An artificial intelligence approach to monitor student performance and devise preventive measures

Ijaz Khan[1,2]* , Abdul Rahim Ahmad[3], Nafaa Jabeur[4] and Mohammed Najah Mahdi[5]

*Correspondence:
ijaz@buc.edu.om
[1] College of Graduate Studies,
Universiti Tenaga Nasional,
Kajang, Malaysia
Full list of author information
is available at the end of the
article

Abstract

A major problem an instructor experiences is the systematic monitoring of students' academic progress in a course. The moment the students, with unsatisfactory academic progress, are identified the instructor can take measures to offer additional support to the struggling students. The fact is that the modern-day educational institutes tend to collect enormous amount of data concerning their students from various sources, however, the institutes are craving novel procedures to utilize the data to magnify their prestige and improve the education quality. This research evaluates the effectiveness of machine learning algorithms to monitor students' academic progress and informs the instructor about the students at the risk of ending up with unsatisfactory result in a course. In addition, the prediction model is transformed into a clear shape to make it easy for the instructor to prepare the necessary precautionary procedures. We developed a set of prediction models with distinct machine learning algorithms. Decision tree triumph over other models and thus is further transformed into easily explicable format. The final output of the research turns into a set of supportive measures to carefully monitor students' performance from the very start of the course and a set of preventive measures to offer additional attention to the struggling students.

Keywords: Artificial intelligence, Student performance prediction, Educational data mining, Machine learning, Decision tree, k-nn

Introduction

Students are the main stakeholders of the educational institutions. The performance of educational institutes plays an important role in producing paramount quality graduates and post-graduates. The modern-day educational institutes are trying to uphold quality and prestige in the education society. In fact, the institutes are more concerned about their prestige as compare to the quality of education (Norris et al., 2008). However, various government and accreditation agencies ensure the educational institutes sustain a high quality learning environment and the concrete procedures of accreditation has compelled the institutions to plan and implement novel procedures to preserve their standards. For instance, the Oman Academic Accreditation Authority (OAAA) and Accreditation Board for Engineering and Technology

(ABET) (Nettleman, 2018) in United States, ensures high-class educational institutes in the country. To preserve their position, the institutes are in quest of innovative practices.

The educational institutions implement novel technologies, for instance, Learning Management Systems (LMS), Intelligent Tutoring Systems (ITS), and online learning platforms which facilitate them to accumulate enormous quantity of data about the students and the learning environment (Gašević et al., 2015). The data may include the documentations of students, their behaviors, performance in assessment tools (exams etc.), interaction with online social forums, demographic data, and administrative data (Khan, 2018). The institutes require innovative practices to make the most of the collected data and augment their decision making approaches. Several computer technologies offers conveniences to renovate the complex material to make it easy to understand and remember (Romanenko et al., 2019). The data mining algorithms apply eminent techniques to the data and extracts momentous information (Alabri et al., 2019). Machine Learning is among the tools possessing the potential to support the education institutions in several states of affairs. The machine learning algorithms make use of the previous data and forecast the likelihood of an event with suitable precision (Khan et al., 2019).

A chief objective of the institutions is to monitor the students' academic performance, in a course, and identify the students with inadequate academic progress (Khan et al., 2019). The instructor may not be able to distinguish the level of students at the start of the course. However, if the struggling students are identified by some means then the instructor can design preventive measures to deal with them. Therefore, sophisticated prediction models are imperative to forecast the final outcome of the student and make it possible for the instructor to take care of the struggling students.

Several prediction models monitor student's performance in online learning (Umer et al., 2017; Yang et al., 2017). The models, for instance (Asogwa & Oladugba, 2015), identify the students who are more likely to drop the course at the beginning of the semester. Several authors design models (Costa et al., 2017; Khan et al., 2019) specifically to track the performance of novice computer language programmers. Numerous prediction models have been implemented (Alfere & Maghari, 2018; Asogwa & Oladugba, 2015; Jishan et al., 2015; Kiu, 2018; Lagman et al., 2019; Liao et al., 2019; Ma et al., 2018; Mondal & Mukherjee, 2018; Oladokun et al., 2008; Pandey & Taruna, 2016) to back up machine learning algorithms as useful tools to predict student's final outcome.

Learning is always local (Drachsler & Greller, 2012). The existing models are constructive solely in the local context. The students within different educational environment may respond in a different way. In this paper, we use machine learning algorithms to develop and choose a prediction model which is able to identify the students with poor performance. The model is further transformed into effortlessly explicable shape. The ultimate outcome of the model turns into a set of supportive measures to carefully monitor student performance and a set of preventive measures useful for inadequate students. The research identifies the key features which influence a student's final outcome. The paper is organized as; Sect. Literature review provides a literature review of artificial intelligence, machine learning algorithms and discusses several student performance prediction models. Research methodology Section provides the 3-step methodology

Khan *et al. Smart Learn. Environ.* (2021) 8:17

Page 3 of 18

used in this research. Section Data preparation, Experimental evaluation, and Model implementation explains each step of methodology in details. Section Model execution and results provide results from the field test and Sect. Conclusion and future work concludes the paper with future aims.

Literature review

Artificial Intelligence (AI) aspire towards providing adequate intelligence to computers so they can think and act in responses, similar to the human being (Lesinski et al., 2016). Unlike computers, human can learn from their experience which enables them to make intellectual decisions according to their individual circumstances. On the other hand, computer has to follow the man-made algorithms to accomplish the required task. Artificial Intelligence aims to lessen this dissimilarity between computer and human by seeking innovative techniques to equip computers with intelligence and enable them to act like human being. The term is often applied to projects which develop systems conferred on humans' distinct intellectual processes, for instance, the ability to think, discover meaning, or learn from previous experience. The AI applications are steadily growing within distinct commercial, service, manufacturing and agricultural industries, making its more prominent (Došilović et al., 2018). Future AI artefacts will be capable to interact with human beings in their native languages, and adapt to their movements and emotions (Lu, 2019).

Machine Learning is one of the AI applications to facilitate systems with the ability to automatically learn and improve from experience without any explicit programming (Mitchell et al., 2013). The prime goal is to enable computers to learn automatically and set the procedures to make future decisions (Nilsson, 2014). Machine Learning algorithms learn from the prearranged data and then make decisions for unseen data. Machine learning uses two major classes of algorithms: supervised learning and unsupervised learning. Supervised learning are either classification or regression algorithms. The classification algorithms comprises of input, output and the aim is to apply an algorithm to identify the mapping function from the input to the output (Qazdar et al., 2019). Each instance consists of independent variables (prediction features) and a dependent variable (prediction class). The algorithms process the entire training dataset and identify the patterns and rules hidden in the data. A model, constructed on the basis of the identified rules, gets unseen instances and classifies them in appropriate classes.

Some of the most widely used supervised learning algorithms are Artificial Neural Networks (ANN), Naive Bayes, k-Nearest Neighbors (k-NN), Support Vector Machines, and Decision Trees. Artificial Neural Networks (ANN) (Mitchell et al., 2013) is derived from the structural and functional features of the biological nervous system (Witten et al., 2016). Naïve Bayes (Domingos & Pazzani, 1997) is based on Bayes theorem of probability to classify the unseen data instance. The main assumption is that the input features are conditionally independent with familiar classification. K-NN stores training dataset in the memory and then compares each instance with the instances it has seen in the training process (Cunningham & Delany, 2007). Support Vector Machines (Suthaharan, 2016) plots the training instances in a n-dimensional space with separating hyperplane; the instances on each side of the hyperplane belong to same class. Decision tree follows a recursive technique to build

Khan *et al. Smart Learn. Environ.*   (2021) 8:17

Page 4 of 18

a tree (Li et al., 2019). Decision tree owns several conventional features making it a dominant choice for classification and prediction (Sunday et al., 2020). In contrast to classification algorithms, Regression algorithms, for instance linear regression, learn from the training dataset and develop model for continuous responses.

The unsupervised algorithms explore the hidden patterns and derive inferences from datasets that consists of input data without labeled classes. Clustering is the most common unsupervised learning algorithm. It identifies concealed patterns and makes cluster of data for exploratory analysis. Some of the popular clustering algorithms include k-Means clustering and Fuzzy clustering (Kassambara, 2017).

Machine Learning classification models are used in pedagogical environment to develop students' performance prediction models. These prediction models forecast the final outcome of the student based on several academic features. The main output of the model identifies the students with high probability of ending with unsatisfactory outcome. Once identified, these students can be forwarded for more auxiliary counseling mechanisms. A wide range of machine learning, particularly supervised algorithms, are used to put into operation the concept of student performance prediction modeling. Several challenges come about while developing machine learning models for educational environment. The training dataset, of a course, may comprise of low number of total instances due to limitation on classroom size. Generally, a smaller number of students end up with unsatisfactory outcome which leads to dataset with uneven ratio between the classes.

Numerous models have been proposed under different educational context to address the student performance prediction. Kausar et al. (2020) made use of ensemble techniques to examine the relationship between students' semester course and final results. The experimental evaluation concludes Random Forest and Stacking Classifiers with achieving the highest accuracy. Orong et al. (2020) used modified Genetic Algorithm (GA) to eliminate excessive features and applied decision tree algorithm to discover the weak students and thus facilitates the institution to design interference measure to raise the student attrition. Chen et al. (2018) built models with decision tree and linear regression with a set of features extorted from the institution's auto-grading system. The research assists the institution to recognize the struggling students and assign teaching hours automatically in a smart way. Saa (2016) proposed a decision tree model to discover the essential features which influence students' academic performance. The data related to students' demographic, academic and social behavior was collected through a survey. Iatrellis et al. (2020) proposed a machine learning approach wherein K-Means algorithm generates a set of coherent clusters and afterward supervised machine learning algorithms are used to train prediction models for predicting students' performance. Maesya and Hendiyanti (2019) developed model to forecast if the student will graduate on time or late than the standard graduation duration. Kiu (2018) examines the correlation between social activities and the final results of the students. Decision tree emerged as an useful tool, however, a weak correlation was examined between the two factors. Kaunang and Rotikan (2018) produced several models based on decision tree algorithm over a data containing student's demographics, academic and family background features collected through questionnaires. Yousafzai et al. (2020) applied decision tree and

regression algorithms over the historic performance of students and proposed a system to forecast students' grades.

There are a number of models which are based on decision tree (Kiu, 2018; Pandey & Taruna, 2016), lazy classifiers (Alfere & Maghari, 2018), Artificial Neural Networks (Asogwa & Oladugba, 2015; Mondal & Mukherjee, 2018; Oladokun et al., 2008). Naïve Bayes (Jishan et al., 2015; Lagman et al., 2019) and Support Vector Machine (Liao et al., 2019; Ma et al., 2018).

The literature review confirms machine learning algorithms as productive tools for developing models to predict student's final outcome. The existing models are useful locally and produce efficient results for single course. Therefore, we develop a prediction model for a course taught at the host institution. The aim of this research is not solely to develop prediction model but to interpret the model into easily understandable form. Further, the interpretation is described as precautionary measure for the students. This model implementation proposes appropriate measures for prior and post model execution.

Research methodology

Figure 1 illustrates the methodology used in this research. It is a 3-phased methodology and analogous kind of operations takes place in each phase. The foremost task is to define and prepare the training dataset. The data preparation deals with cleaning and pre-processing of data. The data cleaning involves elimination of irrelevant features and handling instances having missing values for feature(s). Data Pre-Processing, further improves the quality of dataset so the algorithms can generate improved results. The experimental evaluation phase executes a set of machine learning algorithms over the prepared dataset. Each algorithm produces a prediction model. The produced models are compared through several evaluation metrics and the model which appears robust is chosen for interpretation. The model implementation phase transforms the chosen model to a form easily understandable by the instructor. The concluding step proposes precautionary measures in the light of the transformed model.

Data preparation

Data description

The dataset in this research consists of the student academic records for a course taught et al.-Buraimi University College (BUC), Sultanate of Oman. The data is collected from

Fig. 1 The methodology used in this research
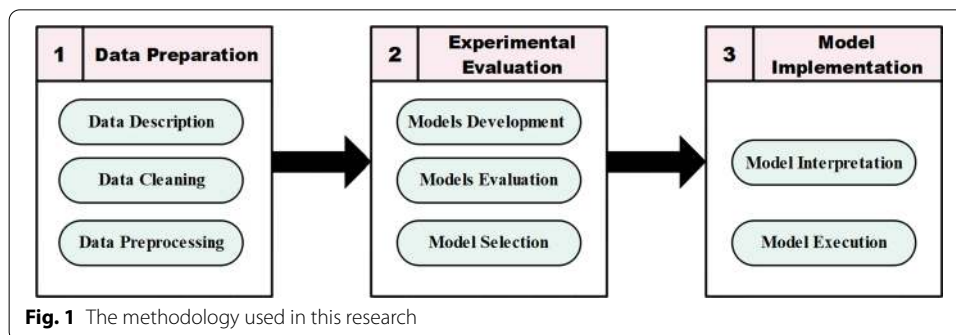
registration department of the college and all the ethical guidelines are followed carefully. The training dataset spans over a time periods of 3 semesters. There are total of 151 instances in the training dataset with 10 prediction features and one prediction class. The prediction class classifies students as either "Low" or "High". Table 1 provides the list of features along with their description.

The data consists of students' academic record from a face to face taught course "Phonetics and Phonology" which deals with the production of speech sounds and sounds patterns by humans without preceding knowledge of English language. The assessment consists of three exams of 15, 15 and 50 marks each, and an assignment of 20 marks. The first exam is taken at the 6th week of the semester and the assignment is usually assigned afterwards. It is necessary for the instructor to identify the struggling students soon after the end of exam-1. Since the first exam carries 15 marks hence, the students still have to work hard for the remaining 85 grade.

Data cleaning

Several features, known as irrelevant features, do not participate in the prediction rather they are associated with the student's privacy. To deal with the ethical and privacy restraints, we eradicate irrelevant features, for instance, student's ID, student name, and course code. Similarly, machine learning algorithms cannot understand and interpret noisy data properly. Noisy data take into account the instances having irrelevant or misleading values to one or more than one feature. Such instance demotes the algorithm's performance and therefore these features have to be dealt carefully. Table 2 demonstrates examples from our dataset with instances with noisy data. Several instances appear with missing value of PreReq_Grades and the last student did not conduct the exam-1 (Grade-1_Cont). The noise is either reduced through several techniques or such instances are removed from the dataset. Once the noisy instances are removed, our training dataset has 151 instances.

Data pre-processing

The training dataset usually consists of a large set of features but using entire set of features might relegate the classification result (Márquez-Vera et al., 2013). It is better to

Table 1 The list of student features with description

Feature	Description	Values
Gen	Gender of the student	Male, Female
Attendance	Attendance in the class	Real (Percent)
Major	Major of the student	Translation, Literature
Year	Year of study	From 1 to 4
Session	Whether student study in morning or evening	Morning, Evening
Grade-1_Cont	Grades in Exam-1	Real (Percent)
CGPA	Cumulative Grade Points Average of student	Real (0–4)
Sponsorship	Whether sponsored for study by government	Yes, No
Dorm	Whether resides in hostel	Yes, No
PreReq_Grades	Grades in the prerequisite subject	Real (Percent)
Final_Grade	Prediction Class	Low and High

Khan *et al. Smart Learn. Environ.*   (2021) 8:17

Page 7 of 18

Table 2 Examples of noisy instances in the dataset

Gender	Attendance	Major	Year	Session	Grade-1_Cont	CGPA	Sponsorship	Dorms	PreReq_Grades	Final_Grade
F	100	Literature	Year-3	Morning	80.0	3.04	Y	N	?	Low
F	81.25	Literature	Year-4	Evening	86.7	3.2	Y	Y	?	High
F	100	Translation	Year-3	Morning	96.7	3.63	Y	N	?	High
F	81.25	Literature	Year-3	Evening	?	3.2	Y	Y	75.5	?

Table 3 Features in descending order with respect to their gain ratio values

Features	GR value
CGPA	0.324522
PreReq_Grades	0.160408
Grade-1_Cont	0.138898
Attendance	0.09145
Gender	0.043883
Dorms	0.012948
Year	0.00544
Session	0.003749
Sponsorship	0.001073
Major	0.000547

Table 4 The significant features and their descriptive analysis

Feature	Minimum	Maximum	Mean	SD
CGPA	1.51	3.93	2.86	0.544
PreReq_Grades	50	99	76.818	12.784
Grade-1_Cont	56.7	100	84.459	9.712
Attendance	77.08	100	94.591	4.739

finalize a subset of features which appears useful in the classification process. The feature selection phase eases the model interpretability, lessen the model complexity, augment the computational efficiency and consequently elude overfitting (Costa-Mendes et al. 2020). Various feature selection algorithms are available for this purpose. To decrease the number of overlapping features, we utilize Gain Ratio Attribute Evaluator Filter with Ranker search methods. Gain Ratio is a type of feature selection algorithm based on the principle of information gain. It produces Gain Ratio (GR) values for the features and a high value indicates the importance of the feature for classification. Table 3 provides the features listed in descending order of their GR values.

We chose the 4 most significant features; CGPA, PreReq_Grades, Grade-1_Cont (exam-1 marks) and attendance. Table 4 demonstrates the descriptive analysis of the chosen features. The data is well distributed within the range of each feature. Since passing the pre-requisite subject is mandatory, therefore, the minimum is 50. The attendance is well distributed between 77 to 100%.

Experimental evaluation

Model development

We make use of Waikato Environment for Knowledge Analysis (WEKA) to perform the classification experiments (Hall et al., 2009). Developed at the University of Waikato, New Zealand, WEKA is an open source software consisting of a wide range of algorithms for data pre-processing, classification, clustering, regression, and association rules.

We selected four widely used machine learning algorithms. From lazy algorithms, we chose k-Nearest Neighbours (k-NN) implemented as IBk in WEKA. RepTree is an

Khan *et al. Smart Learn. Environ.* (2021) 8:17

Page 9 of 18

Table 5 A confusion matrix for binary classification model

		Predicted results	
		High	Low
Actual values	High	True positive (TP)	False negative (FN)
	Low	False positive (FP)	True negative (TN)

Table 6 Confusion matrixes for prediction models

High	Low	Classified as	High	Low	Classified as
k-NN			*Decision tree*		
103	12	High	102	13	High
14	22	Low	8	28	Low
Artificial neural networks (ANN)			*Naïve Bayes*		
103	12	High	99	16	High
11	25	Low	8	28	Low

implementation of decision tree in WEKA. Similarly, we chose Multilayer Perceptron (MLP) which is a class of Artificial Neural Networks (ANN) and the fourth algorithm is Naïve Bayes. We employed tenfold cross-validation (Hastie et al., 2005) in which the training dataset is split into 10 identical length intervals. In each cycle, the nine intervals are used for learning purpose and the tenth for testing the algorithm's performance. It is an iterative process and, in each iteration, a new interval is chosen for testing. Confusion Matrix (Tharwat, 2018), visualizes a classification model. Table 5 illustrates a standard confusion matrix for binary classification model. Table 6 provides the confusion matrixes for the produced prediction models.

Model evaluation

The capability to classify students into their correct classes demonstrates the dominance of the model. To measure the performance of algorithms, we use accuracy, precision, recall, F-Measure and Mathew Correlation Coefficient (MCC). The Classification accuracy evaluates performance of the prediction models on the whole. It gives an idea about how effectively the model correctly identified the True Positive (TP) and True Negative (TN) instances. It is calculated as:

$$Accuracy = \frac{(TP + TN)}{(TP + FN + FP + TN)} \tag{1}$$

Figure 2 compares the accuracies of the produced models. It illustrates decision tree is achieving an accuracy of over 85% while the remaining stays slightly below. This demonstrates that the prediction models are showing an excellent understanding of the training dataset. However, since accuracy single-handedly does not guarantee the superiority of prediction model, especially when the dataset consists of uneven ratio of instances in the prediction classes. Therefore, we compare the models through other metrics as well.

Fig. 2 Accuracy comparison of prediction models

Fig. 3 Comparing the precision and recall of the prediction models
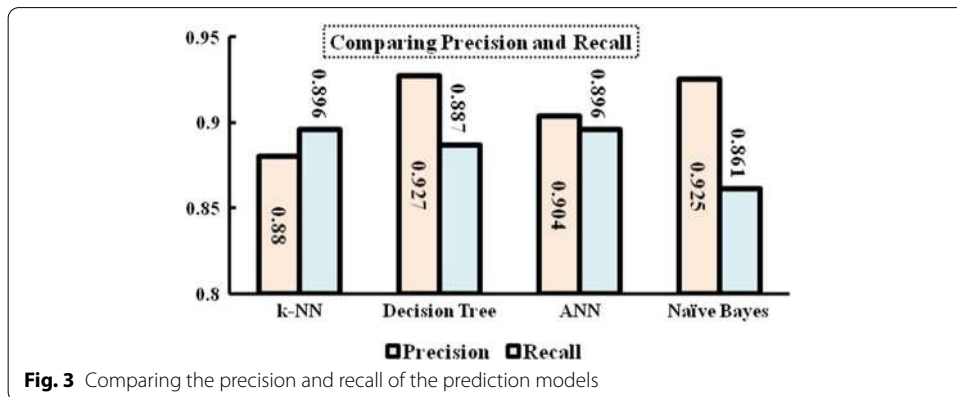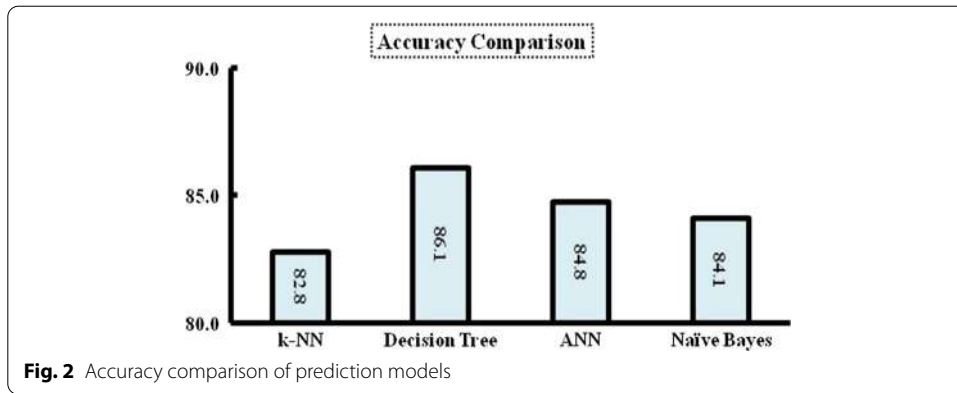
The recall refers to completeness of the model while precision refers to the exactness of the model. Figure 3 compares the precision and recall values of the prediction models. They are calculated as follow

$$Recall = \frac{TP}{(TP + FN)} \tag{2}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{3}$$

Decision tree and Naïve Bayes are achieving the highest precision followed by ANN. However, both k-NN and ANN prevail over in the recall comparison. It shows that decision tree have relative lower recall value, although, it appeared with highest accuracy.

F-Measure is calculated as the harmonic average of precision and recall and thus encompasses the algorithms performance in a single value. It is calculated as:

$$F - Measure = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \tag{4}$$

Fig. 4 Comparing F-Measure value of the prediction models

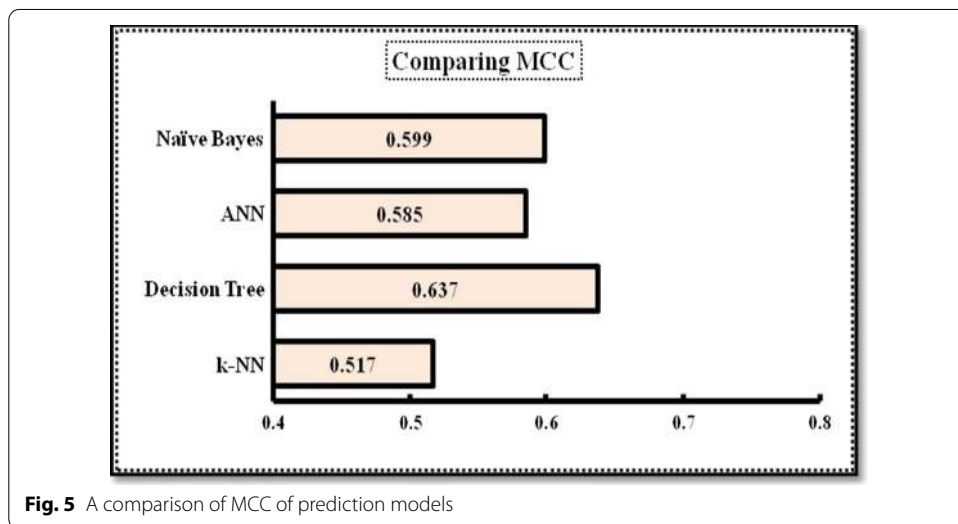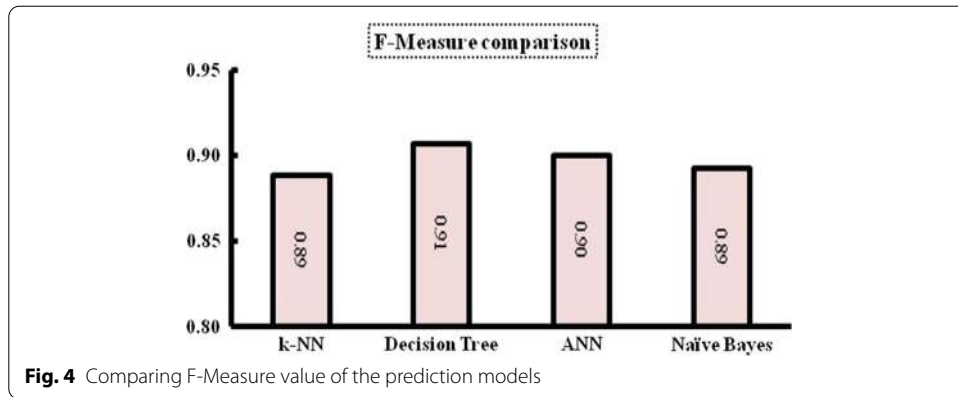**Fig. 5** A comparison of MCC of prediction models

Figure 4 compares the F-Measure values of the developed models. It shows that Decision tree is achieving the highest F-Measure values followed by ANN. K-NN, despite having better recall than Naïve Bayes achieves similar F-Measure value.

Mathew Correlation Coefficient (MCC) (Matthews, 1975) is a reliable statistical rate which evaluate the performance of the classifier in terms of how well it classified the instances in correct classes. It returns a value between -1 (total disagreement between the prediction and observation) and $+1$ (a perfect prediction). It is calculated as follow.

$$MCC = \frac{TP.TN - FP.FN}{\sqrt{(TN + FN).(TP + FP).(TN + FP).(TP + FN)}} \tag{5}$$

To check the ability of algorithms to classify instances in their exact classes, we compare their MCC values in Fig. 5. It illustrates decision tree leaving behind the entire set of models. This concludes that decision tree has high capability to place accurately the instances in their respective classes.
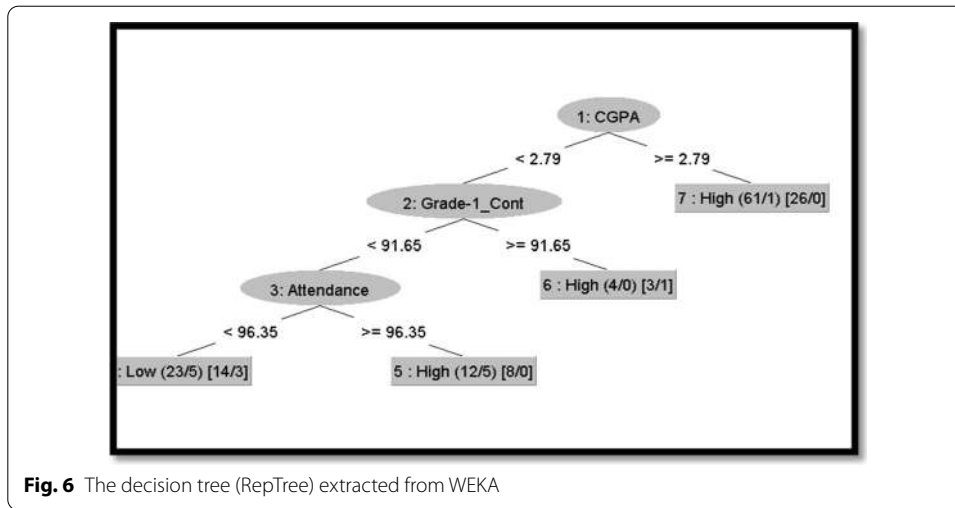
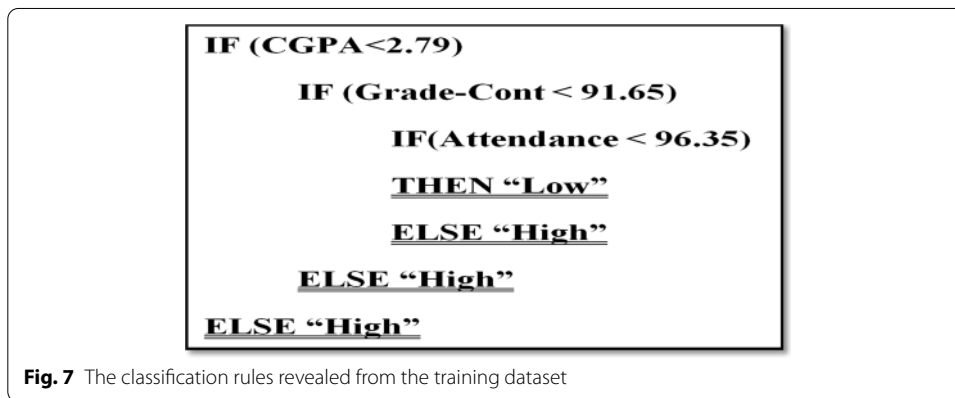**Fig. 6** The decision tree (RepTree) extracted from WEKA

Fig. 7 The classification rules revealed from the training dataset

Model selection

The main purpose of this research is to track the students' academic performance, identify the students with low academic capabilities at precise time and propose precautionary measures. The models evaluation concludes decision tree model leading in almost all the evaluation metrics. Achieving high accuracy illustrates its capability to correctly classify TP and TN instances. A higher precision illustrates that around 93% of its prediction will be correct and high recall indicates the predictions will be identical with the actual outcomes. Similarly, the high F-Measure increases its capability to produce results with high harmonic average of both precision and recall.

A higher MCC shows that decision tree is capable to label the instances correctly within the classes. Even though, the evaluation shows a minor difference in metrics, but the higher MCC value corroborate decision tree preeminence. Moreover, decision tree classifiers provide clear illustration that are easily understandable even by ordinary users (Trabelsi et al., 2019). Therefore, we propose to choose decision tree based model for further evaluation.

Model implementation

Model interpretation

The previous section concludes decision tree as an appropriate model in the current context. Figures 6 and 7 shows the decision tree and the classification rules respectively as extracted from WEKA. The key advantage of decision tree is its ease in understanding and interpretation. The decision tree splits the CGPA, exam-1 marks and attendance. The rules from the decision tree further clarify the model. It shows that CGPA is the prime feature of students' performance followed by the exam-1 marks. It shows that the students having CGPA of 2.79 and above have probability to obtain higher grades. On the other hand, the students having CGPA less than 2.79 but obtaining 91.65% or above grades in the exam-1 also tend to have higher grades. Students with CGPA less than 2.79, with less than 91.65% grades in the exam-1 and attending less than 96.35% classes are in the real danger of producing a lower outcome in the course. The decision tree built tree based up 3 features, bypassing PreReq_Grades.

Model execution

The prediction model is imperfect with no precautionary measures. Therefore, the final phase of this research is to implement the model in field. The goal is to transform the model in easy to understand procedure and design preventive actions for the students with inadequate performance.

Supportive measures (pre-execution procedures)

The output model reveals attendance as one of the key factor and thus the instructor should constantly emphasis over its significance from the start of the course/semester. In order to lessen the consequences of absence the instructor can revise the teaching methodology by adding time slot for quick revision of the previous class. Further, the model indicates the instructor should give extra attention to the student having CGPA below 2.79. These measures can reduce the number of inefficient students prior to the exam-1.

Preventive measures (post-execution procedures)

The instructor executes the model immediately after the exam-1 and gets the list of students classified with their probable final outcome. The students in danger of failing the course will be forwarded for advisory consultations. The instructor may arrange interview sessions with each student and plan precautionary actions compliant with the adapted perspective of every student. Additionally, the instructor may arrange additional classes to bring the at-risk students on right track. The main purpose of additional classes is to revise the course contents and motivate the students.

Model execution and results

At the end of exam-1, the course instructor prepares a prediction dataset. This dataset comprises the students' CGPA, attendance and grades in exam-1. Unlike the training dataset, the last column of the prediction dataset (Final_Grade) is marked with a "?" sign and is supposed to be filled by the prediction model. The instructor provides the

prediction dataset as input to the developed model. The model executes, and predicts the final outcome of each student in the prediction dataset. The output file contains the predicted final outcome of each student as either "Low" or "High". The students having "Low" predicted outcome have higher risk of producing unsatisfactory final result. Therefore, the instructor enlists these students, check their sponsorship status and send them warning along with their shortcomings.

The instructor can forward the students to advisory committee if he/she has severe issues with her progress. This is rigorous for the students with government sponsorship, as the government may cease the sponsorship provided that the students fail to maintain CGPA above 2.0 at the end of academic year. In certain circumstances, the instructor can provide advices or arrange additional class.

The model is tested in real environment for one semester over a class with 25 students. Once exam-1 ended, the model was executed which identified 5 female students in danger of ending the semester with unsatisfactory results. Table 7 shows the list of students and their essential information.

Student-1 has disappointing academic standing as demonstrated by her CGPA, attendance and poor grades in the pre-requisite course. The student is forwarded to the advisory interview where they will investigate reasons for her low CGPA. The instructor provides advises to improve the attendance and offer additional class to improve her academic position. Similarly, student-4 needs an advisory meeting to avoid falling down below CGPA of 2.0 and qualify for extra classes and attendance advices from the instructor. Both students 1 and 4 have government sponsorship so need additional advices.

Student-3 has a poor attendance which probably leads her to poor grades in the first exam despite a healthy CGPA and satisfactory performance in pre-requisite course. The student qualifies for extra classes to improve her position as well as needs instructor advice to sort out the issues restricting her to attend the classes. Student-2 is having poor grades in pre-requisite course and in the first exam qualifies her for additional classes and needs attendance advices from instructor. The only issue with student-5 is her low attendance, which is handled with instructor advices and warning.

Overall, 2 students are forwarded for advisory committee interviews, 4 qualify for additional classes and one needs sever attendance warning and advices. The additional classes, of smaller length, were arranged to provide additional support to the students. Similarly, the students were advised to visit instructor during the office hour of the instructor to gain additional knowledge.

Overall, the students' showed average interest in additional classes or visits to the instructor office. One of the major reasons could be the timings clashes between

Table 7 The list of students with unsatisfactory academic standings

S. no.	Attendance	Grade-1_Cont	CGPA	PreReq_Grades	Sponsorship
1	90.62	76.7	1.96	54	Y
2	89.58	63.3	2.45	50	N
3	79.17	66.7	2.36	75	N
4	89.58	70	2.01	78	Y
5	89.58	83.3	2.16	60	N

student's free time and instructor's office hour. However, the improvement in their attendance enhanced their learning skills and by the end of the course, the students ended up with satisfactory results.

The prediction model appeared constructive and helpful for both the instructor and the students. The students at the risk of producing unsatisfactory results were identified and alerted at the 6th week of the semester. Student at the edge of low CGPA praised the model and an in-time warning motivated them to work hard and come out of the struggling status. In the same way, the students with satisfactory outcome also adopted the attendance advice and eventually, at the end of the semester, negligible number of students appeared with low class attendance. The overall effects are measured with the improvement of the struggling students. Most of the students (4 out 5) were able to produce an acceptable final result.

Conclusion and future work

The instructors are keener to monitor of students' academic growth and provide additional support to the students with inadequate academic progress. The instructor can provide additional support to the struggling students. This research assesses the usefulness of machine learning algorithms to monitor students and notify instructor about the students who are predicted with inadequate result. The prime aim is to identify the struggling students at an early stage of the semester so they will have enough time to rework and attain a satisfactory final result. We pre-processed a dataset having 151 instances and applied a set of machine learning algorithms, explicitly k-NN, decision tree, artificial neural networks, and naïve bayes, to come up with most appropriate prediction model. Decision tree prevails by achieving an accuracy of over 86%, F-Measure of 0.91 and MCC of 0.63. The chosen model is then transformed into a shape easily comprehensible so the instructor can easily view the findings and prepare necessary precautionary procedures.

The interpretation reveals CGPA, grades in exam-1, grades in pre-requisite course and the class attendance as features which quantifies a student's academic position. The instructor must present additional care to the students with lower CGPA and encourage the students about significance of attending the class. The instructor executes the model after the exam-1 and instigates preventive measures to offer additional attention to the struggling students in the form of advisory meetings, arranging additional classes and precautionary actions in compliance with personalized circumstances of the individual struggling student. The field test demonstrates the efficiency of the model and several students are identified with probable unsatisfactory final results. The instructor devised additional procedures to provide personalized support to each student.

In future, we would like to extend the notion and apply the model again after the 2nd exam. This will increase the efficiency of the model and struggling students will get an additional opportunity to rework and prepare well for the forthcoming assessments. Since, the additional classes or visits to instructor's office did not work well; therefore, we plan to append an additional recommendation module to the proposed framework. The recommendation module will automatically send personalized recommendations to

the students in accordance to their current status. We aim to apply the model in other courses and extend the notion to entire set of courses in the institution.

Abbreviations
OAAA: Oman Academic Accreditation Authority; ABET: Accreditation Board for Engineering and Technology; ITS: Intelligent Tutoring Systems; LMS: Learning Management Systems; AI: Artificial intelligence; ANN: Artificial neural networks; k-NN: K-nearest neighbors; WEKA: Waikato Environment for Knowledge Analysis; MLP: Multilayer perceptron; MCC: Mathew correlation coefficient; TP: True positive; FN: False negative; FP: False positive; TN: True negative.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40561-021-00161-y.

> **Additional file 1.** Original dataset file.
>
> **Additional file 2.** Online Resource 2_DataSet_Preprocessed.

Authors' contributions
Each author have made considerable contributions to the design of the work, preparation of data, the experimentations (their validation), analysis (revision and validation of analysis), Each author have drafted the work or revised it for possible improvements. All authors read and approved the final manuscript.

Authors' Information
Ijaz Khan received his bachelors (Hons) degree in Computer Science from University of Peshawar, Pakistan in 2003 and MPhil in Computer Science from Liverpool John Moores University in 2009. He has been working as a lecturer at Al-Buraimi University College, Oman for the last 8 years. Currently, he is studying at the University Tenaga Nasional (UNITEN) Malaysia for his Ph.D., focusing on using machine learning algorithm to develop students' performance prediction models. His research interests include Artificial Intelligence, Machine learning, internet of things and social networking.

Abdul Rahim Ahmad early career was as a lecturer in the Malaysian Polytechnic system. He later joined the satellite television company Measat Broadcast Network Systems (MBNS) broadcast automation team. In 1997, he joined Universiti Tenaga Nasional (UNITEN) as Senior Lecturer. He held various positions in the College of Computing and Informatics (CCI), Research Management Centre (RMC) and the Institute of Energy Policy and Research (IEPRe). He is currently an Associate Professor in CCI. He obtained his bachelor degree in Computer Science from University of Queensland in Australia, masters from Loughborough University, UK and Phds from Universiti Tteknologi Malaysia and University of Nantes, France. His academic and research are in Computer Systems, Networks and Artificial Intelligence. He was active as a committee member of IEEE Malaysia and IEEE Computer Society Malaysia.

Dr. Nafaâ Jabeur is Associate Professorand Director of Research at the GermanUniversity of Technology in Oman (GUtech). He received his PhD and MSc degrees in Computer Science from Laval University, Quebec, Canada in 2006 and 2001, respectively. He has over than fifteen years of experience in the industrial and academic sectors. Nafaâ has participated in several R&D projects,edited2 books, and authored more than 80 research papers in prestigious conferences and high ranked journals.His main research interests include Smart cities, Transportation, Blockchain, Drones, and Artificial Intelligence.

Mohammed Najah Mahdi received his B.Sc. degree in Information Engineering, College of Engineering, Baghdad University, Iraq, in 2002 and M.Sc. in Information Technology from Faculty of Computer Science and Information Technology, University of Malaya (UM) in 2011. He later obtained his Ph.D. in Information and Communication Technology, in 2017. He is currently a Post-Doctoral Researcher at the University Tenaga Nasional (UNITEN) Malaysia. His research interests include faceted search, information overload, exploratory search and information retrieval.

Availability of data and materials
All data generated or analyzed during this study are included in the Additional files 1 and 2. The datasets used during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]College of Graduate Studies, Universiti Tenaga Nasional, Kajang, Malaysia. [2]Information Technology Department, Buraimi University College, Al-Buraimi, Oman. [3]College of Computing and Informatics, Universiti Tenaga Nasional, Kajang,

Malaysia. [4]Computer Science Department, German University of Technology, Muscat, Oman. [5]Institute of Informatics and Computing in Energy, Universiti Tenaga Nasional, Kajang, Malaysia.

References

Alabri, A., Al-Khanjari, Z., Jamoussi, Y., & Kraiem, N. (2019). Mining the students' chat conversations in a personalized e-learning environment. *International Journal of Emerging Technologies in Learning (iJET), 14*(23), 98–124.

Alfere, S. S., & Maghari, A. Y. (2018). Prediction of student's performance using modified KNN classifiers. *Prediction of Student's Performance Using Modified KNN Classifiers.*

Asogwa, O., & Oladugba, A. (2015). Of students academic performance rates using artificial neural networks (ANNs). *American Journal of Applied Mathematics and Statistics, 3*(4), 151–155.

Chen, H. (2018). *Predicting student performance using data from an Auto-grading system.* University of Waterloo.

Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior, 73*, 247–256.

Costa-Mendes, R., Oliveira, T., Castelli, M., & Cruz-Jesus, F. (2020). A machine learning approximation of the 2015 Portuguese high school student grades: A hybrid approach. *Education and Information Technologies*, 1–21.

Cunningham, P., & Delany, S. J. (2007). k-Nearest neighbour classifiers. *Multiple Classifier Systems, 34*(8), 1–17.

Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning, 29*(2–3), 103–130.

Došilović, F. K., Brčić, M., & Hlupić, N. (2018). Explainable artificial intelligence: A survey. *Paper presented at the 2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO).*

Drachsler, H., & Greller, W. (2012). Confidence in learning analytics.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10–18.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: Data mining, inference and prediction. *The Mathematical Intelligencer, 27*(2), 83–85.

Iatrellis, O., Savvas, I. K., Fitsilis, P., & Gerogiannis, V. C. (2020). A two-phase machine learning approach for predicting student outcomes. *Education and Information Technologies*, 1–20.

Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics, 2*(1), 1.

Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1): Sthda.

Kaunang, F. J., & Rotikan, R. (2018). Students' Academic Performance Prediction using Data Mining. In *Paper presented at the 2018 third international conference on informatics and computing (ICIC).*

Kausar, S., Oyelere, S., Salal, Y., Hussain, S., Cifci, M., Hilcenko, S., Iqbal, M., Wenhao, Z., & Huahu, X. (2020). Mining smart learning analytics data using ensemble classifiers. *International Journal of Emerging Technologies in Learning (iJET), 15*(12), 81–102.

Khan, I., Al Sadiri, A., Ahmad, A. R., & Jabeur, N. (2019). Tracking student performance in introductory programming by means of machine learning. In *Paper presented at the 2019 4th MEC international conference on big data and smart city (ICBDSC).*

Khan, S. (2018). Modern Internet of Things as a challenge for higher education. *IJCSNS, 18*(12), 34.

Kiu, C.-C. (2018). Data mining analysis on student's academic performance through exploration of student's background and social activities. In *Paper presented at the 2018 fourth international conference on advances in computing, communication & automation (ICACCA).*

Lagman, A. C., Calleja, J. Q., Fernando, C. G., Gonzales, J. G., Legaspi, J. B., Ortega, J. H. J. C., Ramos, R. F., Solomo, M. V. S., & Santos, R. C. (2019). Embedding naïve Bayes algorithm data model in predicting student graduation. In *Paper presented at the proceedings of the 3rd international conference on telecommunications and communication engineering.*

Lesinski, G., Corns, S., & Dagli, C. (2016). Application of an artificial neural network to predict graduation success at the United States Military Academy. *Procedia Computer Science, 95*, 375–382.

Li, M., Xu, H., & Deng, Y. (2019). Evidential decision tree based on belief entropy. *Entropy, 21*(9), 897.

Liao, S. N., Zingaro, D., Thai, K., Alvarado, C., Griswold, W. G., & Porter, L. (2019). A robust machine learning technique to predict low-performing students. *ACM Transactions on Computing Education (TOCE), 19*(3), 1–19.

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. *Journal of Management Analytics, 6*(1), 1–29.

Ma, X., Yang, Y., & Zhou, Z. (2018). Using machine learning algorithm to predict student pass rates in online education. In *Paper presented at the proceedings of the 3rd international conference on multimedia systems and signal processing.*

Maesya, A., & Hendiyanti, T. (2019). Forecasting student graduation with classification and regression tree (CART) algorithm. In *Paper presented at the IOP conference series: Materials science and engineering.*

Márquez-Vera, C., Morales, C. R., & Soto, S. V. (2013). Predicting school failure and dropout by using data mining techniques. *IEEE Revista Iberoamericana De Tecnologias Del Aprendizaje, 8*(1), 7–14.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure, 405*(2), 442–451.

Mitchell, R., Michalski, J., & Carbonell, T. (2013). *An artificial intelligence approach*: Springer.

Mondal, A., & Mukherjee, J. (2018). An Approach to predict a student's academic performance using Recurrent Neural Network (RNN). *International Journal of Computers and Applications, 181*(6), 1–5.

Nettleman, C. A., III. (2018). An assessment of ABET-accredited undergraduate land surveying and geomatics programs in the United States. *Surveying and Land Information Science, 77*(2), 105–114.

Nilsson, N. J. (2014). *Principles of artificial intelligence*: Morgan Kaufmann.

Norris, D., Baer, L., Leonard, J., Pugliese, L., & Lefrere, P. (2008). Action analytics: Measuring and improving performance that matters in higher education. *EDUCAUSE Review, 43*(1), 42.

Oladokun, V., Adebanjo, A., & Charles-Owaba, O. (2008). Predicting students academic performance using artificial neural network: A case study of an engineering course.

Orong, M. Y., Caroro, R. A., Durias, G. D., Cabrera, J. A., Lonzon, H., & Ricalde, G. T. (2020). A predictive analytics approach in determining the predictors of student attrition in the higher education institutions in the Philippines. In *Paper presented at the proceedings of the 3rd international conference on software engineering and information management*.

Pandey, M., & Taruna, S. (2016). Towards the integration of multiple classifier pertaining to the Student's performance prediction. *Perspectives in Science, 8*, 364–366.

Qazdar, A., Er-Raha, B., Cherkaoui, C., & Mammass, D. (2019). A machine learning algorithm framework for predicting students performance: A case study of baccalaureate students in Morocco. *Education and Information Technologies, 24*(6), 3577–3589.

Romanenko, V., Tropin, Y., Boychenko, N., & Goloha, V. (2019). Monitoring student performance using computer technology. *Slobozhanskyi herald of science and sport, 7*(2(70)), 36–39.

Saa, A. A. (2016). Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications, 7*(5), 212–220.

Sunday, K., Ocheja, P., Hussain, S., Oyelere, S., Samson, B., & Agbo, F. (2020). Analyzing student performance in programming education using classification techniques. *International Journal of Emerging Technologies in Learning (iJET), 15*(2), 127–144.

Suthaharan, S. (2016). Support vector machine *Machine learning models and algorithms for big data classification* (pp. 207–235): Springer.

Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*.

Trabelsi, A., Elouedi, Z., & Lefevre, E. (2019). Decision tree classifiers for evidential attribute values and class labels. *Fuzzy Sets and Systems, 366*, 46–62.

Umer, R., Susnjak, T., Mathrani, A., & Suriadi, S. (2017). On predicting academic performance with process mining in learning analytics. *Journal of Research in Innovative Teaching & Learning*.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann.

Yang, T.-Y., Brinton, C. G., Joe-Wong, C., & Chiang, M. (2017). Behavior-based grade prediction for MOOCs via time series neural networks. *IEEE Journal of Selected Topics in Signal Processing, 11*(5), 716–728.

Yousafzai, B. K., Hayat, M., & Afzal, S. (2020). Application of machine learning and data mining in predicting the performance of intermediate and secondary education level student. *Education and Information Technologies*, 1–21.

Publisher's Note