The Academy of Sciences
Islamic Republic of Iran

Iranian Society of Engineering Education

The 6th Iran International Conference on Engineering Education,
19-21 November, 2019, Ferdowsi University of Mashhad, Mashhad, Iran.

A heuristic method for curriculum planning based on students' interest

B.Behkamal, B.Bolourian Haghighi, A. Tavakoli Haghighi

ABSTRACT

Nowadays scheduling courses in universities could be a challenge for university authorities due to time interferences, students' preferring and teachers' and student's limitations. This paper addresses the inefficient course schedule challenge and proposes a practical heuristic method to overcome it. Discovering students' favorite courses and courses that are picked together usually could be informative for university authorities. In this paper we try to discover students' favorite courses, using their educational data and correlation analysis and frequent pattern mining, and then, using this knowledge to help scheduling courses so that all students can have the optimal curriculum.

Keywords

Educational data mining, Learning analytics, Course scheduling, students' interest.

1. INTRODUCTION

Educational data mining is a hot field of study nowadays. With the fast-paced growth of educational data, it is possible to extract useful knowledge out of them. Knowing the optimal course schedule can be helpful for new students. Discovering trends can be illuminating for the university to plan each course perfectly. Learning about students learning model can be productive for teachers as they can use the best method of teaching. Effects of Massive Open Online Courses (MOOCs) can be evaluated and improved.

In recent years, a field of study has been created to use data capability for developing cities (named Smart city). Of course, we are far away from reaching smart cities goals, and thus more efforts from researchers are needed, and we need to pay attention to these researches and use them in city management. On the other hand, teachers face the challenge of knowing how much the students of a class are prepared for the lesson to teach the best set of topics. Authorities of the college have to consider many different factors when they are filling the empty time slots of the day and week with courses, and sometimes, the planning is not the best version of what it can be. Students do not always follow the official educational path that is advised to them by the university. Some of them might fail a course and would have to take it again next semester, causing them to fall behind the path. However, they can graduate too as they have some freedom of choice in picking courses. Detecting the path that is most taken by these kinds of students can be helpful to guide them back on track. These were some of the challenges that can be solved with the use of educational data generated by students, authorities and educational systems. Most of Massive Open Online Course (MOOC) systems have recommendation systems that help new students to pick a suitable set of course according to their preferences. In this paper, we try to figure out what courses are taken together to discover the best schedule for semesters.

The rest of this paper is organized as follows. The motivation of this work is described with an example scenario in section 2. An overview of related works is presented in section 3. Our proposed methodology is described in section 4. We analyze our results in section 5. Section 6 is dedicated to the conclusion, and the references are listed in section 7.

2. MOTIVATION

In universities, students can choose their preferred courses to form a wide variety of choices although they

B.Behkamal is Assistant Professor at Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: behkamal@um.ac.ir, tell: 09155027726).
B.Bolourian Haghighi is PhD student at Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: b.bolourian@mail.um.ac.ir, tell: 09155255832).
A. Tavakoli Haghighi is Master student at Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: amirtavakolihaghighi@mail.um.ac.ir, tell: 0939031158).

should not pick two courses that have the same day and time since they interfere with each other. This may cause some students to give up their preferred course, and they are forced to pick something that they are not willing to take which will affect their performance during the semester. On the other hand, authorities of the university are unaware of all of the students' specific situation and preferring. They just follow a pattern mentioned by the university and try not to schedule some courses with interference that are known to be picked together over a long time.

By analyzing the educational information of students over the years, we can extract the knowledge to see what courses usually are picked together and what courses are picked more than the others. This way, we can discover the actual educational path that the student would take instead of old routine that university published which may be overlooked by some students causing them to face the problem of insufficient schedule and performance reduction. By detecting the best curriculum, we can advise the authorities in charge of course scheduling to arrange the classes in the way that courses that are picked packed together, do not interfere with each other.

3. LITERATURE REVIEW

Nowadays Educational data mining (EDM) has become a popular subject of study because the volume of data is growing and the information that can be extracted from that data is quite important and useful, and it cannot be overlooked. The EDM can be divided into three groups based on the purpose of the study, including curriculum planning, result prediction, and concept mining which are explained as follows.

3.1 Curriculum planning and scheduling

This topic is about optimizing the process of scheduling the courses of the university and recommending the best course combination to students to enhance the performance and have a positive effect on students' grade at the end of the semester [1].

Aher proposed a framework for recommending courses in a MOOC system [2]. They used WEKA and machine learning algorithms to train a model to recommend courses. They used the data stored in the Moodle database, an open-source course management system. In this system, every click is registered in the log file, and by analyzing it, popular courses that are chosen together can be detected. First, the set was preprocessed, and those who did not pick any course were removed. Then the Apriori algorithm in WEKA was used to extract association rules and was presented in a table as the results. The author announced the study of the use of other data mining algorithms as future works.

In another work done by Bahargam in 2015, two tasks were said to be important in optimizing the performance of understanding concepts in the classroom or any MOOC systems [3]. One of them is grouping the students to benefit from corporations and the second one is to find the optimal course schedule. An NP-Hard problem was discussed to solve this challenge. They used a recursive greedy algorithm, with the idea of doing the job with maximum profit, was used. They used the BUCS data set, which included the information of 394 graduate students of Computer Science College of Boston University and also, 41 course title. For the missing values, they used the Graded Response model to generate the values. Another dataset, in bigger volume, was generated using GRM, so that the performance of this method can be studied on bigger volumes of data. The results indicated that the random algorithm had the lowest performance and with the growth of the group number, the profit grows correspondingly. If time increases, the value of profit for K-means become closer to their value. The results of generated dataset indicated that their method works better than K-means.

In another paper by Jugo in 2016, knowing the educational paths better and performing proper recommendation to students to increase the quality of learning was discussed [4]. The novelty of their work is in two parts: the first step was data preparation; path detection and path evaluation are purely independent. The second step was to test and evaluate the method on a real dataset. With a formula, the profit of all paths in the dataset is calculated, and then it is given to the SPAN algorithm. This method was evaluated in 2 fields with 31 and 69 courses with the cooperation of 30 and 20 students. The percentage of those following the recommended path was 5 to 47.

Backenkohler in 2018, proposed a modular approach which was a combination of predicting with the awareness of grades and time information and course orders [5]. The basis of their work is to use collaborative filtering to build a model which consist of two parts: course dependency graph and grade prediction. Thus, this model predicts not only the expected performance (such as students' grades) but also the expected preparedness (how can one benefit from a specific order of courses). To train model they used the educational data of the students of the computer science department of Saarland University. Their unique idea was to consider performance and preparedness independently and assigning different weights to each of these parameters. The challenge with this dataset is that students can register in the course, later after the beginning of the semester and the delay will not be registered. In this model, a personalized concept graph will be constructed which has two important components: concept dependency graph which shows the effects, of course, A on course B, and performance prediction which is calculated with collaborative filtering. After evaluation, results indicated that the concept graph shows better accuracy in recommending course than predicting grades. Also, students are more likely to follow the previous student paths rather than to pay attention to their abilities.

Zafra proposed a hybrid method in 2018, which was a combination of collaborative filtering and content-based filtering to detect the most popular courses to recommend to students in university course recommendation system. Importance and weight of each course were computed automatically with the genetic algorithm. The study was done on a dataset containing more than 1700 evaluation completed by graduate students of the computer science department of Cordoba University in 2016 and 2017 with more than 63 tiles of courses. For evaluation, two experiments were done. First, an optimization in course weights was achieved; then, the results were compared with other methods. As for future work, the author suggests studying a set of courses wider than the one in this study to discover better weighing for courses and parameter optimization and also, this method has the potential to be used in other fields of study [6].

Since our main goal is related to curriculum planning, we briefly compare the works done on that topic in Table I.

3.2 Concept Mining

A usual challenge faced by teachers is to figure out a proper ordering of course concepts to present in virtual or real classes. Therefore, concept mining is one of the tasks of EDM which helps to discover information such as the best order or combination of concepts. On the other hand, concept mining can also be useful for students as well. By presenting a ranking and categorizing lessons and concepts to new students, it will become easier for them to pick them or plan how much time is suitable to spend on each of them [7]. We talk about two works that are done on this subject.

Zhu proposed a method in 2017 to rank the concepts of a MOOC class so that students can know before taking the course or during it, how important are the lessons thought in the class. A framework was proposed that automatically detects concepts of the lesson and ranks them. Their framework has three parts: first, by using LDA, concepts are extracted from subtitles. Then with a Pagerank algorithm, the importance of words is calculated, and in the last step, the transforming function applies the value of Importance to the words and importance of concepts are calculated. To evaluate and compare their method, they used Random algorithm, Bag of words, TF-IDF and TextRank as baselines. For future works, authors consider using learning patterns of students in the concept graph [8].

Asaad proposed a method in 2018 for constructing concept graphs to present the succession of concepts in a course in a MOOC system using lectures transcripts [9]. They proposed two measures to discover concepts ordering and constructing a directed graph for it: the bridge ensemble measure and the global direction measure. The bridge ensemble measure detects the overlap between concepts and the occurrence of two concepts in one time-window and also in which lecture a concept was talked about for the first time. The global direction measure

Table I. Comparison of the works studied on the subject of curriculum planning

Ref.	Year	Goal	Dataset	Metrics	Algorithms	Tools
[2]	2011	To extract the best course combination	82 courses categorized in 13 with 45 record and 15 features	- Support - Confidence	- Apriori	- WEKA - Moodle
[3]	2015	Effects of grouping students and optimizing schedules on the profit gained from class	- BUCS - BUCSSynth	- Profit gained	- K-means	- Python - Scipy
[4]	2016	Detecting the best educational path	Two datasets containing students' information	- Click - Follows	- USCAN	-
[5]	2018	Building recommending a model with student performance and course dependency	Information of student of the computer science department of Saarland University	- Mean Absolute Error (MAE) - Root Mean Square Error (RMSE)	- Collaborative Filtering - Mann-Whitney U-Test	-
[6]	2018	A multi-criteria method to recommend the best course combination	Results of more than 1700 student evaluation	- Root Mean Squared Error (RMSE) - Discount Cumulative Gain - Reach Time	- Genetic Algorithm - Collaborative Filtering - Content-Based Filtering - Exhaustive Search	- Apache Mahout - JCLEC Library

considers the time of concepts being discussed whether globally or locally in each lecture. The first measure has three components: the bridge, the sliding window, and the first lecture indicator. Bridge refers to the concepts that are preliminary to some other concepts and on should know them first to understand the next one. The sliding window is for the time which two concepts are related but always talked about together. If such a relation is not discovered within a lecture, it will be missed because it is not between two different lectures. The third component marks the first lecture that a concept is spoken about. The authors used TF to determine what concept is talked about in each lecture. The second measure is used to give directions to graph more decisively. If concept A is discussed before concept B, it is more likely that A is preliminary to B. to evaluate the study they used the course information retrieval and search engines of Coursera to plot the course graph, and they used AutoPhrase to extract phrases. As for future works, they plan to compare the performance of the constructed graph using different methods with the comments of students and teacher included.

3.3 Result prediction

One of the useful information for the university is to know the results that a student will get. Using this information, the university can invest in promising students and warn those who are threatened by the danger of failing a course and guide them for improving their situation. Some of the works done on this subject are briefly discussed below [10].

Salazar published a work in which they talked about extracting knowledge such as success, failure, retention, and desertion from student information using clustering and decision rules using data mining algorithms. Works done in this paper can be divided into steps discussed below: first is creating a data warehouse by preprocessing data and filtering the unwanted data out. Second is to detect useful correlations between variables. The third is to cluster the data and choose the clusters to discover hidden rules. Forth is to extract decision rules by applying the decision tree to chosen clusters. The fifth is to evaluate the extracted rules with the help of an expert, and the last step is to propose useful recommendations to improve the university academic system. They used IUS students' information from 1986 to 1999. They calculated two features, desertion and retention using time and added them to feature set. To analyze the results, numeric data were categorized and labeled. 331 rules were extracted, and then, they were evaluated by experts and graded 7.94 out of 10. Some recommendation was proposed based on the discovered rules. Some of which were educational consulting and analyzing educational situation during the semester. The authors proposed to use the system developed in this work to develop an expert system for future works [11].

In another work by Strecht in 2015, some of the most popular classification and clustering algorithms were evaluated and compared in predicting success or failure and grades of students. The former is considered a classification and the latter is a regression problem. They created a model for each course. They used Proto University data which consisted of around 700 courses. SVM and decision tree gave the best result for classification and SVM, Random Forest, AdaBoost.R2 gave the best results for regression. They used a model that only predicted failure as the baseline model for the classification problem. Baseline model for the regression problem predicted the average grade for the dataset. They evaluated the models with K-fold Cross-Validation. For regression RMSE was metric, and for Classification, they calculated The F1 Score. As for future work, they plan to predict the range of grades as a compliment for this work. Also, using an automated approach to optimize algorithms parameters is an idea for future works [12].

Mori and Chan published a work in 2016 in which they analyzed the behavior and performance of students using high-level features and the random forest algorithm. They

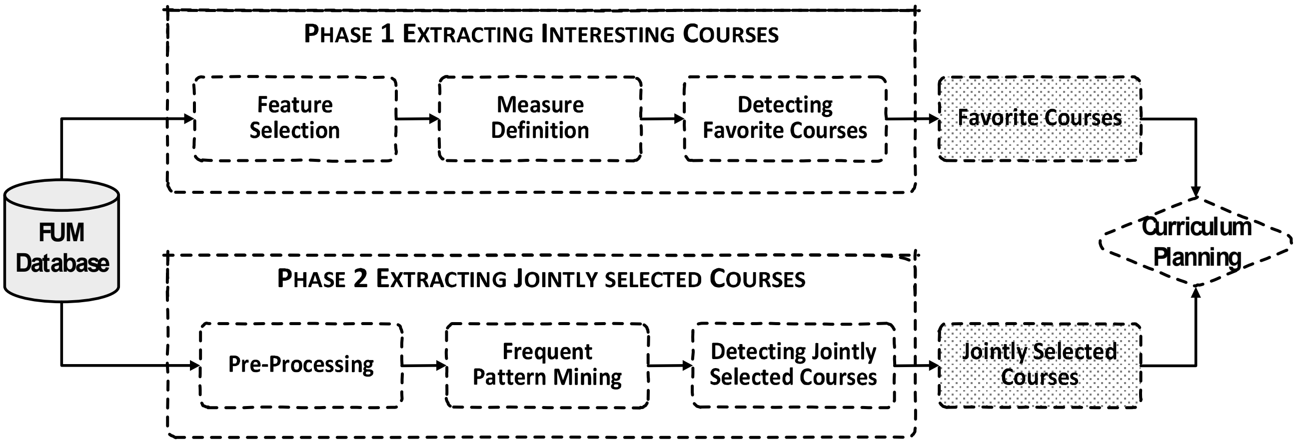

Fig. 1. The proposed method.

introduced features derived from the courses' syllabus and sequential patterns. They used the random forest with cross-validation. Considering a course including ten periods, the evaluation showed that the proposed model has the accuracy of more than 70% in the first period and more than 90% in the first five periods. Features form the log file considered as low-level features, and features that are during activities are high-level features. The first step is to create features representing student behavior. Second is to find relations between features using machine learning algorithms, and the third step is to construct the model. They used SPAM to detect high-level features considering behavioral sequences. In the random forest, the root node holds the most important features and other nodes are used to analyze the combinations of behaviors. They used the information of Concept and principle of behavior analysis course offered in 2013 as dataset which included the information of 110 students. They used the original decision tree and the decision tree algorithm without post pruning as the baseline method. According to the results, their approach showed improvements [13].

Yee-king proposed a collaborative system to annotated educational videos in 2016. They tested the system with one of the courses in the Coursera website, and KNN and student click log was used to predict and categorize student grades. The study was done on the information of 993 students who completed all assignments of the creative programming course. They used a collaborative annotating tool for videos and asked the student to develop the code they were given and create a five-minute video of themselves running the program and explaining their code, three times during the semester. They would load their video to the system and could be viewed and annotated by everyone. The system recorded the clicks and mouse-overs as log files, and vectors with 16 features were created to predict student grades. The result showed that the most reliable feature was playing which is logged when a user clicks on a comment on the timeline to open the discussion thread [14].

4. PROPOSED METHOD

n this section, we try to analyze the favorite courses and schedule them in such way that those favorite courses that are usually picked together, do not interfere with each other and then recommend them to the students making the process easier for them. To do so, we will first describe the used dataset in the present study. After presenting data preprocessing tasks, the proposed method will be described in details based on the steps presented in Fig. 1.

Table II. The features used in this study

#	Name	Type	Desc
1	Stno	Nominal	Student Number
2	FldCode	Nominal	Field Code
3	EduSecCode	Nominal	Education Code
4	CouCode	Nominal	Course Code
5	Grade	Numeric	Grade
6	CurrentStatus	Nominal	Current Status
7	MaxCapacity	Numeric	Max Capacity
8	PLesName	Nominal	Persian Course Name

4.1 FUM dataset description

In this work, the educational dataset of Ferdowsi University of Mashhad (FUM) is used for our experiments. FUM is currently one of the nation's top three universities and also the largest center of higher education in North-East of Iran. As a result, FUM dataset contains comprehensive information of students of various degrees (associate, bachelor, master, and Ph.D.) in 537 different fields of study [15]. This dataset has been collected by the Center of Information and Communication Technologies of FUM and anonymously published for research and educational purposes during 2001-2018.

4.2 Extracting interesting courses

As shown in Fig.1, the first phase of the proposed method contains three tasks. The first step is selecting the required features from FUM dataset. Then, six measures are defined to select interesting courses from students' points of view, and at the last step, the most favorite courses are extracted using proposed measures.

4.2.1 Feature selection

We have selected a portion of FUM dataset containing master and doctoral courses. We select Masters and Ph.D. students because they have more freedom of choice when it comes to picking courses without any limitation. The candidate features used in this work are extracted from three tables of Student information, Grades, and Courses. Table II illustrates these features.

After selecting the features mentioned above, we performed data cleaning and preprocessing tasks. The details of data preparation are summarized as below:

1- First, the required features are selected from Student information, Grades, and Courses tables. In this step, some features such as 'Stno', FldCode, and EduSecCode are selected from the Student information table. Also, 'CouCode', Grade, and 'CurrentStatus' are selected from Grade table. Finally, from the Courses table, CouCode and PLesName are selected. Afterward, the join operator is applied to obtain a unique table.

2- In this step, according to EduSecCode that represents students' grade, the master and doctoral students are classified into two groups. So, the results are classified as Professional Doctorate, Doctor of Philosophy, and Master.

3- Among all candidate features, only Grade and MaxCapacity features are numeric. By looking in details, there is no outlier data in these features.

4.2.2 Measure definition

In this step, to discover the type of graduate students' favorite courses, six measures are defined to select interest courses as follows:

A. **The number of students who picked the course:** In this step, the number of students who picked the course is computed. So, this measure is calculated having both fields of 'CouCode' and the number of students.

B. **The GPA of students in each course**: According to this measure, the average grades of students for each course is computed. It should be noted that the courses which are selected by 50 students or fewer are not taken under consideration (Even if the course has a higher AVG score).

C. **The variety of students who selected the course (based on the field of study):** In this way, a new table is generated to show the number of students with a different subject who registered in each course. It seems that the course can be considered as a favorite if at least, five students in the same subject select it. Hence, if, for example, if twelve students from three different field of study select a course, it can be a measure to show that this course is helpful or interesting for these three fields of study.

D. **The number of students who drop the courses during the semester:** The 'CurrentStatus' field is used to know the number of students who dropped the course during the semester. This way, we know the number of students who select and drop the course. It should be noted that we cannot decide the favorite course, only based on the number of students who drop a course. So, a new feature as HperC is computed. This feature is calculated as the ratio of the number of students who dropped the course and those who have registered in it.

E. **The number of times that a course is offered:** Another measure that can be effective to discover favorite courses is the number of times it is offered. For example, the course can be classified as offered regularly, randomly, or in even-odd semesters. To this end, the number of times a course is offered is computed.

F. **The average of master and doctoral students who register in a course:** This measure is calculated by dividing the number of all student by the number of offered courses.

4.2.3 Detecting Favorite Courses

Since our method was based on a heuristic methodology, there was no evaluation metric or previously achieved results to compare evaluate our work by comparing it with them. To overcome this challenge, we considered each measure independently, then sorted the result and extracted the top 10 to construct a table containing 10 top results of each feature analyzed independently. Then we searched for courses that two or more of measures had in common. In other words, we calculated the mean of all measures for top 10 results and used it as the label. In this way, we achieved a sense of effectiveness for each feature.

4.3 Extracting jointly selected courses

In the second phase of the proposed method, the frequent pattern mining methods are employed to discover frequent courses, and then the jointly selected courses are extracted.

4.3.1 Pre-processing

In this step, first the 'course code' selected from Table II for frequent pattern mining. In this way, the take courses process and course represent transaction set and items, respectively.

4.3.2 Frequent Pattern Mining

To discover frequent courses in masters and doctoral which are picked together, we need to extract association rules from this dataset, and to this end, we used Apriori and Eclat algorithms.

4.3.3 Detecting Jointly Selected Courses

We constructed transactions for each field showing which courses are picked by students and after applying Apriori and Eclat to list them by confidence and lift, our results were very similar. To this end, first, we extracted the two frequent Itemsets. Results showed that one of these two courses are in the thesis. Then we analyzed the three frequent itemsets. It's notable that increasing k more than 3 reduced the accuracy.

4.4 Curriculum planning

As mentioned above, the favorite courses and jointly selected courses are extracted. With the help of the result of these phases, the education department will be planning to present course in such way that the interesting courses can be easily chosen by the students in the future semester without any time conflict. In other words, considering discovered these courses in two phases, curriculum scheduling will be done in a way that they do not interfere with each other.

5. RESULTS

Masters and Ph.D. students have more freedom of choice when it comes to picking courses. Because of this, we cannot extract the preliminaries from extracted rules, but we can discover which courses are frequently picked together. According to discovered patterns presented in Table 3, the thesis courses pick count is equal to all other courses, and all counts are 117. A reason for this result is that the thesis can be picked with any other courses without any limitation. So, to select the most frequent 20itemset among all courses, we need to extract frequent 3-itemset and then remove the 'thesis' course. By looking at details, the courses A and D are chosen three times, so, these items are considered as

Table 3. An example of jointly selected courses

Course Names
{Thesis, Course A, Course B}
{Thesis, Course C, Course B}
{Thesis, Course D, Course A}
{Thesis, Course A, Course D}
{Thesis, Course F, Course A}
{Thesis, Course G, Course K}
{Thesis, Course A, Course D}
{Thesis, Course A, Course B}

frequent. Hence, an appropriate strategy for course scheduling is to prevent any time conflict for these courses.

6. CONCLUSION

In this paper, we proposed a heuristic method to discover students' favorite courses based on features extracted from educational data. We first analyzed the effects of six heuristically proposed measures to discover the favorite courses from students' point of view. Then we extract frequent jointly selected courses using Apriori and ECLAT algorithms. By comparing two sets of courses obtained from the first and second phases, curriculum scheduling can be done in a way that favorite and frequently selected courses do not have any time conflict.

To better analyzing the results, one of our future research direction is to evaluate the interestingness of courses based on surveys completed by students themselves. In this way, we can define the interestingness level as a label and select the most effective measures by correlation analysis method.

7. REFERENCES

[1] A. Dutt and M. A. Ismail and T. Herawan, "A Systematic Review on Educational Data Mining," IEEE Access, vol. 5, pp. 15991-16005, 2017.

[2] Aher, Sunita B and Lobo, LMRJ, "A Framework for Recommendation of courses in E-learning System," International Journal of Computer Applications, vol. 35, pp. 21--28, 2011.

[3] Bahargam, Sanaz, and Erdos, D'ora and Bestavros, Azer and Terzi, Evimaria, "Personalized Education; Solving a Group Formation and Scheduling Problem for Educational Content," International Educational Data Mining Society, p. 4, June 2015.

[4] Jugo, Igor, and Kovacic, Bozidar and Slavuj, Vanja, "Guiding Students Towards Frequent High-Utility Paths in an Ill-Defined Domain," Ninth International Conference on Educational Data Mining, pp. 599--600, 29 July 2016.

[5] Michael Backenköhler, Felix Scherzinger, Adish Singla, Verena Wolf, "Data-Driven Approach Towards a Personalized Curriculum," Eleventh International Conference on Educational Data Mining, p. 6, 2018.

[6] Esteban, Aurora and Zafra, Amelia and Romero, Crist, "A Hybrid Multi-Criteria Approach Using a Genetic Algorithm for Recommending Courses to University Students," International Educational Data Mining Society, 2018.

[7] Bakhshinategh, Behdad, Zaiane, Osmar R., ElAtia, Samira, Ipperciel, Donald, "Educational data mining applications and tasks: A survey of the last ten years," Education and Information Technologies, vol. 23, pp. 537--553, 01 Jan 2018.

[8] Zhu, Jie, and Li, Xiang and Wang, Zhuo and Zhang, Ming, "*An effective framework for automatically generating and ranking topics in Mooc videos*," Educational Data Mining, vol. 2017, pp. 150--155, 25 June 2017.

[9] ALSaad, Fareedah, and Boughoula, Assma and Geigle, Chase and Sundaram, Hari and Zhai, ChengXiang, "Mining MOOC Lecture Transcripts to Construct Concept Dependency Graphs," Eleventh International Conference on Educational Data Mining, p. 7, 2018.

[10] Bogarín, Alejandro and Cerezo, Rebeca and Romero, Cristóbal, "*A survey on educational process mining*," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, p. e1230, 2018.

[11] Salazar, A and Gosalbez, J and Bosch, I and Miralles, R and Vergara, L, "A case study of knowledge discovery on academic achievement, student desertion and student retention," Information Technology: Research and Education, vol. 17, pp. 150-154, 2004.

[12] Stretch, Pedro; Cruz, Luís; Soares, Carlos; Mendes-Moreira, João; Abreu, Rui, "A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance," International Educational Data Mining Society, p. 4, 2015.

[13] Mori, Makoto, and Chan, Philip, "Identifying Student Behaviors Early in the Term for Improving Online Course Performance," Ninth International Conference on Educational Data Mining, pp. 611--612, 2016.

[14] Yee-King, Matthew and Grimalt-Reynes, Andreu and d'Inverno, Mark, "Predicting student grades from online, collaborative social learning metrics using K-NN," Ninth International Conference on Educational Data Mining, pp. 654--655, 2016.

[15] Fatemi, Mohammad Rasool, Behnam Bakhshi, Alireza Zamani, and Behshid Behkamal. "*A scenario-based approach for the behavior analysis of talented students*." In 2017 7th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 384-389. IEEE, 2017.