An Overview on the Use of Educational Data Mining for Constructing Recommendation Systems to Mitigate Retention in Higher Education

Thiago Nazareth de Oliveira
Instituto de Computação
Universidade Federal Fluminense (UFF)
Niterói, Brazil
thiagooliveira@id.uff.br

Flavia Bernardini
Instituto de Computação
Universidade Federal Fluminense (UFF)
Niterói, Brazil
fcbernardini@ic.uff.br

José Viterbo
Instituto de Computação
Universidade Federal Fluminense (UFF)
Niterói, Brazil
viterbo@ic.uff.br

2021 IEEE Frontiers in Education Conference (FIE) | 978-1-6654-3851-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/FIE49875.2021.9637207

***Abstract*—In higher education, many students exceed the expected time to complete their undergraduate programs. This delay is called retention, which can lead to program abandonment. STEM undergraduate programs, in particular, have higher retention and dropout rates when compared to other non-STEM programs. The students in such programs end up exchanging or dropping out the programs before graduating, causing waste in economic, social and academic terms. In this context, Recommendation Systems can be used to support students and managers in choosing disciplines, contributing for them achieving better academic performance and thus aiming to improve student learning and engagement and mitigate retention. For constructing these Recommendation Systems, Educational Data Mining techniques, including machine learning algorithms, can be used to identify and predict retention situations and contribute to reducing their occurrence. The aim of this paper is to present a Systematic Literature Review (SLR) for identifying the use of Educational Data Mining methodologies, techniques and tools to implement Recommendation Systems with a focus on preventing student retention in higher education programs. We selected studies available in digital libraries that are international references in publications of scientific articles, in order to answer the following research question: What machine learning methods were used in Recommendation Systems in the context of Educational Data Mining? Among the various studies found to reduce student retention rates, most used methods to predict student grades. We observed that there are many papers proposing the use of machine learning methods for predicting failure in disciplines, either through regressors or classifiers. However, just a few studies have proposed Recommendation Systems to assist students in choosing subjects at the time of enrollment for the next term, which indicates a large area for the development of further future work in this field.**

***Index Terms*—Recommendation Systems, Educational Data Mining, Retention in Higher Education, Machine Learning**

I. Introduction

Student retention, also called graduation delay, related to the time exceeded for completion of a given course, is a problem that affects Higher Education Institutions (HEI) [1]. This problem occurs to a great extent in STEM degree programs, which have high retention and dropout rates when compared for other non-STEM programs [2]. Students in such programs end up switching or abandoning the program before graduating, generating economic, social and academic waste [3]. Thus, detecting in advance which studies are subject to retention and their causes is very important so that educational managers (course coordinators, unit directors, among others) can foresee this situation, being able to propose effective actions so that this does not occur.

Recently, researchers in Educational Data Mining (EDM) have developed methods for exploring data in education [4]. One of the applications is the detection of retention and its causes. In this context, tools to support decision making can be built using data mining techniques to mitigate retention and improve student learning and involvement, contributing to better academic performance.

On the other hand, Recommendation Systems (RSs) have been succesfully used in many domains and they are present in the user's daily life, such as shopping websites, apps for streaming of movies and music, recommendations from friends on social networks, among others [5]. In general, RSs employ the use of software tools and techniques to suggest items to a user in his(her) application context [6]. RSs have also been used in higher education to prevent retention, such as the proposal of Almutairi, Sidiropoulos and Karypis [7] for predicting students final grades in courses that have not yet been taken, in order to help them for selecting courses of the current semester, increasing their chances of success. In literature, we found many studies presenting the state of the art of the use of EDM in distance education platforms, such as [8], some of them focus on steps of EDM for analysing students performance, and only a few point out the possibility of using these results for implementing RSs [9] [10], but not deepening their discussion. So, we have not found any literature review study that shows the state of the art of using RSs for retention problems in higher education thorugh the use of EDM technologies.

The objective of this article is to present a Systematic Literature Review (SLR) on the methodologies, techniques and tools used in EDM, with a focus on machine learning algorithms, which have been used to build and implement RS's, with the intention of mitigating the retention of higher education

students, contributing to reduce its occurrence. To achieve our objective, we present in Section II the methodology for conducting a systematic literature review. In Section III we present a brief description of the studies that were found. In Section IV we present the answers to the research questions. Finally, in Section V we present the conclusions of this article.

II. Methodology for Conducting Systematic Literature Review

In this Systematic Review of the Literature (SRL), we use digital libraries that are international references in publications of scientific articles, in order to answer the following research question: In the context of SRs for mitigation of retention in higher education, what methods of learning machine were used? As a secondary research question, we also seek to understand what were the general objectives of the studies found in the context of EDM? We used a total of 5 libraries, the *ACM Digital Library*, *IEEE Digital Library*, *Science @ Direct*, *Scopus* and *Springer Link*. We collect complete documents on these bases using the following string search: *("Educational data mining" OR "Educational mining") AND ("Retention") AND ("Recommendation systems").*

We found a total of 163 works, with years of publication between 2012 and 2020. The *Scopus* returned 104 studies in the search; *Springer Link* returned 42; a *Science @ Direct*, 13; a *IEEE Digital Library*, 3; and *ACM Digital Library*, only 1. Inclusion criteria (address RS's and EDM; focus on student academic performance; and address retention) and exclusion (duplicates; which do not address RS's; which focus on selecting students for higher education, and focused on distance learning) to identify the most relevant citations for the purpose of the current review. Studies with a focus on distance learning were excluded because, in this context, e-learning platforms are used and, from them, many features are extracted that are not available in face-to-face teaching, such as: information from students regarding the delivery of in activities, participation in forums, online time on the tool, iteration with tutor, partial notes among others. Figure 1 shows our PRISMA Flow Diagram [11]. The flow diagram describes the phases of the SLR, mapping the number of articles identified, included and excluded.

After applying these criteria, 31 scientific papers remained. Of these, 23 are primary studies, that is, studies that propose a methodology or method as a scientific contribution; and 8 are secondary studies, that is, literature review studies. A summary of these studies are presented in the next section. Figure 2 shows a pie chart with the number of studies per type.

III. Brief Description of Selected studies

In the primary studies analyzed, Johri *et al* [12] present a simplified data analysis proposal to support decision making by managers in retention problems. The authors do their case study at a large public university to examine the retention rate in engineering and science colleges, with the aim of providing information to help administrators and teachers develop policies that can decrease retention rates.

On the other hand, many studies describe a scenario of grade prediction and prediction of dropout, which can lead to retention, using machine learning and optimization, without being associated with the implementation of recommendation mechanisms. Hence, these predictions can be used by students, teachers and managers to support decision making in educational settings. In this line, there are studies that present proposals based on (i) supervised machine learning (13 studies); (ii) supervised machine learning for regression problems (1 study); and (iii) unsupervised machine learning (2 studies) — we describe these studies below. And, at the end of this section, we present the 8 secondary articles that were found in our literature review.

A. Supervised Machine Learning

In this context, Buenaño, Gil and Luján [13] aim to use machine learning techniques to predict student grades using grade history in subjects already taken. They used the decision tree induction algorithm to estimate whether a student will succeed or not, justifying the use of decision trees because they are interpretable by human beings. Polyzou and Karypis [14] aim to predict a student's performance at the end of the semester before the end of the course, focusing on underperforming students. The authors classify students into two groups: those who are likely to successfully complete a discipline or activity; and those who seem to have difficulties. After identifying the last group, action strategies can be devised to increase their likelihood of success. Polyzou and Karypis [15] extended the previous study. The grade prediction problem was also formulated as a binary classification task — two groups of students are formed according to the performance in the discipline. However, the authors extended the experiments using several learning algorithms: decision tree induction, construction of linear support vector machines, *Random Forest* and *Gradient Boosting*, having this last obtained the best performance. Alper and Cataltep [16] aim to predict the success of students (pass or fail), using several supervised learning algorithms, in different subjects in the Computer Engineering course. Student grade history data and additional indicators set by course coordinators are used. To accomplish this task, the authors compared several supervised learning algorithms, namely: *Bayesian Logistic Regression*, *Minimun Redundancy Maximum Relevance feature selection algorithm*, *Naive Bayes*, *Multilayer Perceptron*, *SVM* and *Logistic Regression*. Al-Saleem *et al* [17] use EDM techniques to extract knowledge from previous students to predict the future performance of students in disciplines. The authors use decision tree algorithms and present a case study by developing a system for predicting student grades in real time. In the system, students can view grades scheduled for one or more subjects. The goal is to help students choose subjects that they are most likely to succeed. Bañeres and Serra [18] present a system to support teachers and students, predicting, in real time, the chances of passing a discipline based on a different set of indicators, such as approved subjects or acquired skills. The objective of the system is to improve the quality of the

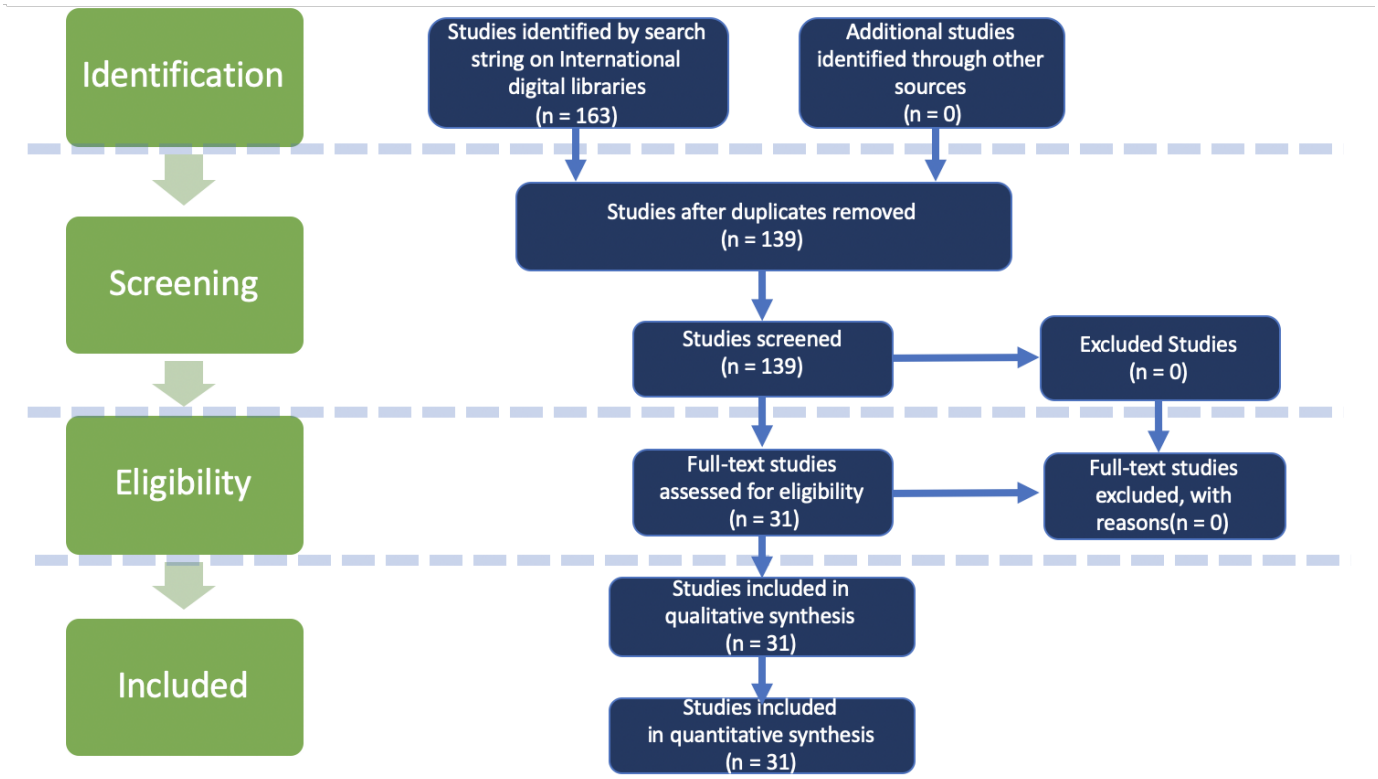

Fig. 1. Information flow through the steps of the review using PRISMA Flow Diagram (Adapted from [11]).

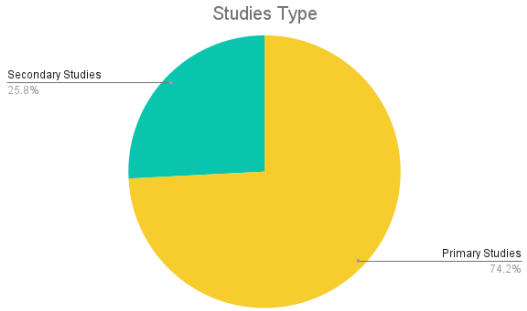

Fig. 2. Number of studies per type

models and information of the Virtual Learning Environment (VLE), which contains the attempts, achievements and actions performed by the students. The system consists of a decision tree model, trained using different data sources. Miguéis *et al* [19] classify students into segments, based on the average grade and the time required to complete the degree. The authors use a discretization algorithm to categorize student performance into five classes, corresponding to different levels of academic performance. Then, they build a model to predict student performance based on the categorization performed. One paper only modeled this problem as regression: In the context of using supervised machine learning for regression problems, Hu *et al* [20] build a specific regression model for predicting a student's grade in the discipline. The authors used data on students, subjects and instructors in order to assist in making decisions about future choices of subjects to be taken. Hence, they proposed a hybrid model to predict student performance in the next period, taking into account not only course scores (with the claim that grades alone may not fully capture all factors that affect student performance) but also other information related to students. Use random forest to find the importance of each attribute used in the model, to understand which are the factors that stand out the most for academic success. Two studies use machine learning to select attributes that lead to retention. Ghanem and Alobaidy [21] discuss the problem of redundant attributes for use in predicting student performance. Hence, they propose the use of decision trees to select the main attributes that affect student performance. Finally, they use a classifier built with the *Naive Bayes* algorithm to predict student performance, using only the attributes identified by the resulting decision tree, excluding irrelevant attributes. Marwaha and Singla [22] provide modeling to identify the most influential attributes that can predict student performance. The authors predict the pass or fail, as well as the student's performance in the first case (excellent, good or average). The authors compare several predictive machine learning algorithms using different attributes, belonging to the academic, demographic, social and behavioral categories, which can influence student performance.

B. Supervised Machine Learning for more complex classifications

Sharabiani *et al* [23] aim to predict the grades of engineering students in three course subjects that have a high failure rate, often leading to students dropping out: General Physics II, Calculus II and C / C ++ Programming. The authors are based on the fact that, normally, students who do not do well in these three subjects change their course and leave the Faculty of Engineering. The authors used Bayesian network models, which take into account the statistical interdependence between variables, to model the problem of predicting student grades in these disciplines. The authors compared the results with decision tree models, *Naive Bayes* and *K closest neighbors*. The accuracy of the Bayesian network model compared to the other models is significantly better. Hu and Rangwala [24] apply Markovian models to predict student performance in the following year. The authors present a case study, applying such models to identify students at risk using historical data from the analyzed students' grades Cano and Leonard [25] present an alert system based on the results of a classifier built by Interpretable Genetic Programming, building comprehensible classification rules. Several views built using these rules are offered. The goal is to identify students at risk of failing or dropping out of a course.

C. Unsupervised Machine Learning

Iam-On and Boongoen [26] present a clustering approach to explore categories and characteristics of students, aiming to increase the quality of analytical results by developing a descriptive model that can complement the predictive side. By obtaining clusters of students' academic profiles, the authors report that it is possible to discover the main characteristics and relationships between variables of interest. To generate a set of clusters, the clustering R-KM algorithm is used. In the German project *LAPS*, presented by Hinkelmann and Jordine [27], the *Apriori* algorithm was used to build sets of association rules. This project was developed and used to support universities in the challenge of identifying students who can benefit from the prior support of the university's student support and learning center, in order to improve academic results. Membership rules are used to analyze students' study progressions. These results are compared with the grades achieved by students in their program of studies so far. The authors state that since the progression of an enrolled student will not be statistically different from students who have completed or failed their program of study, the comparison can be used to make an individual analysis of the risk of failure or the possibility of student success .

Finally, 7 papers propose and analyze the behavior of RS's, of which 3 focus on specific proposals based on the development of recommendation mechanisms. Almutairi, Sidiropoulos and Karypis [7] aim to predict students' final grades in subjects that have not yet been taken to help students in selecting subjects in the semesters. The authors propose an RS's using collaborative filtering and matrix factoring. Huang *et al* [28] present an RS of optional subjects, based on score prediction, proposing a new collaborative filtering algorithm. The authors argue that the selection of optional subjects is a critical task for university students, due to the large number of subjects available that may be unfamiliar. Improper choices can lead to student withdrawal. Goga, Kuyoro and Goga [29] present an RS that can predict the academic performance of students in the first year, in addition to recommending the necessary actions to be taken to help them, thus guiding the institution's management in making decisions about intervention strategies when there is need. To build this system, the *Random Forest* algorithm was used because it presented better accuracy results compared to other tested algorithms. A form was also applied to experts to classify some factors, found in the literature, that can influence student performance, based on their perspective on whether the factor is very influential, influential, less influential or has no influence on performance academic. The factors identified in the literature are sex, average family income, mother's qualification, father's educational qualification, parents' marital status, mother's occupation, father's occupation, family size, ethnicity, religion, education sponsor, age at which student entered university, grade of high school, grade of the exam of University enrollment and accumulated average of the first year. Respondents considered all factors identified to be very influential. Since the variables were established from the literature and the interview with experts in the field, Machine Learning algorithms for the construction of the recommendation system were applied using student data. Using the random tree algorithm, an intelligent recommendation system was built to predict student performance in the first year. As presented, the system uses the following two categories of information: 1) family factors and 2) previous educational performance before entering a higher education institution. Vaidhehi and Suchithra [30] propose a framework for the construction of an intelligent recommendation system for choosing subjects. The recommendation mechanism maps the student, his learning style, the subjects and the context of the learning environment, and suggests the appropriate subjects for the student in a personalized way. The authors argue that the guidance systems of choice of subjects minimize the retention rates of the educational institution and improve student performance. Ren, Ning and Rangwala [31] propose Additive Latent Effect models in the Factor Matrix structure to predict the grade that a student must obtain in a discipline not yet taken. The latent effects of the academic level of the discipline instructor and the student are inserted, together with the global latent factor of the student for the task of predicting the grade more accurately. Ren, Ning and Rangwala [32] aim to predict the grade that a student must obtain in a discipline that is available for enrollment in the next period. The authors consider that the student's grade in a given discipline is determined by two factors. The first factor is the student's competence in relation to the topics, content and requirements of the discipline. The second factor is the student's previous performance in relation to other subjects. In this study, matrix factorization algorithms were used for the note prediction task. Sweeney, Lester and Rangwala [33] also

presents an approach to predicting student discipline scores for the next enrollment period in a traditional university setting. The authors compared several matrix factoring techniques. The authors restricted the experiments to the configuration of collaborative filtering, using only the students' grades for the predictions.

D. Secondary Studies

We also collected with our search string eight papers presenting literature reviews. Cui *et al* [10] present two main categories of studies that focused on predicting student performance in the context of higher education. Of the 121 reviewed articles published between 2002 and 2018, most studies (86) focused on predicting performance at the level of disciplines in undergraduate or graduate courses (first category). The second category (a total of 35 studies) grouped studies to predict student results at the program level of undergraduate or graduate courses, that is, the student's overall academic performance. Hellas *et al* [34] present a review on predicting student performance. They present different performance concepts and summarize what these studies described about the factors and methods used to predict performance. First, the authors aim to identify categories and trends already identified in existing literature reviews to contribute to a new understanding of the literature in the area. 13 reviews and research papers were selected, selected from a total of 147 published between 2010 and 2018. Of these articles analyzed, five summarize the factors for predicting performance, two summarize the methods for predicting performance and four offered some explanation of the meaning performance or what was being predicted. None of the researches contributed to the intersection of these three problems. Then, the authors carried out a systematic review of the literature, which totaled 357 studies analyzed to answer the proposed research questions, which were: what is the current state of the art in predicting student performance and what is the quality of these studies.

Manjarres *et al* [35] present a review of the studies, published between 2008 and 2015, focused on the use of EDM techniques. The authors classify the 127 studies found in three ways: the first by themes related to the educational domain (for example: dropout and retention, performance analysis and evolution of students, generation of educational recommendations); the second for data mining techniques that were used; and the third relate the data mining techniques used in each educational domain theme presented. Thus, the authors seek to provide elements to support decision making about which of these data mining techniques can be used in particular situations. Martins, Migueis and Fonseca [36] present a brief description of the literature regarding the most cited scientific studies in the field of EDM, aiming to value and affirm this area as a tool for analyzing and building strategies in the decision-making processes of academic institutions. The authors present a summary of the systematic reviews found in the literature and what are the most used techniques in the implementation of EDM tasks. The authors do not explain the period of analysis of the studies. Dutt, Ismail and Herawan [37] present a systematic literature review on the clustering algorithm and its applicability and usability in the context of EDM. The authors found 166 studies between 1983 and 2016, among which 35 applied predominantly the clustering approach for EDM. Bin Mat *et al* [9] present a literature review in order to give a brief overview of how academic analysis has been used in educational institutions, what tools are available and how the institution can predict student performance. They provide a summary of the tools used and what input data is needed for those tools. Bakhshinategh *et al* [38] present a review of the existing literature on EDM and the tasks introduced in each one. The authors study the examples found in the research, as well as the publications of recent years in the *Journal of Educational Data Mining*. The applications of the EDM techniques used were grouped into categories and subcategories for similar purposes. 13 categories of applications of these techniques were identified. Four applications found in the review are grouped under "Student modeling", six under "Decision Support Systems" and three are presented as "Others" because they differ from other applications. The studies found were published between 2009 and 2016. Peña-Ayala [39] presents a review of the EDM literature, where the authors analyze 240 papers, divided into two categories: 222 papers describe used approaches and 18 describe used tools. The analyzed papers were published between 2010 and 2013. The authors grouped the papers that describe approaches used in EDM in: student modeling (43), student behavior modeling (48), student performance modeling (46), evaluation, support and feedback for students (21) and finally curriculum, domain knowledge, sequencing and support from teachers (19). The studies describing the tools used in the EDM area were categorized into: extraction, learning support and attribute engineering (classical steps of data mining) (4), visualization (6) and support analysis (8). For the authors, support analysis aims to evaluate the behavior and performance of the students during their interaction with the computer-based educational systems.

IV. Answering Research Questions

Our main purpose in this study was to answer our research question: "In the context of SRs for mitigation of retention in higher education, what methods of learning machine were used?". For answer this question, we observed that one study used *collaborative filtering algorithms and matrix factoring*, two used only *collaborative filtering* and three used the *Random Forest* machine learning algorithm. With that, we could observe that only 4 of the 31 study selected in the SLR was really proposing a new RS. The 27 study that did not implement RS's were found by the string search because they had the term *Recommendation Systems* in their text, but citing the possibility of using these systems or mentioned this approach in the descriptions of related studies.

To answer the secondary question, which aims to understand the general objectives of the study found in the context of EDM, the Table I presents the main objectives we found in the analyzed primary studies. We also list, for each ob-

TABLE I
LIST OF OBJECTIVES AND STUDIES FOUND

Objective	Studies	Total
Performance Prediction	[13], [15], [25], [19], [14], [24], [20], [23], [16], [17], [18], [27]	12
Students Categorization	[26]	1
Attribute selection	[21], [22]	2
Data analysis	[12]	1
New RS	[7], [28], [29], [30],	4
Performance Forecast for RS	[31], [32], [33]	3
Secondary Studies	[10], [34], [35], [36], [37], [9], [38], [39]	8

jective, the studies found. The main objects were students performance prediction (Performance prediction); Students Categorization; Attribute Selection; Data Analysis; proposal for new RSs (New RS); proposal of performance forecasting to build RS (Performance Forecast for RS); and Secondary Studies. Among the studies presented in this table, three of them have proposals for the use of RS's applied in the context of higher education, in order to help students in the task of choosing subjects, mandatory or optional. For this, these studies used the forecast of grades in the subjects not yet done to make recommendations, with the objective of improving the academic performance of students. One paper proposes a framework for building a RS of disciplines based on the mapping of student information, their learning style, the disciplines and the context of the learning environment. Thus, we conclude that all the studies analyzed in our SLR that propose RSs have a unique objective: recommending disciplines for students. However, RSs could also be proposed for recommending (i) to undergraduate program coordinators which students may failure in their courses subscriptions; or (ii) to professors for helping them to identify and prevent which students may failure in their courses.

Regarding to secondary studies, most of them aim to understand what tools, techniques and what problems have been attacked in the context of educational data (such as performance prediction for retention and evasion treatment). Our study focused on finding studies related to proposing RSs for handling student retention. Table **??** presents the scope we identified of the analyzed secondary studies. Section III-D describe each of these secondary studies in more details. Although analysis of student performance prediction was not in our search string, this scope emerged in our study, as we seen in Section III. We could observe in our table that none of the secondary studies focused on our main three themes (scope): EDM, Retention and RS.

V. CONCLUSIONS AND FUTURE WORK

In this study, several studies related to EDM, retention and RS were presented. We focus on the context of higher education. Our objective was to understand and present the state of the art of the SR proposal for the prevention of student retention in HEIs. The papers were selected from 5 databases of internationally known journals. After applying inclusion and exclusion criteria, we selected 31 of the 163 studies that resulted from the applied search. We present an overview of all selected studies. The vast majority of them have the objective of predicting student grades, in subjects not yet taken or in complete courses. The studies found used several different features to model the problems they proposed to solve, such as: academic (year of enrollment, grades, courses, delivery of activities), demographic (ethnicity, age, nationality) and social (family income, profession and parents' education level). To evaluate the generated models, several evaluation metrics were used, such as, for example, confusion matrix, accuracy, precision and ROC curve. Few studies, among those selected, propose RS's to recommend subjects for students to enroll in the next school term. This indicates a large field still open for the development of further future work in this area.

We observed that the use of machine learning, in the context of EDM, can be used to build RS's. In addition, in a previous study by our research group, [40] proposed methods for mining *Directed Acyclic Graphs (DAGs)* to discover retention patterns in undergraduate programs. The authors argue that the representation of graphs as DAGs allows the use of efficient algorithms to find paths with associated costs. The proposed methods aim to find the longest path in degree programs and the most expensive path in student grade report, helping to indicate the relationship between completing the minimum time in a degree program and the most expensive path in grade report of students. This is also an approach that can be used to build RS's for identification and prevention of retention.

REFERENCES

[1] A. d. V. C. Campello and L. N. Lins, "Metodologia de análise e tratamento da evasão e retenção em cursos de graduação de instituições federais de ensino superior," *XXVIII Encontro Nacional de Engenharia De Produção, RJ, 13p*, 2008.

[2] A. Sithole, E. T. Chiyaka, P. McCarthy, D. M. Mupinga, B. K. Bucklein, and J. Kibirige, "Student attraction, persistence and retention in stem programs: Successes and continuing challenges." *Higher Education Studies*, vol. 7, no. 1, pp. 46–59, 2017.

[3] R. L. L. Silva Filho, P. R. Motejunas, O. Hipólito, and M. B. d. C. M. Lobo, "A evasão no ensino superior brasileiro," *Cadernos de pesquisa*, vol. 37, no. 132, pp. 641–659, 2007.

[4] R. Baker, S. Isotani, and A. Carvalho, "Mineraçao de dados educacionais: Oportunidades para o brasil," *Revista Brasileira de Informática na Educação*, vol. 19, no. 02, p. 03, 2011.

[5] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," in *Recommender systems handbook*. Springer, 2011, pp. 1–35.

[6] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, "Recommender systems," *Physics reports*, vol. 519, no. 1, pp. 1–49, 2012.

TABLE II
LIST OF OBJECTIVES AND STUDIES FROM SECONDARY STUDIES

studies	Scope				
	EDM overview	Analysis of Student Performance Prediction	Focus on Retention in EDM	Focus on Clustering	Focus on RSs
[10]	✗	✓	✗	✗	✗
[34]	✗	✓	✓	✗	✗
[9]	✗	✓	✓	✗	✗
[35]	✓	✗	✗	✗	✗
[36]	✓	✗	✗	✗	✗
[38]	✓	✓	✗	✗	✗
[39]	✓	✓	✗	✗	✗
[37]	✓	✓	✗	✓	✗

[7] F. M. Almutairi, N. D. Sidiropoulos, and G. Karypis, "Context-aware recommendation-based learning analytics using tensor and coupled matrix factorization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 5, pp. 729–741, 2017.

[8] P. Moreno-Marcos, C. Alario-Hoyos, P. J. Muñoz-Merino, and C. Kloos, "Prediction in MOOCs: A review and future research directions," *IEEE Trans Learning Tech*, vol. 12, no. 3, pp. 384–401, 2018.

[9] U. Bin Mat, N. Buniyamin, P. M. Arsad, and R. Kassim, "An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention," in *Proc 2013 IEEE 5th Conf Engineering Education (ICEED)*. IEEE, 2013, pp. 126–130.

[10] Y. Cui, F. Chen, A. Shiri, and Y. Fan, "Predictive analytic models of student success in higher education," *Information and Learning Sciences*, 2019.

[11] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, P. Group *et al.*, "Preferred reporting items for systematic reviews and meta-analyses: the prisma statement," *PLoS medicine*, vol. 6, no. 7, p. e1000097, 2009.

[12] A. Johri, H. Rangwala, J. Lester, and O. Almatrafi, "Board# 65: Retention and persistence among stem students: A comparison of direct admit and transfer students across engineering and science," in *2017 ASEE Annual Conference & Exposition*, 2017.

[13] D. Buenaño-Fernández, D. Gil, and S. Luján-Mora, "Application of machine learning in predicting performance for computer engineering students: A case study," *Sustainability*, vol. 11, no. 10, p. 2833, 2019.

[14] A. Polyzou and G. Karypis, "Feature extraction for classifying students based on their academic performance," *International Educational Data Mining Society*, 2018.

[15] ——, "Feature extraction for next-term prediction of poor student performance," *IEEE Transactions on Learning Technologies*, vol. 12, no. 2, pp. 237–248, 2019.

[16] M. E. Alper and Z. Cataltepe, "Improving course success prediction using abet course outcomes and grades." in *CSEDU (2)*, 2012, pp. 222–229.

[17] M. Al-Saleem, N. Al-Kathiry, S. Al-Osimi, and G. Badr, "Mining educational data to predict students' academic performance," in *Proc Int Workshop ML and DM in Pat Recog*, 2015, pp. 403–414.

[18] D. Bañeres and M. Serra, "On the design of a system to predict student's success," in *Conference on Complex, Intelligent, and Software Intensive Systems*. Springer, 2017, pp. 274–286.

[19] V. L. Miguéis, A. Freitas, P. J. Garcia, and A. Silva, "Early segmentation of students according to their academic performance: A predictive modelling approach," *Decision Support Systems*, vol. 115, pp. 36–51, 2018.

[20] Q. Hu, A. Polyzou, G. Karypis, and H. Rangwala, "Enriching course-specific regression models with content features for grade prediction.(2017)," 2017.

[21] A. S. Ghanem and H. Alobaidy, "Data mining for intelligent academic advising from noisy dataset," in *Proc 2018 Int Conf Innov Intel Inf, Comp and Tech (3ICT)*. IEEE, 2018, pp. 1–5.

[22] A. Marwaha and A. Singla, "A study of factors to predict at-risk students based on machine learning techniques," in *Proc. Int. Com. Cont. and Dev*. Springer, 2020, pp. 133–141.

[23] A. Sharabiani, F. Karim, A. Sharabiani, M. Atanasov, and H. Darabi, "An enhanced bayesian network model for prediction of students' academic performance in eng. prog." in *Proc. 2014 IEEE G. Eng. Educ. Conf. (EDUCON)*, 2014, pp. 832–837.

[24] Q. Hu and H. Rangwala, "Course-specific markovian models for grade prediction," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2018, pp. 29–41.

[25] A. Cano and J. D. Leonard, "Interpretable multiview early warning system adapted to underrepresented student populations," *IEEE Trans Learning Tech*, vol. 12, no. 2, pp. 198–211, 2019.

[26] N. Iam-On and T. Boongoen, "Generating descriptive model for student dropout: a review of clustering approach," *Human-centric Computing and Information Sciences*, vol. 7, no. 1, p. 1, 2017.

[27] M. Hinkelmann and T. Jordine, "The laps project: Using machine learning techniques for early student support," in *Utilizing Learning Analytics to Support Study Success*. Springer, 2019, pp. 105–117.

[28] L. Huang, C.-D. Wang, H.-Y. Chao, J.-H. Lai, and S. Y. Philip, "A score prediction approach for optional course recommendation via cross-user-domain collaborative filtering," *IEEE Access*, vol. 7, pp. 19 550–19 563, 2019.

[29] M. Goga, S. Kuyoro, and N. Goga, "A recommender for improving the student academic performance," *Procedia-Social and Behavioral Sciences*, vol. 180, pp. 1481–1488, 2015.

[30] V. Vaidhehi and R. Suchithra, "An enhanced framework to design intelligent course advisory systems using learning analytics," in *Proc. Int. Conf. Data Eng. and Commun. Tech.* Springer, 2017, pp. 723–732.

[31] Z. Ren, X. Ning, and H. Rangwala, "Ale: Additive latent effect models for grade prediction," in *Proc. 2018 SIAM Int. Conf. Data Mining*, 2018, pp. 477–485.

[32] ——, "Grade prediction with temporal course-wise influence," *arXiv preprint arXiv:1709.05433*, 2017.

[33] M. Sweeney, J. Lester, and H. Rangwala, "Next-term student grade prediction," in *2015 IEEE International Conference on Big Data (Big Data)*. IEEE, 2015, pp. 970–975.

[34] A. Hellas, P. Ihantola, A. Petersen, V. V. Ajanovski, M. Gutica, T. Hynninen, A. Knutas, J. Leinonen, C. Messom, and S. N. Liao, "Predicting academic performance: a systematic literature review," in *Proc 23rd Annual ACM Conf Innov and Tec in Comp Science Edu*, 2018, pp. 175–199.

[35] A. V. Manjarres, L. G. M. Sandoval, and M. S. Suárez, "Data mining techniques applied in educational environments: Literature review," *Digital Education Review*, no. 33, pp. 235–266, 2018.

[36] M. P. Martins, V. L. Migueis, and D. Fonseca, "Educational data mining: A literature review," in *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2018, pp. 1–6.

[37] A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," *IEEE Access*, vol. 5, pp. 15 991–16 005, 2017.

[38] B. Bakhshinategh, O. R. Zaiane, S. ElAtia, and D. Ipperciel, "Educational data mining applications and tasks: A survey of the last 10 years," *Education and Information Technologies*, vol. 23, no. 1, pp. 537–553, 2018.

[39] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert systems with applications*, vol. 41, no. 4, pp. 1432–1462, 2014.

[40] J. d. J. Costa, F. Bernardini, D. Artigas, and J. Viterbo, "Mining direct acyclic graphs to find frequent substructures—an experimental analysis on educational data," *Information Sciences*, vol. 482, pp. 266–278, 2019.