Southeast Europe Journal of Soft Computing ISSN 2233 – 1859
Available online: http://scjournal.ius.edu.ba

Vol 10, No. 1, March 2021, pp. 55-70

Supervised Learning Algorithms in Educational Data Mining:

A Systematic Review

[1*] Alaa Khalaf Hamoud, [2] Jasim Mohammed Dahr, [3] Ihab Ahmed Najim, [4] Mohammed B. M. Kamel,
[5] Ali Salah Hashim, [6] Wid Aqeel Awadh, and [7] Aqeel Majeed Humadi

[1,5,6] College of Computer Science and Information Technology,
University of Basrah,
Basrah, Iraq.
[2] Directorate of Education in Basrah,
Basrah, Iraq.
[3] College of Computer and Mathematic Sciences,
Tikrit University,
Iraq.
[4] Institute for Data Science, Cloud Computing and IT Security,
Hochschule Furtwangen University,
Germany.
[7] College of software engineering,
Islamic Azad University of Khorasgan,
Iran.

**Corresponding Author*: alaa.hamoud@uobasrah.edu.iq

Article Info

Article history:

Article received on 20 March 2021

Received in revised form 26 March 2021

Keywords:

Educational Data Mining; Supervised Learning Algorithms; Supervised Methods; Systematic Literature Review.

ABSTRACT: Academic institutions always looking for tools that improve their performance and enhance individuals' outcomes. Due to the huge ability of data mining to explore hidden patterns and trends in the data, many researchers paid attention to Educational Data Mining (EDM) in the last decade. This field explores different types of data using different algorithms to extract knowledge that supports decision-making and academic sector development. The researchers in the field of EDM have proposed and adopted different algorithms in various directions. In this review, we have explored the published papers between 2010-2020 in the libraries (IEEE, ACM, Science Direct, and Springer) in the field of EDM are to answer review questions. We aimed to find the most used algorithm by researchers in the field of supervised machine learning in the period of 2010-2020. Additionally, we explored the most direction in the EDM and the interest of the researchers. During our research and analysis, many limitations have been examined and in addition to answering the review questions, some future works have been presented.

1. INTRODUCTION

Due to growth in the produced electronic data in the universities and the huge improvements in the algorithms that discover and extract patterns and information from the large electronic data, it has been a necessity to adopt and apply those algorithms to obtain meaningful information from the produced data in the universities. Educational Data Mining (EDM) can be defined as one of the academic fields that focuses on exploring new and useful information in data that is stored in educational settings. The explored information is used to support and develop cognitive theories of teaching and learning [1]. In fact, it can be defined as the application of techniques or methods of data mining to educational data that come from schools universities, or any educational environment to address important educational questions [2]. EDM focuses specifically on methods that explore the distinct or unique types of data obtained from educational data repositories to use these methods to better understand students and the settings in which they learn [3]. The process of EDM converts preliminary data coming from educational environments to valuable information that could be very effective in practice and educational research [4]. However, the data mining applications in the educational field are limited [5].

EDM exploits different data mining algorithms, statistical analysis, and machine learning algorithms over different types of data sets with a different number of dimensions in the educational sector. In this field, the data is taken from different systems and turned into information that has a great impact on educational practice and research. The main goal is to resolve issues related to education by analyzing these data. This process does not differ from other data mining approaches in following and applying key steps of implementing the data mining model [4]. Stakeholders in the field of EDM may be affected by or affect the EDM outcome. Stakeholders include all the persons involved in the academic sector from students, pupils, learners to teachers, educators, and instructors. Educational researchers, course developers, academic administrators are also involved in this field besides academic companies, and organizations [6]. There are many objectives and tasks in EDM [7-9], among those are increasing the organization's profitability, reducing learners' failure, improving academic learning progress, finding the factors that affect learners' progress, improving understanding of learners' behavior, understanding the learners' methods of learning, improving both students and domain model, implementing assessment of learning system performance, implementing recommendation systems for course adaptation, predicting students' performance student/teacher skills modeling and discovering/predicting learners' dropouts. Many other objectives can be involved as well, such as grouping similar students based on their characteristics similarity, enhancing learning process time, resources, curriculum, and schedule. EDM systems differ from each other in the ways they present/receive the knowledge and give access to the learners. The system types may be offline, online, and intelligent systems. Data mining algorithms have been examined and proved their accuracy in the field of EDM. Time-series forecasting, clustering, regression, classification, association rules mining, and machine learning have been quietly used and provide good predictions based on educational data. The machine learning field where the algorithm learns from the data, also widely used in this field. Many other approaches such as data warehouse [10-13], OLAP [14], fuzzy inference system over the internet of things [15-18], and cloud computing [19] can be adapted in the field of EDM. The integration of DM algorithms with machine learning and other fields may produce and powerful platform that presents fast accurate answers to stakeholders' questions [9].

There are several basic categories of methods or procedures that are suggested in EDM: clustering, prediction, discovery within models, relationship mining, and distillation of data for human judgment [20]. The primary goal of a prediction procedure is to name the data object and predicting a class. Predicting the educational outcomes for students is one of the main application areas of prediction [21]. There are several types of research in the area of prediction that have been conducted on different levels: at the level of the degree, at the level of course, or the level of the teaching system. In this systematic literature review, the authors focus on the studies that have been used supervised learning algorithms in EDM in the category of prediction.

Recently, supervised learning has received a lot of attention for its ability to accurately predict a variety of complex areas. One of the most important areas for accurately predicting outcomes is educational data. Supervised learning is also called supervised machine learning and is part of artificial intelligence and machine learning. It is defined by using a classified dataset to train approaches that accurately predict results or classify data.

The rest of the research is organized in the following order: the second section presents the previous studies in our field of research while the details of the methodology followed in the paper. The fourth section states our findings from the literature review. The fifth section discusses the results in the previous step. The final section provides a summary of the research, conclusion, and future work.

2. LITERATURE REVIEW

Data mining has been applied for several different purposes, including the prospecting of educational environments in the field of prediction. For instance, Fernandes, Holanda, Victorino, Borges, Carvalho, and Van Erven in [22] and during the 2015 and 2016 school terms conducted a predictive analysis of the academic performance of students in public schools of the Federal District of Brazil. The proposed technique is based on the conventional Cross-Industry Data Mining Standard Process (CRISP-DM) and uses a dataset collected from the Federal District of Brazil's State Department of Education repository. Initially, we performed a descriptive statistical analysis to gain insight into the data. Subsequently, two datasets were collected. The first dataset includes variables that were collected before the beginning of the school year, and the second includes academic variables that were collected two months after the start of the semester. Gradient Boosting Machine (GBM) classification models were developed to predict the academic results of student success for each dataset at the end of the school year. Results revealed that while the attributes' grades 'and absences were the most significant to predict student performance academic results at the end of the year, the analysis of the demographic attributes shows that neighborhood, school, and age are also possible measures of academic success or failure of a student.

Asif, Merceron, Ali, and Haider in [21], listed that they used data mining techniques to research graduate students' results, concentrating on two facets of their performance. First, predict the academic achievements of students at the end of a four-year program of study. Second, study the standard progress and integrate them with the outcomes of the forecasts. Two significant groups of students have been identified: low performance and high performance. Results suggest that by concentrating on a limited number of courses that measure especially good or bad success, prompt alerts and encouragement for low-profile students and guidance and opportunities for high-performing students may be offered.

In [23], a quantitative analysis on the usefulness of EDM methods is provided to show the effectiveness of the early prediction of the students that are likely to fail in the introductory programming courses. This research was differentiated as follows: first, examining the efficacy of such techniques in detecting students who are likely to fail early enough to take steps to minimize the rate of failure; second, evaluating the effects of pre-processing data and fine-tuning algorithms on the effectiveness of these techniques. In their research that has been done on two distinct and independent databases on introductory programming courses accessible from a Brazilian Public University, the researchers assessed the efficacy of four estimation techniques: one from distance education and the other from on-campus. The results have shown that the techniques tested in their study are capable of detecting students who are likely to fail early, improving the effectiveness of some of these techniques after applying pre-processing and/or fine-tuning algorithms, and statistically substantially outperforming the support vector machine technique.

3. RESEARCH METHOD

Systematic Literature Review (SLR) is characterized by covering a set of previous studies related to a specific topic and drawing accurate and comprehensive results based on the particular strategy. Systematic reviews have become widely used as a method to summarize research evidence rather than expert commentaries and narrative reviews [24]. A significant difference between literature review and other methods of analysis such as narrative reviews and meta_analysis) is that it follows a rigorous approach in restricting literature to collect accurate information [25, 26]. A narrative review is a more incidental and less comprehensive method, it is describing a group of previous studies arranged thematically and eliciting comprehensive results from the trends impressions. Meta-analysis is a statistical procedure band for merging outcomes from quantitative studies to interpret general trends and it often concentrates on a single variable relationship.

The systematic literature review should always have a protocol [27]. Consequently, the researchers in this paper relied on a protocol that was accomplished depending on a series of steps located by Kitchenham and Charters [28]. Systematic review steps are shown in figure 1.

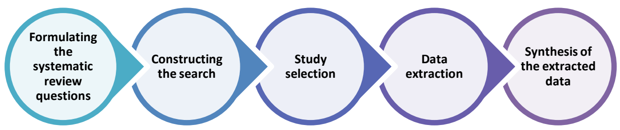


Figure 1: Systematic literature review steps

3.1 Formulating the systematic review questions

The first step in a systematic literature review is the formulation of research questions that the study strives to answer. In this research to identify the most important algorithms that are based on supervised machine learning and particularly are used in prediction, the researchers identified the first question:

What are the most algorithms of supervised machine learning that widely used in prediction?
The main objective of the systematic literature review is to answer the definite research question, instead of mentioning summaries about the field in which the researchers are interested in general [29]. The first research question was formulated to summarize the literature and create research on all algorithms used in the field of prediction. Certainly, the previous studies used many algorithms in the prediction field, that led researchers to formulate other research question:
What is the supervised learning algorithm most commonly used in the field of prediction?
Based on exploring the EDM field, a lot of sectors will be presented. This will led us to find and explore the different sectors of EDM which took the researcher's interest. Accordingly, the next question will be:
What is the sector that faced a lot of attention in the field of EDM?

3.2 Constructing the search

To search for previous studies, it is necessary to specify a set of relevant keywords that will be adopted in the research, in addition to specifying the digital libraries available to researchers used to search to answer the research questions.

3.2.1 Search process

The researchers focused on the group of digital libraries that can be accessed for research, as follows: IEEE electronic library [30-62], ACM digital library [63-82], Science Direct [21, 22, 83-104], Springer Link, and Wiley online library [105-116]. The choice of digital libraries was made according to the available libraries to the researchers.

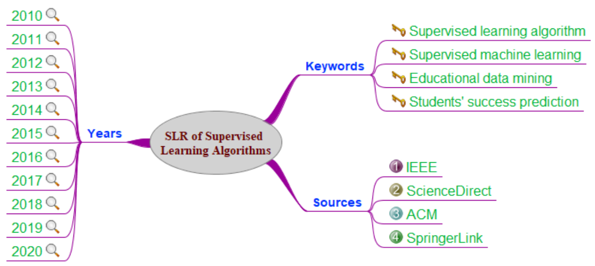


Figure 2: Search strategy of systematic literature review

3.2.2 Terminology

Terminology is a keyword or a set of words that will be used to get related studies on a specific topic. Therefore, researchers defined a set of research terms to define previous studies to obtain the required answers to the research questions. The basic search strings include the following terms: “supervised machine learning”, “supervised learning algorithm”, “Educational data mining”, “students’ success prediction” and “prediction system”.

3.3 Study selection

To determine which studies will be included and others to be excluded, a set of criteria will be applied to exclude some previous studies. These criteria include: All studies unrelated to the research topic will be excluded, all studies within the study period will be included, all studies outside the period (from 2010 to 2020) will be excluded, studies that do not include an algorithm related to the supervised learning will be excluded, one study will be selected if the study is duplicated in more than one digital library, concentrate on studies written in English only, and other studies that do not include prediction in EDM will also be excluded. Figure 2 illustrates the search strategy of systematic literature review.

3.4 Data Extraction

This step includes a detailed table for each study that was included according to the criteria in the systematic review [117]. The purpose of this table is to document the information obtained from preliminary studies accurately and transparently by researchers. This study will rely on the strategy presented by Salvado, Nakasone, and Pow-Sang in data extraction [118]. The strategy form includes the following details: (1) Title of study, (2) Author(s), (3) Publication type, (4) Date of extraction, and (5) Name of the digital library in which the study was found, and extra information related to the accuracy of algorithms and scope of research. The numerical data that will be produced from this form is of utmost importance to summarize the results of the previous set of studies.

3.5 Synthesis of the extracted data

It is the last stage of the systematic review steps that includes collecting and summarizing the results obtained from the included studies. It consists of three rounds: In the first round, all studies relevant to the field of research will be included. In the second round, studies that include supervised algorithms with educational prediction systems, in general, will be determined. Finally, in the third round the studies that meet all the inclusion criteria, in addition to the ability to summarize their results to answer the research question accurately will be determined. There are 90 studies extracted in the methodology stages that meet all criteria. Figure 3 shows details regarding the selected studies that were found during the search process, and all the extract information is detailed in the below link: https://drive.google.com/file/d/1mju2nVlh4aKsDP-BwApJ52KF5icuiTXY/view?usp=sharing

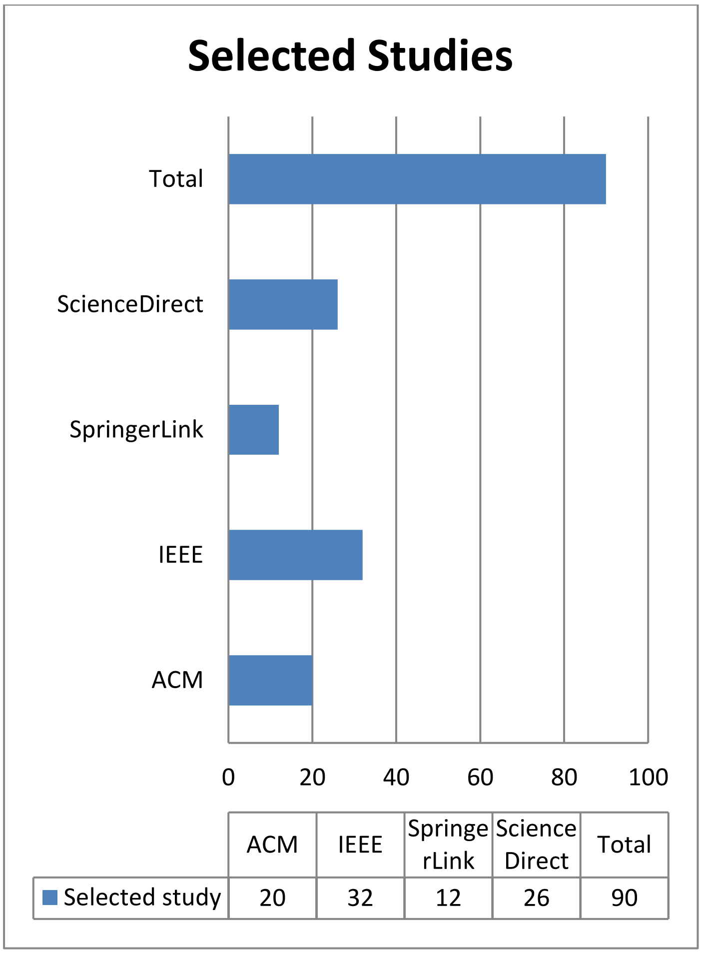


Figure 3: The total number of studies selected

4. RESULTS AND DISCUSSION

To understand the supervised learning algorithms used in the field of prediction that is used in educational data mining, systematic review questions were established to define the relevant algorithms for supervised learning in the field of prediction. The research questions were answered by examining or reviewing the literature on educational data mining. The first research question was answered by exploring research related to prediction within educational data containers and finding the algorithms used for this purpose. The results extracted during the application of the steps of a systematic review of the literature found that there was a set of algorithms used for prediction, the algorithms are as follows: decision tree, artificial neural network, support vector machine, logistic regression, ZeroR, k-nearest neighborhood, linear classifier, ensemble model, genetic programming, conditional random fields, Naive Bayes, association rules mining as depicted in Figure 4.

Southeast Europe Journal of Soft Computing
Available online: http://scjournal.ius.edu.ba

ISSN 2233 – 1859

Vol 10, No. 1, March 2021, pp. 55-70

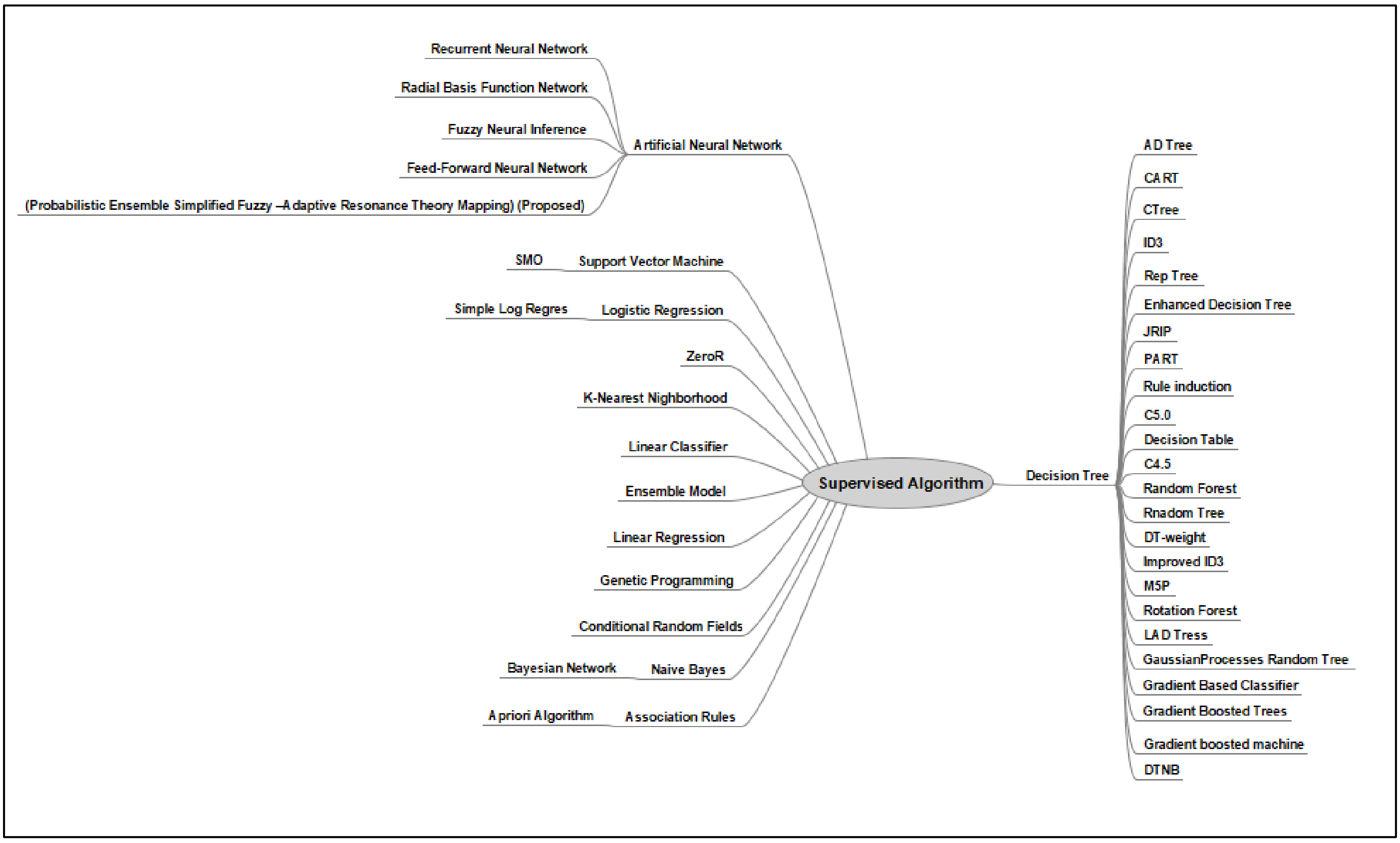


Figure 4: Supervised learning algorithms used in prediction field

The data attributes can be divided into two types: Continuous and categorical. The continuous variables hold the data with numeric values, and some mathematical operations such as integer, interval-scaled, and ratio-scaled variables can be performed on them. The categorical variables hold the data with nominal values where the mathematical operations and ordering such as nominal, binary, and ordinal variables cannot be performed on them [119, 120].
Regression is a data mining technique for predicting the numeric values in the class label. The regression analysis is one of the used statistical methods with numeric class prediction. It also includes the identification process of distribution trends in the training dataset [121]. A simple form of regression is the linear regression in which given two variables, it predicts the relationship between those two variables.
Classification is a process of dividing objects and assigning them into exclusive and exhaustive categories called classes [120]. While both regression and classification are used in the prediction, but they differ in the method of prediction as the first is used to predict continuous values while the latter is used to assign the objects into discrete categories. In classification, each object should be assigned to one and only one class. The problems where the goal is to find a specific value in a particular feature (class label) in the data can be solved by classification. The approaches of solving such data mining problems are supervised where there is a relationship between the class label and the remaining features [119]. The model in the classification is built based on analyzing the train data set (a set of objects where the class labels are previously known) to find and predict the unknown class label of objects. The conducted model can be presented in different forms such as mathematical forms, decision trees, (IF Then) classification rules, and neural networks [121, 122]. Below some used algorithms in regression and classification are discussed.
Decision Tree (DT) algorithms are used for classification where the algorithms construct a flow-chart-like tree structure. Each non-leaf internal node in the tree refers to a test on a variable and each branch is the test outcome(s). The external terminal node represents the class label prediction. The algorithm used to build the tree selects the best variable to perform

splitting into classes [123]. Different algorithms are developed and examined to handle categorical and numeric attributes in different sectors such as ID3, C4.5, RepTree, Random Forest, Hoeffding Tree, Decision Stump, LMT, Random Tree [124, 125].

Naïve Bayes (NB) is one of the classification algorithms that perform quite well in different domains. Naïve assumption comes from assuming the conditional independencies among variables. Bayes comes from using the Bayesian theorem a mathematical probability theory to find the possible classification [126].

The Bayesian theorem is stated as follows [127]:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Where the P(A) is a probability of an event (A), P(B) is a probability of an event (B) and P(A|B) is a probability of an event (A) conditional to the probability of the event (B). With NB algorithm, the prior probability is combined with conditional probabilities in a formula that can be used to calculate the possible probabilities to get classifications of a large dataset. Despite the assumption that the effect of a value of one feature or attribute on a given probability of class label is independent of the values of other features, the NB algorithm, in general, gives acceptable results [120].

Support Vector Machine (SVM) is developed by Vapink based on the principle of theoretical learning. This algorithm is used in many regression and classification fields besides outlier detection. It represents the base concept of risk minimization. In SVM, the kernel is used to map the original input space into a high-dimensional dot space. This space represents the feature space that defines the optimal hyperplane. The hyperplane is determined by data points (support vectors). Although SVM is not often used in problem-domain, it can provide a very strong generalized output for problems of classification [128, 129]. The output of the linear form of SVM is computed as:

$$o = \vec{x} \,.\, \vec{a} - b$$

Where $\vec{a}$ is the input vector and $\vec{x}$ is the normal vector to the hyperplane. Sequential Minimal Optimization (SMO) proposes a fast and easy method to solve the problem of quadratic optimization by optimizing the minimum subset through each iteration. The overall problem is separated into small collections of problems to be solved analytically. Small subsets of the training dataset are managed where a linear memory is needed. SMO performs faster than SVM due to regular chunking of the SVM algorithm to scale between cubic and linear [130].

Artificial Neural Networks (ANN) is one of the most important algorithms where a neuron has constructed the network to perform deep learning. Three layers represent the overall network (input, hidden, and output) layers where these layers consist of many units (represent neurons) to simulate the human brain work. The neurons in the layers are connected and the direction of data can determine the ANN algorithm. These units can detect the interactions amongst features. These units can perform feature detection even if they are independent. Many algorithms are developed in this field, the most used algorithms are feed-forward and recurrent neural networks [124].

K-nearest neighbors (KNN) algorithm is one of the simplest machine learning algorithms where each object will have a grade based on the majority vote of its neighbors. First, the object is assigned to the closest class amongst all its neighbors. K is a small positive number where if it sets to 1, then the object is assigned to the closest class from its neighbors. When K is an odd number in the binary classification (two final classes), the tied votes elimination will be easy [131].

Association Rules Mining (ARM) is used to measure the association amongst items in the dataset. The result association rules take the form of IF Then where the left part of the rule is called antecedent while the right part is called consequent [132, 133]. There are two important measurements (support and confidence) used to measures the accuracy of the resulting rules. The support value measures the appearance of the items while the confidence measures the accuracy of the rule.

$$\text{Support}(\{X\} \rightarrow \{Y\}) = \frac{Transactions\ containing\ both\ X\ and\ Y}{Totall\ number\ of\ transactions}$$

$$\text{Confidence}(\{X\} \rightarrow \{Y\}) = \frac{Transactions\ containing\ both\ X\ and\ Y}{Transactions\ containing\ X}$$

These factors may be predefined by minimal values called thresholds. Generally, association rules and classification rules do not differ, except that the association rules can be used to predict any attribute or a combination of attributes. The second difference is classification rules intend to be used together while association rules are not. The rules in ARM can express many regularities in the dataset which represent the prediction of many different classes. The most important algorithm in this field is the Apriori algorithm [134].

Linear and Logistic regression are determined by the class label. When the class label and all attributes are numeric values, then the linear is the natural regression technique to consider. Linear is a stable technique in statistics where the idea is to define the class as a combination of weighted attributes:

$$x = w_0 + a_1w_1 + a_1w_2 + a_3w_3 + \ldots + a_kw_k$$

Where a represent the attribute, w represents the weight, and x represents the class.

The result set from the calculating values from the training dataset is a numeric weighted set that can be used to measure the class label of new predicted instances. For decades, linear regression is widely used and proved its accuracy, simplicity, and easiness in statistical applications. However, linearity is the disadvantage of this algorithm. When a non-linear dependency is exhibited in the data, it can found the best fitting line which is represented by the mean square difference. Linear regression is widely used in the classification of data sets with numeric attributes. Linear or nonlinear regression can be used for classification. When a regression is performed for each class, the output for the training instances is set to 1 where the instance belongs to the class and the output is set to 0 for the instances that do not belong to the class. This outputs a class that is the linear regression. However, the value may be improper and take an observation out of range [0, and 1]. Logistic regression is another statistical technique that overcomes the problems of restricted values and least square regression. Instead of the approximate value, the value will be approximated when the value is exceeded, and hence the logistic regression model builds a model based on linear regression of the transformed target feature [135].
The clustering techniques are apart from unsupervised machine learning approaches where the class label is unknown [136]. The data is considered as objects and these objects are grouped into clusters based on similarity among these objects. The similarity is measured by the distance among objects. Cluster quality may be determined by the diameter or centroid distance. The maximum distance between two objects represents the diameter, while the average distance or the Euclidian distance may measure the centroid distance. Clustering techniques can also be used in the field of supervised machine learning by determining the class label. Many clustering algorithms such as K-means and hierarchical algorithms are proved their reliability and accuracy [137].
In this study, researchers also strived to answer the second research question related to the most used supervised algorithm in the field of prediction. A different set of supervised algorithms have been used to measure the accuracy of the data in a literature review (selected studied). These algorithms were used in the selected studies to measure accuracy in many scopes of research: predicting student performance, predicting student behavior, predicting the probability of students' degree completion, predicting a job after graduation, predicting future events, predicting to improve students final grade…etc. All the papers included in this review published in the period 2010-2020 are either journal or conference papers. Figure 5 lists these papers with their category according to year.

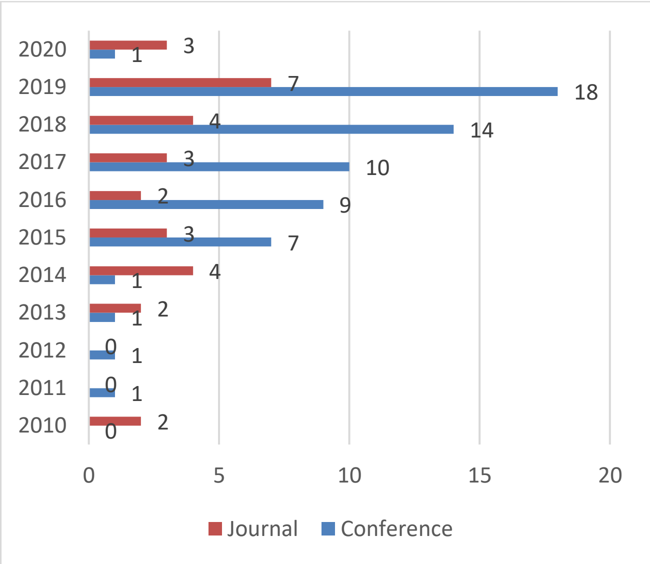
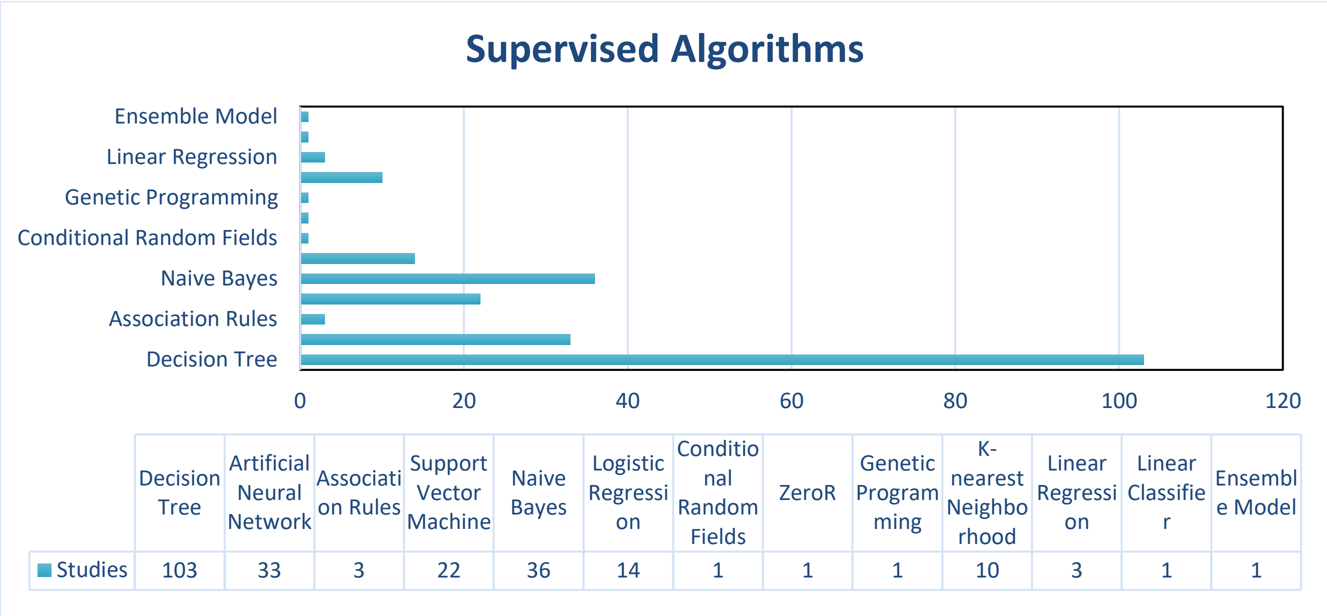


Figure 5: Journal vs conference EDM papers

Conference papers take a high percentage of the published papers. In the year 2010, two conference papers are published, while in the year 2011, and 2012 only one journal paper was published each year. In the year 2013, two journal papers and one conference paper are published, while in 2014, four journal papers and one conference paper are published. In 2015, seven conference papers and three papers are published, while in 2016 nine conference papers and two journal papers are published. In the years 2017, 2018, 2019, and 2020, ten conference papers and three journal papers, fourteen conference papers and four journal papers, eighteen conference papers and seven journal papers, and three journal papers along with one conference paper are published, respectively.
Concerning the second research question, the vast majority of studies adopted the use of DT algorithms, as they were used more than 100 times in various studies, and in some cases, more than one DT algorithm was used in the same study. ANN and NB algorithms used fewer than DT algorithms, where they were used 33 and 36 times, respectively. SVM and LR algorithms were used 22 and 14 times, respectively. K-nearest

Neighborhood algorithm was used 10 times in the studies. The use of the rest of the algorithms has been varied from one to three times. Figure 6 illustrates the use of algorithms in the selected studies.



	Decision Tree	Artificial Neural Network	Association Rules	Support Vector Machine	Naive Bayes	Logistic Regression	Conditional Random Fields	ZeroR	Genetic Programming	K-nearest Neighborhood	Linear Regression	Linear Classifier	Ensemble Model
Studies	103	33	3	22	36	14	1	1	1	10	3	1	1

Figure 6: The use of algorithms in the selected studies

Figure 7 illustrates the usage of data mining algorithms, namely DT, LR, SVM, ARM, ANN, NB, and KNN, according to the covered years. In the first five years, the usage of data mining algorithms was less than the next years. In the period 2015-2019, the data mining algorithms usage increased proportionally, while this usage reduced in 2020. In 2019, the usage of all algorithms reached its highest number, and among other approaches, DT was the most used one.

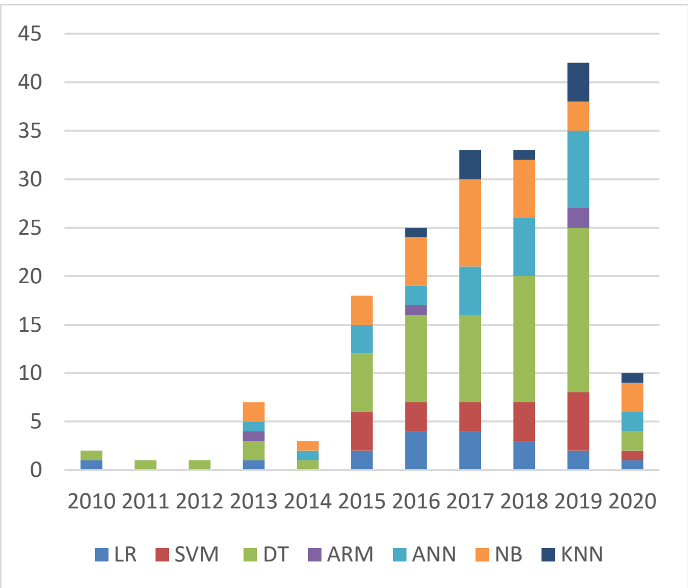


Figure 7: Data mining algorithms according to years

It can be seen that DT is used every year from 2010-2020 in the EDM comparing with other algorithms that are used in some period of the covered years.

Since the EDM field holds many sectors, it is necessary to explore the researches according to the educational sector. This step will help to find the most important sector which the researchers concentrated on in their works. Predicting students' performance, analyzing and predicting students' dropout, predicting students' actual degrees, student job after graduation, and analyzing and identifying students' activities and background are the most important sectors in the educational field. Exploring students' profiles, the response of the learners, predicting both educational future events and instructors' performance, enhancing learners adaptation system, improving students' skills and performance are the other sectors that the researchers worked on. Table 1 lists all the educational sectors cited in the studies.

Table 1: EDM sectors

Reference	Educational Sector
[22, 30, 31, 35, 37, 42-46, 48, 49, 51, 52, 54, 58, 59, 62, 65-71, 74, 76, 78-81, 83, 85, 86, 88, 89, 93, 94, 96, 97, 100, 101, 103-105, 107, 111, 112, 114-116, 138, 139]	Predicting students' performance.

[22, 55, 60, 72, 73, 77, 98, 140]	Analyzing and predicting students' dropout.
[33, 34, 53, 57, 82, 109]	Predicting student' degrees.
[36]	Predicting of a job after graduation.
[38, 50, 56, 90, 91, 95]	Analyzing students' activities and background.
[39]	Identifying students' profiles.
[40]	Identifying the fast response learners.
[42]	Predicting educational future events.
[61, 87, 102]	Enhancing learners adaptation system.
[21, 63, 64, 84]	Improving students' skills and performance.
[91]	Predicting instructors' performance.
[99, 106, 108, 113]	Predicting course selection and completion.
[110]	Prediction graduation rate.

The most sector that got the researchers' interest was predicting students' performance followed by analyzing and predicting students' dropout, predicting students' degrees, analyzing students' activities and background, predicting course selection and completion, and improving students' skills and performance. The other sectors such as predicting job after graduation, identifying students' profiles, identifying the fast response learners, predicting educational future events, enhancing learners adaptation system, predicting instructors' performance, and predicting graduation rate took less attention of the researchers.

In this study, researchers sought to answer SLR questions related to supervised machine learning algorithms used in prediction in the educational field, in addition to clarifying which algorithms are mostly used. The period of conducting the study was 10 years for the published papers from 2010 to 2020 according to specific criteria for selecting studies. The first question was answered by reviewing all the algorithms used in prediction. As for the second methodology question, the study found that algorithms of DT, ANN, and NB are the most popular algorithms among supervised learning algorithms in the field of prediction.

5. CONCLUSION AND DISCUSSION

EDM is one of the most important disciplines that explores and discovers hidden patterns in educational data. This field exploits different data mining algorithms, statistical analysis, and machine learning algorithms over different types of data sets with a different number of dimensions in the educational sector. The main goal behind implementing systems for predicting students' performance is to get highly accurate results with high responsive speed based on educational data. The nature and size of source data, number of features, size of noise within data, outliers, and dirty data are the most important factors that affect the classifier accuracy. Besides, choosing the right algorithm to handle the data also affects the accuracy. The nature of data enforce the analysts to perform data preprocessing to improve the data quality and then the knowledge conduced. Since the EDM field holds many sectors, it is necessary to explore the researches according to the educational sector.

The vast majority of studies adopted the use of DT algorithms, as they were used more than 100 times in various studies. In several studies, more than one DT algorithm was used in the same study. ANN and NB algorithms used fewer than DT algorithms. These two algorithms were used 33 and 36 times, respectively. As for the SVM and LR algorithms, they were used 22 and 14 times, respectively. The K-nearest Neighborhood algorithm was used 10 times in the studies.

As a comparison between journal and conference papers, the conference papers are the most published category. In the period 2015 to 2019, the conference papers increased proportionally, where this period reached its highest number in 2019. In 2019, the usage of all algorithms reached the peak where the DT was the most used algorithm among all other algorithms. It has been noticed that in the EDM, DT is used by different researchers in all years during the ten years of the studied period comparing to other algorithms that have been used in some years during the studied period.

EDM holds many sectors and disciplines, the sector of predicting students' performance toke the researchers' interest followed by analyzing and predicting students' dropout, predicting student' degrees, analyzing students' activities and background, predicting course selection and completion, and improving students' skills and performance. The other sectors such as predicting job after graduation, identifying students' profiles, identifying the fast response learners, predicting educational future events, enhancing learners adaptation system, predicting instructors' performance, and predicting graduation rate took less attention of the researchers.

The field of EDM is very important as it answers an urgent need to explore the variables that affect the learners' progress and present that in their works. This effect can be measured according to some features, data set size and number of instances in each work. Many researchers did not clarify the basic information about the educational data used in implementing their models. The relationship among size, data type, and the number of features of the educational dataset affects the accuracy of the model. In the future, a study can be made to explore the relationship between educational data type, size, features with the model accuracy.

REFERENCES

[1] R. S. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," JEDM| Journal of Educational Data Mining, vol. 1, pp. 3-17, 2009.

[2] C. Romero and S. Ventura, "Data mining in education," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 3, pp. 12-27, 2013.

[3] M. C. Mihaescu and P. S. Popescu, "Review on publicly available datasets for educational data mining," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, p. e1403, 2021.

[4] C. Romero and S. Ventura, "Educational data mining: a review of the state of the art," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, pp. 601-618, 2010.

[5] J. Ranjan and K. Malik, "Effective educational process: a data-mining approach," Vine, 2007.

[6] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, p. e1355, 2020.

[7] P. Ihantola, A. Vihavainen, A. Ahadi, M. Butler, J. Börstler, S. H. Edwards, et al., "Educational data mining and learning analytics in programming: Literature review and case studies," Proceedings of the 2015 ITiCSE on Working Group Reports, pp. 41-63, 2015.

[8] A. Charitopoulos, M. Rangoussi, and D. Koulouriotis, "On the use of soft computing methods in educational data mining and learning analytics research: A review of years 2010–2018," International Journal of Artificial Intelligence in Education, vol. 30, pp. 371-430, 2020.

[9] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Mining in educational data: review and future directions," in Joint European-US Workshop on Applications of Invariance in Computer Vision, 2020, pp. 92-102.

[10] A. Hamoud, H. Adday, T. Obaid, and R. Hameed, "Design and Implementing Cancer Data Warehouse to Support Clinical Decisions," International Journal of Scientific & Engineering Research, vol. 7, pp. 1271-1285, 2016.

[11] A. K. Hamoud, A. S. Hashim, and W. A. Awadh, "CLINICAL DATA WAREHOUSE: A REVIEW," Iraqi Journal for Computers and Informatics, vol. 44, 2018.

[12] A. Hamoud and T. Obaid, "Building Data Warehouse for Diseases Registry: First step for Clinical Data Warehouse," International Journal of Scientific & Engineering Research, vol. 4, pp. 636-640, 2013.

[13] A. Hamoud and T. A. S. Obaid, "Design and Implementation Data Warehouse to Support Clinical Decisions Using OLAP and KPI," Department of Computer Science, University of Basrah, 2013.

[14] A. Hamoud and T. Obaid, "Using OLAP with Diseases Registry Warehouse for Clinical Decision Support," International Journal of Computer Science and Mobile Computing, vol. 3, pp. 39-49, 2014.

[15] A. M. Humadi and A. K. Hamoud, "Liver Hepatitis Diagnosing based on Fuzzy Inference System," Journal of Southwest Jiaotong University, vol. 54, 2019.

[16] A. Humadi and A. Hamoud, "Online Real Time Fuzzy Inference System Based Human Health Monitoring and Medical Decision Making," International Journal of Computer Science and Information Security, vol. 15, pp. 197-204, 2017.

[17] H. N. H. Al-Hashimy, A. K. Hamoud, and F. A. Hussain, "The Effect of Not Using Internet of Things in Critical life Situations in the Health Field and the Effect on Iraqi Profitability: Empirical Study in Basra," Journal of Southwest Jiaotong University, vol. 54, 2019.

[18] W. A. Awadh, A. S. Hashim, and A. K. Hamoud, "A REVIEW ON INTERNET OF THINGS ARCHITECTURE FOR BIG DATA PROCESSING," Iraqi Journal for Computers and Informatics, vol. 46, pp. 11-19, 2020.

[19] W. A. Awadh, A. S. Hashim, and A. Hamoud, "A Review of Various Steganography Techniques in Cloud Computing," University of Thi-Qar Journal of Science, vol. 7, pp. 113-119, 2019.

[20] R. Baker, "Data mining for education," International encyclopedia of education, vol. 7, pp. 112-118, 2010.

[21] R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," Computers & Education, vol. 113, pp. 177-194, 2017.

[22] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. Van Erven, "Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil," Journal of Business Research, vol. 94, pp. 335-343, 2019.

[23] E. B. Costa, B. Fonseca, M. A. Santana, F. F. de Araújo, and J. Rego, "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses," Computers in Human Behavior, vol. 73, pp. 247-256, 2017.

[24] P. Hemingway and N. Brereton, "What is a systematic review', What is Series," Bandolier, April, 2009.

[25] J. H. Littell, J. Corcoran, and V. Pillai, Systematic

reviews and meta-analysis: Oxford University Press, 2008.
[26] H. Cooper, A. B. Allen, E. A. Patall, and A. L. Dent, "Effects of full-day kindergarten on academic achievement and social development," Review of educational research, vol. 80, pp. 34-70, 2010.
[27] A. O'Connor, K. Anderson, C. Goodell, and J. Sargeant, "Conducting systematic reviews of intervention questions I: writing the review protocol, formulating the question and searching the literature," Zoonoses and public health, vol. 61, pp. 28-38, 2014.
[28] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.
[29] D. Gough, S. Oliver, and J. Thomas, An introduction to systematic reviews: Sage, 2017.
[30] C. Jalota and R. Agrawal, "Analysis of Educational Data Mining using Classification," in 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 243-247.
[31] K. I. M. Ramaphosa, T. Zuva, and R. Kwuimi, "Educational data mining to improve learner performance in Gauteng primary schools," in 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), 2018, pp. 1-6.
[32] O. Sukhbaatar, K. Ogata, and T. Usagawa, "Mining Educational Data to Predict Academic Dropouts: a Case Study in Blended Learning Course," in TENCON 2018-2018 IEEE Region 10 Conference, 2018, pp. 2205-2208.
[33] R. Ahuja and Y. Kankane, "Predicting the probability of student's degree completion by using different data mining techniques," in 2017 Fourth International Conference on Image Information Processing (ICIIP), 2017, pp. 1-4.
[34] K. Bunkar, U. K. Singh, B. Pandya, and R. Bunkar, "Data mining: Prediction for performance improvement of graduate students using classification," in 2012 Ninth International Conference on Wireless and Optical Communications Networks (WOCN), 2012, pp. 1-5.
[35] S. S. Athani, S. A. Kodli, M. N. Banavasi, and P. S. Hiremath, "Student academic performance and social behavior predictor using data mining techniques," in 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 170-174.
[36] P. Rojanavasu, "Educational data analytics using association rule mining and classification," in 2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON), 2019, pp. 142-145.
[37] F. J. Kaunang and R. Rotikan, "Students' Academic Performance Prediction using Data Mining," in 2018 Third International Conference on Informatics and Computing (ICIC), 2018, pp. 1-5.
[38] C.-C. Kiu, "Data Mining Analysis on Student's Academic Performance through Exploration of Student's Background and Social Activities," in 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), 2018, pp. 1-5.
[39] J. d. O. Kelly, A. G. Menezes, A. B. de Carvalho, and C. A. Montesco, "Supervised learning in the context of educational data mining to avoid university students dropout," in 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), 2019, pp. 207-208.
[40] V. Mhetre and M. Nagar, "Classification based data mining algorithms to predict slow, average and fast learners in educational system using WEKA," in 2017 International Conference on Computing Methodologies and Communication (ICCMC), 2017, pp. 475-479.
[41] K. Parmar, D. Vaghela, and P. Sharma, "Performance prediction of students using distributed data mining," in 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015, pp. 1-5.
[42] H. Göker, H. I. Bülbül, and E. Irmak, "The estimation of students' academic success by data mining methods," in 2013 12th International Conference on Machine Learning and Applications, 2013, pp. 535-539.
[43] B. Guo, R. Zhang, G. Xu, C. Shi, and L. Yang, "Predicting students performance in educational data mining," in 2015 International Symposium on Educational Technology (ISET), 2015, pp. 125-128.
[44] T. Devasia, T. Vinushree, and V. Hegde, "Prediction of students performance using Educational Data Mining," in 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), 2016, pp. 91-95.
[45] N. Ketui, W. Wisomka, and K. Homjun, "Using classification data mining techniques for students performance prediction," in 2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON), 2019, pp. 359-363.
[46] M. P. Martins, V. L. Migueis, and D. Fonseca, "A data mining approach to predict undergraduate students' performance," in 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), 2018, pp. 1-7.
[47] R. Patil, S. Salunke, M. Kalbhor, and R. Lomte, "Prediction system for student performance using data mining classification," in 2018 Fourth International Conference on Computing Communication Control and Automation

(ICCUBEA), 2018, pp. 1-4.
[48] B. Al Breiki, N. Zaki, and E. A. Mohamed, "Using Educational Data Mining Techniques to Predict Student Performance," in 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), 2019, pp. 1-5.
[49] C. E. L. Guarín, E. L. Guzmán, and F. A. González, "A model to predict low academic performance at a specific enrollment using data mining," IEEE Revista Iberoamericana de tecnologias del Aprendizaje, vol. 10, pp. 119-125, 2015.
[50] N. Buniyamin, U. bin Mat, and P. M. Arshad, "Educational data mining for prediction and classification of engineering students achievement," in 2015 IEEE 7th International Conference on Engineering Education (ICEED), 2015, pp. 49-53.
[51] K. Shaukat, I. Nawaz, S. Aslam, S. Zaheer, and U. Shaukat, "Student's performance in the context of data mining," in 2016 19th International Multi-Topic Conference (INMIC), 2016, pp. 1-8.
[52] H. Goker and H. I. Bulbul, "Improving an early warning system to prediction of student examination achievement," in 2014 13th International Conference on Machine Learning and Applications, 2014, pp. 568-573.
[53] M. Wati, W. Indrawan, J. A. Widians, and N. Puspitasari, "Data mining for predicting students' learning result," in 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT), 2017, pp. 1-4.
[54] M. H. Rahman and M. R. Islam, "Predict Student's Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques," in 2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE), 2017, pp. 1-4.
[55] B. Perez, C. Castellanos, and D. Correal, "Applying data mining techniques to predict student dropout: a case study," in 2018 IEEE 1st Colombian Conference on Applications in Computational Intelligence (ColCACI), 2018, pp. 1-6.
[56] V. Shanmugarajeshwari and R. Lawrance, "Analysis of students' performance evaluation using classification techniques," in 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), 2016, pp. 1-7.
[57] H. Almayan and W. Al Mayyan, "Improving accuracy of students' final grade prediction model using PSO," in 2016 6th International Conference on Information Communication and Management (ICICM), 2016, pp. 35-39.
[58] M. S. Vyas and R. Gulwani, "Predictive analytics for E learning system," in 2017 International Conference on Inventive Systems and Control (ICISC), 2017, pp. 1-4.
[59] M. Nicoletti, M. Marques, and M. Guimaraes, "A data mining approach for forecasting students' performance," in 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), 2018, pp. 1-7.
[60] E. C. Medina, C. B. Chunga, J. Armas-Aguirre, and E. E. Grandón, "Predictive model to reduce the dropout rate of university students in Perú: Bayesian Networks vs. Decision Trees," in 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), 2020, pp. 1-7.
[61] T. Chellatamilan, M. Ravichandran, R. Suresh, and G. Kulanthaivel, "Effect of mining educational data to improve adaptation of learning in e-learning system," 2011.
[62] W. Quan and Q. Zhou, "Predicting the students with mental health risk by using Internet access logs," in 2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), 2018, pp. 140-140.
[63] C. Ma, B. Yao, F. Ge, Y. Pan, and Y. Guo, "Improving prediction of student performance based on multiple feature selection approaches," in Proceedings of the 2017 International Conference on E-Education, E-Business and E-Technology, 2017, pp. 36-41.
[64] A. Ahadi, V. Behbood, A. Vihavainen, J. Prior, and R. Lister, "Students' syntactic mistakes in writing seven different types of SQL queries and its application to predicting students' success," in Proceedings of the 47th ACM Technical Symposium on Computing Science Education, 2016, pp. 401-406.
[65] T. Sinha and J. Cassell, "Connecting the dots: Predicting student grade sequences from bursty MOOC interactions over time," in Proceedings of the second (2015) ACM conference on learning@ scale, 2015, pp. 249-252.
[66] W. Chango, R. Cerezo, and C. Romero, "Predicting academic performance of university students from multi-sources data in blended learning," in Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems, 2019, pp. 1-5.
[67] R. A. Rustia, M. M. A. Cruz, M. A. P. Burac, and T. D. Palaoag, "Predicting Student's Board Examination Performance using Classification Algorithms," in Proceedings of the 2018 7th International Conference on Software and Computer Applications, 2018, pp. 233-237.
[68] S. Altaf, W. Soomro, and M. I. M. Rawi, "Student Performance Prediction using Multi-Layers Artificial Neural Networks: A Case Study on Educational Data Mining," in Proceedings of the 2019 3rd International Conference on Information System and Data Mining, 2019, pp. 59-64.
[69] L. M. Barbosa Manhães, S. M. S. da Cruz, and G. Zimbrão, "Towards automatic prediction of student performance in STEM undergraduate degree programs," in Proceedings of the 30th Annual ACM

Symposium on Applied Computing, 2015, pp. 247-253.

[70] W. Puarungroj, N. Boonsirisumpun, P. Pongpatrakant, and S. Phromkhot, "Application of data mining techniques for predicting student success in English exit exam," in Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, 2018, pp. 1-6.

[71] P. Chaudhury, S. Mishra, H. K. Tripathy, and B. Kishore, "Enhancing the capabilities of student result prediction system," in Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, 2016, pp. 1-6.

[72] J. Figueiredo, N. Lopes, and F. J. García-Peñalvo, "Predicting student failure in an introductory programming course with multiple back-propagation," in Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality, 2019, pp. 44-49.

[73] H. Chanlekha and J. Niramitranon, "Student performance prediction model for early-identification of at-risk students in traditional classroom settings," in Proceedings of the 10th International Conference on Management of Digital EcoSystems, 2018, pp. 239-245.

[74] C. I. P. Benablo, E. T. Sarte, J. M. D. Dormido, and T. Palaoag, "Higher Education Student's Academic Performance Analysis through Predictive Analytics," in Proceedings of the 2018 7th International Conference on Software and Computer Applications, 2018, pp. 238-242.

[75] B. Sekeroglu, K. Dimililer, and K. Tuncal, "Student performance prediction and classification using machine learning algorithms," in Proceedings of the 2019 8th International Conference on Educational and Information Technology, 2019, pp. 7-11.

[76] F. Okubo, T. Yamashita, A. Shimada, and H. Ogata, "A neural network approach for students' performance prediction," in Proceedings of the seventh international learning analytics & knowledge conference, 2017, pp. 598-599.

[77] K. Kang and S. Wang, "Analyze and predict student dropout from online programs," in Proceedings of the 2nd International Conference on Compute and Data Analysis, 2018, pp. 6-12.

[78] S. N. Liao, D. Zingaro, K. Thai, C. Alvarado, W. G. Griswold, and L. Porter, "A robust machine learning technique to predict low-performing students," ACM Transactions on Computing Education (TOCE), vol. 19, pp. 1-19, 2019.

[79] M. V. Amazona and A. A. Hernandez, "Modelling Student Performance Using Data Mining Techniques: Inputs for Academic Program Development," in Proceedings of the 2019 5th International Conference on Computing and Data Engineering, 2019, pp. 36-40.

[80] S. Sisovic, M. Matetic, and M. B. Bakaric, "Mining student data to assess the impact of moodle activities and prior knowledge on programming course success," in Proceedings of the 16th International Conference on Computer Systems and Technologies, 2015, pp. 366-373.

[81] A. Daud, N. R. Aljohani, R. A. Abbasi, M. D. Lytras, F. Abbas, and J. S. Alowibdi, "Predicting student performance using advanced learning analytics," in Proceedings of the 26th international conference on world wide web companion, 2017, pp. 415-421.

[82] Y. Arora, A. Singhal, and A. Bansal, "PREDICTION & WARNING: a method to improve student's performance," ACM SIGSOFT Software Engineering Notes, vol. 39, pp. 1-5, 2014.

[83] M. Ashraf, M. Zaman, and M. Ahmed, "An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches," Procedia Computer Science, vol. 167, pp. 1471-1483, 2020.

[84] J. D. Gobert, Y. J. Kim, M. A. Sao Pedro, M. Kennedy, and C. G. Betts, "Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld," Thinking Skills and Creativity, vol. 18, pp. 81-90, 2015.

[85] C. Romero, M.-I. López, J.-M. Luna, and S. Ventura, "Predicting students' final performance from participation in on-line discussion forums," Computers & Education, vol. 68, pp. 458-472, 2013.

[86] W. Xing, R. Guo, E. Petakovic, and S. Goggins, "Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory," Computers in Human Behavior, vol. 47, pp. 168-181, 2015.

[87] A. A. Yahya and A. Osman, "Using Data Mining Techniques to Guide Academic Programs Design and Assessment," Procedia Computer Science, vol. 163, pp. 472-481, 2019.

[88] P. Kaur, M. Singh, and G. S. Josan, "Classification and prediction based data mining algorithms to predict slow learners in education sector," Procedia Computer Science, vol. 57, pp. 500-508, 2015.

[89] C. Burgos, M. L. Campanario, D. de la Peña, J. A. Lara, D. Lizcano, and M. A. Martínez, "Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout," Computers & Electrical Engineering, vol. 66, pp. 541-556, 2018.

[90] S. Natek and M. Zwilling, "Student data mining solution–knowledge management system related to higher education institutions," Expert systems with applications, vol. 41, pp. 6400-6407, 2014.

[91] N. Thai-Nghe, L. Drumond, A. Krohn-Grimberghe, and L. Schmidt-Thieme, "Recommender system for predicting student performance," Procedia Computer Science, vol. 1, pp. 2811-2819, 2010.

[92] A. M. Ahmed, A. Rizaner, and A. H. Ulusoy, "Using

data mining to predict instructor performance," Procedia Computer Science, vol. 102, pp. 137-142, 2016.
[93] S. Oeda and M. Chieda, "Visualization of Programming Skill Structure by Log-Data Analysis with Decision Tree," Procedia Computer Science, vol. 159, pp. 582-589, 2019.
[94] H. Zeineddine, U. Braendle, and A. Farah, "Enhancing prediction of student success: Automated machine learning approach," Computers & Electrical Engineering, vol. 89, p. 106903.
[95] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," Computers & Education, vol. 103, pp. 1-15, 2016.
[96] M. Bennett, L. Bormann, S. Lovan, and B. Cobb, "Preadmission predictors of student success in a baccalaureate of science in nursing program," Journal of Nursing Regulation, vol. 7, pp. 11-18, 2016.
[97] T. A. Cardona, "Predicting Student Retention Using Support Vector Machines," Procedia Manufacturing, vol. 39, pp. 1827-1833, 2019.
[98] A.-S. Hoffait and M. Schyns, "Early detection of university students with potential difficulties," Decision Support Systems, vol. 101, pp. 1-11, 2017.
[99] A. A. Kardan, H. Sadeghi, S. S. Ghidary, and M. R. F. Sani, "Prediction of student course selection in online higher education institutes using neural network," Computers & Education, vol. 65, pp. 1-11, 2013.
[100] S. Helal, J. Li, L. Liu, E. Ebrahimie, S. Dawson, D. J. Murray, et al., "Predicting academic performance by considering student heterogeneity," Knowledge-Based Systems, vol. 161, pp. 134-146, 2018.
[101] G. Lesinski and S. Corns, "Multi-objective evolutionary neural network to predict graduation success at the United States military academy," Procedia Computer Science, vol. 140, pp. 196-205, 2018.
[102] E. Howard, M. Meehan, and A. Parnell, "Contrasting prediction methods for early warning systems at undergraduate level," The Internet and Higher Education, vol. 37, pp. 66-75, 2018.
[103] G. Badr, A. Algobail, H. Almutairi, and M. Almutery, "Predicting students' performance in university courses: a case study and tool in KSU mathematics department," Procedia Computer Science, vol. 82, pp. 80-89, 2016.
[104] Y. Altujjar, W. Altamimi, I. Al-Turaiki, and M. Al-Razgan, "Predicting critical courses affecting students performance: a case study," Procedia Computer Science, vol. 82, pp. 65-71, 2016.
[105] A. Yağci and M. Çevik, "Prediction of academic achievements of vocational and technical high school (VTS) students in science courses through artificial neural networks (comparison of Turkey and Malaysia)," Education and Information Technologies, vol. 24, pp. 2741-2761, 2019.
[106] C. Beaulac and J. S. Rosenthal, "Predicting university students' academic success and major using random forests," Research in Higher Education, vol. 60, pp. 1048-1064, 2019.
[107] M. Ezz and A. Elshenawy, "Adaptive recommendation system using machine learning algorithms for predicting student's best academic program," Education and Information Technologies, pp. 1-14, 2019.
[108] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, "Using machine learning to predict student difficulties from learning session data," Artificial Intelligence Review, vol. 52, pp. 381-407, 2019.
[109] M. F. Musso, C. F. R. Hernández, and E. C. Cascallar, "Predicting key educational outcomes in academic trajectories: a machine-learning approach," Higher Education, pp. 1-20, 2020.
[110] Y. Nieto, V. García-Díaz, C. Montenegro, and R. G. Crespo, "Supporting academic decision making at higher educational institutions using machine learning-based algorithms," Soft Computing, vol. 23, pp. 4145-4153, 2019.
[111] Z. Nudelman, D. Moodley, and S. Berman, "Using bayesian networks and machine learning to predict computer science success," in Annual Conference of the Southern African Computer Lecturers' Association, 2018, pp. 207-222.
[112] D. Canagareddy, K. Subarayadu, and V. Hurbungs, "A machine learning model to predict the performance of university students," in International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering, 2018, pp. 313-322.
[113] J. Nouri, K. Larsson, and M. Saqr, "Identifying factors for master thesis completion and non-completion through learning analytics and machine learning," in European Conference on Technology Enhanced Learning, 2019, pp. 28-39.
[114] S. Almutairi, H. Shaiba, and M. Bezbradica, "Predicting Students' Academic Performance and Main Behavioral Features Using Data Mining Techniques," in International Conference on Computing, 2019, pp. 245-259.
[115] K. Agarwal, E. Maheshwari, C. Roy, M. Pandey, and S. S. Rautray, "Analyzing Student Performance in Engineering Placement Using Data Mining," in Proceedings of International Conference on Computational Intelligence and Data Engineering, 2019, pp. 171-181.
[116] S. Sivakumar and R. Selvaraj, "Predictive modeling of students performance through the enhanced decision tree," in Advances in electronics, communication and computing, ed: Springer, 2018, pp. 21-36.
[117] M. Petticrew and H. Roberts, Systematic reviews in the social sciences: A practical guide: John Wiley & Sons, 2008.
[118] C. Salvador, A. Nakasone, and J. A. Pow-Sang, "A

systematic review of usability techniques in agile methodologies," in Proceedings of the 7th Euro American Conference on Telematics and Information Systems, 2014, pp. 1-6.
[119] C. C. Aggarwal, Data mining: the textbook: Springer, 2015.
[120] M. Bramer, Principles of data mining vol. 180: Springer, 2007.
[121] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques third edition," The Morgan Kaufmann Series in Data Management Systems, vol. 5, pp. 83-124, 2011.
[122] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "Practical machine learning tools and techniques," Morgan Kaufmann, p. 578, 2005.
[123] I. A. Najm, A. K. Hamoud, J. Lloret, and I. Bosch, "Machine Learning Prediction Approach to Enhance Congestion Control in 5G IoT Environment," Electronics, vol. 8, p. 607, 2019.
[124] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student Performance Prediction Model based on Supervised Machine Learning Algorithms," in IOP Conference Series: Materials Science and Engineering, 2020, p. 032019.
[125] A. Hamoud, "Selection of best decision tree algorithm for prediction and classification of students' action," American International Journal of Research in Science, Technology, Engineering & Mathematics, vol. 16, pp. 26-32, 2016.
[126] A. Hamoud, A. Humadi, W. A. Awadh, and A. S. Hashim, "Students' success prediction based on Bayes algorithms," International Journal of Computer Applications, vol. 178, pp. 6-12, 2017.
[127] N. Ye, Data mining: theories, algorithms, and examples: CRC press, 2013.
[128] W. S. Noble, "What is a support vector machine?," Nature biotechnology, vol. 24, pp. 1565-1567, 2006.
[129] L. Wang, Support vector machines: theory and applications vol. 177: Springer Science & Business Media, 2005.
[130] L. J. Cao, S. S. Keerthi, C. J. Ong, J. Q. Zhang, U. Periyathamby, X. J. Fu, et al., "Parallel sequential minimal optimization for the training of support vector machines," IEEE Trans. Neural Networks, vol. 17, pp. 1039-1049, 2006.
[131] O. Kramer, Dimensionality reduction with unsupervised nearest neighbors: Springer, 2013.
[132] A. Hamoud, "Applying association rules and decision tree algorithms with tumor diagnosis data," International Research Journal of Engineering and Technology, vol. 3, pp. 27-31, 2017.
[133] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in Proceedings of the 1993 ACM SIGMOD international conference on Management of data, 1993, pp. 207-216.
[134] A. S. Hashima, A. K. Hamoud, and W. A. Awadh, "Analyzing students' answers using association rule mining based on feature selection," Journal of Southwest Jiaotong University, vol. 53, 2018.
[135] A. DeMaris, "A tutorial in logistic regression," Journal of Marriage and the Family, pp. 956-968, 1995.
[136] A. K. HAMOUD, "CLASSIFYING STUDENTS'ANSWERS USING CLUSTERING ALGORITHMS BASED ON PRINCIPLE COMPONENT ANALYSIS," Journal of Theoretical & Applied Information Technology, vol. 96, 2018.
[137] R. Xu and D. Wunsch, Clustering vol. 10: John Wiley & Sons, 2008.
[138] C. Verma, Z. Illés, and V. Stoffová, "Age group predictive models for the real time prediction of the university students using machine learning: Preliminary results," in 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019, pp. 1-7.
[139] H. Guruler, A. Istanbullu, and M. Karahasan, "A new student performance analysing system using knowledge discovery in higher educational databases," Computers & Education, vol. 55, pp. 247-254, 2010.
[140] G. J. Baars, T. Stijnen, and T. A. Splinter, "A model to predict student failure in the first year of the undergraduate medical curriculum," Health Professions Education, vol. 3, pp. 5-14, 2017.