi.org/10.1109/TE.2015.2427116

8. F. Ricci, L. Rokach, B. Shapira, Introduction to Recommender Systems Handbook (Springer, Boston, MA, 2015). https://doi.org/10.1007/978-0-387-85820-3_1

9. R. Burke, *Hybrid Web Recommender Systems,* in The Adaptive Web. Lecture Notes in Computer Science, volume 4321 (Springer, Berlin, 2007). https://doi.org/10.1007/978-3-540-72079-9_12.

10. G. Karypis*, Improving Higher Education: Learning Analytics & Recommender Systems Research,* in Proceedings of the Eleventh ACM Conference on Recommender Systems, SIGCHI, August 2017, Como, Italy (2017). https://doi.org/10.1145/3109859.3109870

11. M. Goga, S. Kuyoro, N. Goga, Procedia-Social and Behavioral Sciences, **180**, 1481–1488 (2015). https://doi.org/10.1016/j.sbspro.2015.cog02.296

12. J. Lin, H. Pu, Y. Li, J. Lian, Procedia Computer Science, **129**, 449–453 (2018). https://doi.org/10.1016/j.procs.2018.03.023

13. C. Obeid, I. Lahoud, H. El Khoury, P.A. Champin, *Ontology-based Recommender System in Higher Education,* in Companion Proceedings of The Web Conference, IW3C2, April 2018, Lyon, France (2018). https://doi.org/10.1145/3184558.3191533

14. A. Tagarelli, R. Interdonato, Mining Lurkers in Online Social Networks: Principles, Models, and Computational Methods (Springer, Cham, 2018). https://doi.org/10.1007/978-3-030-00229-9

15. T. Van Mierlo, J Med Internet Res, **16(2)**, e33 (2014). https://doi.org/10.2196/jmir.2966