This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier 10.1109/ACCESS.2017.DO

Early Predicting of Students Performance in Higher Education

ESSA ALHAZMI[1,2], ABDULLAH SHENEAMER[1,3]
[1]Faculty of Computer Science & Information Technology, Jazan University, Jazan 45142 KSA
[2](e-mail: esalhazmi@jazanu.edu.sa)
[3](e-mail: asheneamer@jazanu.edu.sa)
Corresponding author: Abdullah M. Sheneamer(e-mail: asheneamer@jazanu.edu.sa).

This paragraph of the first footnote will contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456."

ABSTRACT Students learning performance is one of the core components for assessing any educational systems. Students performance is very crucial in tackling issues of learning process and one of the important matters to measure learning outcomes. The ability to use data knowledge to improve education systems has led to the development of the field of research known as educational data mining (EDM). EDM is the creation of techniques to investigate data gathered from educational settings, allowing for a more thorough and accurate understanding of students and the improvement of educational outcomes for them. The use of machine learning (ML) technology has increased significantly in recent years. Researchers and teachers can use the measurements of success, failure, dropout, and more provided by the discipline of data mining in education to predict and simulate education processes. Therefore, this work presents an analysis of students performance using data mining methods. The paper presents both clustering and classification techniques to identify the impact of students performance at early stage with on the GPA. For the clustering technique, the paper uses dimensionality reduction mechanism by T-SNE algorithm with various factors at early stage such as admission scores and first level courses, academic achievement tests (AAT) and general aptitude tests (GAT) in order to explore the relationship between these factors and GPA's. For the classification technique, the paper presents experiments on different machine learning models on predicting student performance at early stages using different features including courses' grades and admission tests' scores. We use different assessment metrics to evaluate the quality of the models. The results suggest that educational systems can mitigate the risks of students failures at the early stages.

INDEX TERMS Graph mining; Students' performance prediction; Student academic performance; Early prediction; Data mining

I. INTRODUCTION

Education is an important element and plays a significant role in our society. Information and communication technology has affected many fields of research, specifically in the education field. For example, as we have seen in many countries used various e-Learning environments [1] due to the recent pandemic COVID19.

A higher education institution considers the academic performance of students as one of the most important issues regarding presenting quality education to its students [2], [3].

Understanding the significant factors in students performance at early stage of their education is complex. Various effective tools have been used to overcome the students' performance challenges in academia. However, these tools may not be easy to generalized in all circumstances of education. In the recent years, with the advances of the application of technologies to forecasting studentsâ performance, there are still gaps to be filled in order to analysis and improve the accuracy of students performance using new features and data mining methods and present both clustering and classification tech-

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

niques to identify the impact of students performance at early stage with on the GPA.

The learning process includes a lot of student performance. Identifying students who are more likely to have poor academic success in the future requires making predictions about student performance. If the data has been transformed into knowledge, it may be useful and used in predictions. As a result, the information might improve the quality of education and learning and help students in achieving their academic objectives. Data mining techniques are used in the study area known as educational data mining (EDM) to analyze information derived from educational backgrounds [4]. EDM implementation also aids in the planning of strategies for raising student performance. As a result, it will improve teaching and learning and the students' experience in the educational institution.

Academic success is important because it is strongly linked to the positive outcomes we value. One of the academic success factors is the academic students performance in the college or university. The cumulative academic achievement for each student still indicates the success of every college or university. Also, the other factors we can use in analysis and predict the academic students performance are aptitude test, GPA of secondary school and the name of the school which the student graduated from. We believe that the performance of the students in the first year in college can be used as a factor to predict the performance of student in the rest of years of his/her studies. These factors lead to early remedy for students and take actions to improve student performance Artificial intelligence techniques have been applied on educational data to reveal the significant reasons behind student performance. The contributions of the paper are as follows:

- We propose a framework for predicting students performance using student's academic performance and his/her social relationships features.
- We use admission scores, his/her first level courses scores and academic achievement test (AAT) and general aptitude test (GAT).
- We explore a new way of using admission, his/her first level courses scores, and AAT and GAT by tSNE dimensionality reduction. To the best of our knowledge this attempt is a first of its kind to use features from both admission scores and his/her first level courses scores to early predict student's performance using machine learning.
- We also explore a new way of using increasing the threshold of relocating which is to compute the absolute difference between a grade and following grade after or before.
- We use a state-of-the-art classification models to evaluate the effectiveness of our proposed idea.

We organize the paper as follows: Section II provides the literature related to students performance prediction techniques used in the field of Education. Section III provides details of the used dataset includes data caracterization and correlations. Section IV provides the research methodology followed by the paper. We evaluate and analysis our proposed method and report findings in Section V. Finally, we conclude our work in Section VII.

II. RELATED WORK

Helal *et al.* [5] proposed the concept of using heterogeneity to create better prediction models. Four popular machine learning algorithmsâJRip, sequential minimum optimization, C4.5, and Naive-Bayesâwere used to create these models. The results of the experiment showed that employing student subpopulations to predict academic performance is both successful and promising. Additionally, it demonstrated that both rule-based and tree-based algorithms offered more clear explanations, making them more effective in an educational environment [5]. This study did not consider the combined features corresponding to a particular module.

Xu *et al.* [6] proposed a technique utilizing three popular machine learning algorithms to predict student performance using Internet usage data. From the real Internet usage data of 4000 students, they collected, computed, and normalized parameters like online time, Internet traffic volume, and connection frequency. Their findings indicated that it is possible to predict students' academic success using data on Internet usage.

Dien *et al.* [7] proposed a method to predict student performance using several deep learning methods. Twenty one features from their model are used as convolutional neural network input. The dataset was obtained from the information system of a multidisciplinary university in Vietnam, and experimental findings on the dataset revealed successful prediction. This work used traditional and simple features to build a studentâs performance prediction model for predicting student performance.

Giannakas *et al.* [1] proposed a binary classification framework based on deep neural networks (DNN), and the most crucial features that affected the result were extracted. Different activation functions (Sigmoid, ReLu, and Tanh) and optimizers were used to evaluate the framework (Adagrad and Adadelta). The experimental findings indicate that, when the Adadelta and Adagrad optimizers were utilized, the prediction accuracy of the framework was 76.73% and 82.39%, respectively. The learning performance was 80.76% and 86.57% overall.

Ha *et al.* [8] used various machine learning algorithms to determine the final grade average of students using a variety of criteria, including personal characteristics, university entrance scores, a gap year, and their first- and second-year academic performance. The dataset that was used was obtained from the university's student management information system and a survey of graduates from three different years. The findings demonstrated a connection between the factors and students' academic achievement.

Rastrollo-Guerrero *et al.* [9] studied and analyzed almost 70 papers to show different modern techniques applied for predicting students' performance. They have studied current research on predicting student behavior in a class environment. They came to the conclusion that there is a substantial tendency to predict university student performance from the analysis of these papers. Due to its ability to deliver accurate and trustworthy outcomes, supervised learning has become a popular strategy for predicting students' behavior. On the other hand, due to the low accuracy of predicting students' conduct in the circumstances analyzed, unsupervised learning is a technique that is unattractive to researchers.

Mustafa Agaoglu [10] explains how classifier models are created using four different classification techniques: decision tree algorithms, support vector machines, artificial neural networks, and discriminant analysis. Using performance criteria for accuracy, precision, recall, and specificity, their results are compared over a given dataset composed of student replies to an actual course evaluation questionnaire.

In order to determine the most crucial factors for ensuring the academic performance of engineering students, Gonzalez-Nucamendi et.al [11] describe the determination of student profiles based on the constructs of multiple intelligences and on learning and affective techniques.

Alshanqiti and Namoun [12] proposed hybrid regression model that optimizes the prediction accuracy of student academic performance, an optimized multi-label classifier that predicts the qualitative values for the influence of various factors associated with the obtained student performance and applied combining three dynamically weighted techniques, namely collaborative filtering, fuzzy set rules, and Lasso linear regression. They need to claim their approach generality to confirm the reliability of their model. Also, they need using real academic datasets.

In order to predict a student's placement in the IT business based on their academic achievement in class ten, class twelve, graduation, and backlog to date in graduation, Maurya et al. [13] have presented a few supervised machine learning classifiers.

Alsalm et al. [14] proposed an approach to predict the studentâs grades using the mathematics and Portuguese language course grades data set and applied Deep learning model. The model of this work is only validated using two datasets and need to be validated on other large size balanced datasets.

Ahmed et al. [15] proposed an approach to predict university students' performance in final exams using the algorithm (GBDT) which is a machine learning technology used for regression, classification, and ranking tasks, and is part of the Boosting method family.

Liu and Niu [16] proposed a new approach which is called the Multi-Agent System (MAS). It is used to propose an Agent-based Modeling Feature Selection(ABMFS) model, and the selected feature subset effectively removes the features that are irrelevant to the prediction results. Then, they applied the Deep Learning techniques to construct a Convolutional Neural Network (CNN) based structure to predict student performance.

Evangelista and Descargar [17] provided a method for improving the performance prediction of several single classification algorithms by employing them as base classifiers of heterogeneous ensembles and homogeneous ensembles (bagging and boosting) (voting and stacking). Their model needs to perform optimization techniques to find out the algorithm parameters and configuration.

Quy et al. [18] assess seven well-liked group fairness metrics for issues predicting student performance. On five educational datasets, they run trials with four traditional machine learning models and two fairness-aware machine learning techniques. Their studay is only limited to public schools not student academics in universities

To predict undergraduate students' final test grades, Yağcı [19] proposed a prediction model. The final exam grades were divided into four groups using a comparison of a variety of machine learning methods, including RF, SVM, LR, NB, ANN, and k-nearest neighbor (KNN). These groups were "32.5," "32.5-55," "55-77.5," and " 77.5." The grades of 1854 students who took Turkish Language-I were taken into account in the proposed model using data that was gathered from a Turkish state university. The department, faculty, and midterm exam scores from the student's tree characteristics have been used to forecast the final exam grades. In comparison to other classification models, RF and ANN performed the best, classifying final exam grades with an area under the curve (AUC) of 86% and 74% accuracy, respectively.

III. DATASET DESCRIPTION

In this study, we use a students records' dataset for the five consecutive years. The dataset provides more than 275 thousands records for almost five thousands students including their identification, instructors, gender, sections, program, admission criteria and time, graduation time and GPA, courses' details and their grades. This dataset is being used in this paper to understand the factors that influence the academic performance of the students.

A. DATA CHARACTERIZATION

From this dataset, we explored various attributes that support our task in quantifying and prediction the students performance in early stage.

Table 1: Dataset Description

Record	Description
Student ID	Student identification number (numeric: unique number)
Sex	student's sex(binary: female or male)
Campus name	building of a college which student is related to (nominal: a college and the related building is situated)
Name of the course	A course title (nominal: course title which a student is taking)
Course level	The number by which a course is designated indicates the level of the course (numeric: identified by three to four digits)
Section number	The course section number (numeric: identified by one to four digits)
Lecturer ID	Lecturer identification number (numeric: unique number)
Semester year	Three semesters per academic year (numeric: Year+number of semester 1,2, 3)
Major	student's major (nominal: CS, IT, CNET)
Grade	Student's grade in a specific course (numeric: 0-100)
Student status	(nominal: regular, graduate, discontinuous, discontinuous
Admission year	The semester that student's admitted to his/her college (numeric: year+number of semester 1,2, 3)
Graduate year	The semester that the student is expected to graduate (numeric: year+number of semester 1,2, 3)
Expected graduate year	student's graduated from his/her college (numeric: year+number of semester 1,2, 3)
Secondary school GPA	Student's secondary school GPA (numeric: 0-100%)
General Aptitude Test (GAT) Scores	This test is equivalent to the Arabic version of GAT and is based on mathematical and verbal skills (numeric: from 0 to 100)
Academic Achievements Test Scores	Analyzing achievement tests applied by the Ministry of Education on students in Math, Science, and Arabic subjects(numeric: from 0 to 100)
School name	Name of secondary school that the students graduated from (nominal: name of the secondary school)
Student GPA	Student's GPA when he/she graduated from the college (numerical: from 1 to 5)
Student semester GPA	Last semester of student's GPA (numerical: from 1 to 5)

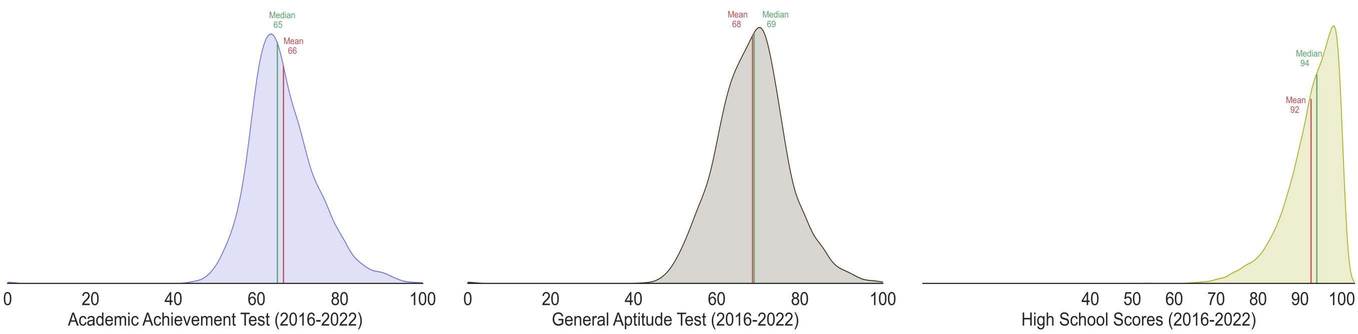

Figure 1: Density distribution functions of the admission criteria including the scores of the academic achievement test, the general aptitude test, and the final grade of high school.

By analyzing admission criteria, Figure 1 shows the density distribution function of the admission criteria. The admission criteria include the academic achievement test, the general aptitude test, and the final grade of high school. These scores are calculated in a weighted ratio to satisfy the admission requirements. The academic achievement test scores are normally distributed showing the mean and median of 66 and 65, consecutively. Similarly, the general aptitude test scores show mean and median of 68 and 69, consecutively. In contrast, the distribution of the final grade of high school scores are left-skewed with mean and median 92 and 94, consecutively.

When comparing the admission criteria with respect to gender during the five years, Figure 2 shows the slight variations between them. It is clearly noticed that the male scores are more stable than female scores in all the admission criteria.

In addition to the admission criteria, figures 3 and 4 describe the most courses are the obstacles for the students. It is clearly noticed that Math courses are the most challenging courses. The English courses are considered the second challenging courses for the students.

IV. PROPOSED METHOD

A. PREPARING THE DATA

In this phase, we use a sample of students' academic records at our CS college to explore the significant factors in students performance. Our target is to present the data of students records and features in early stage to predict their performance in the final stage (final accumulative GPA).

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

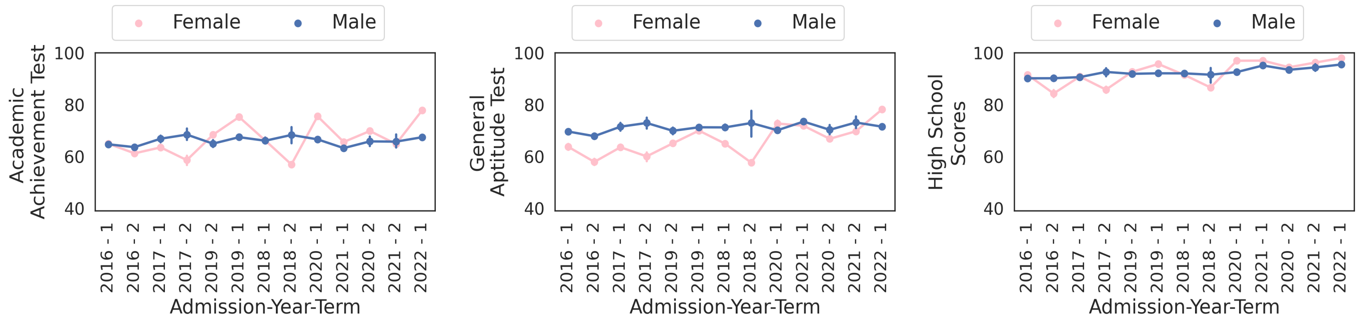

Figure 2: Gender comparison in admission criteria over admission years and terms including the scores of the academic achievement test, the general aptitude test, and the final grade of high school.

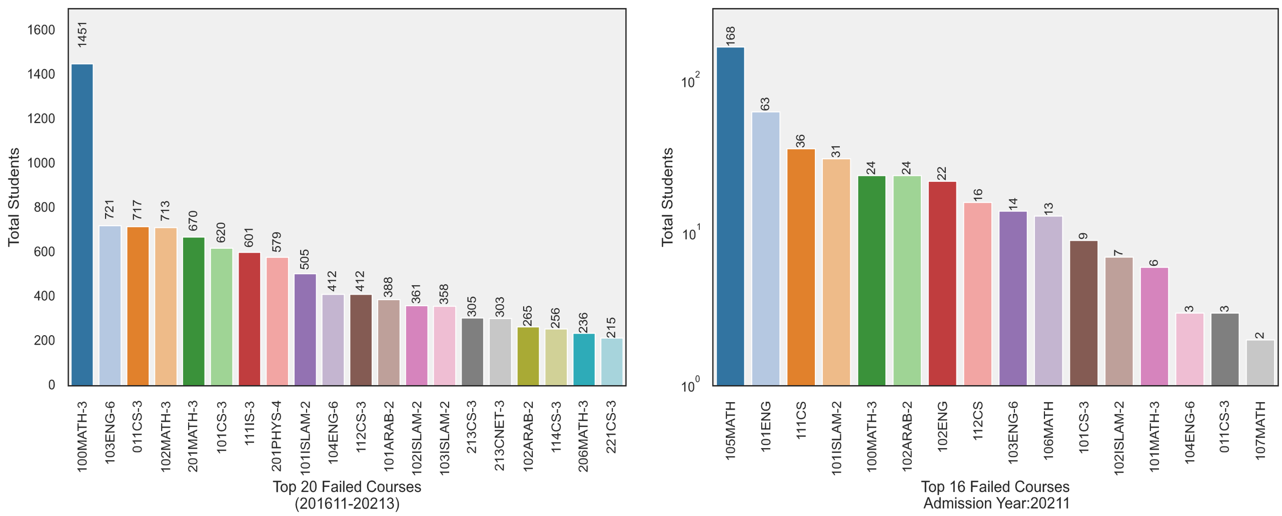

Figure 3: Total failed students in courses (left:during the five years, right: in 2021)

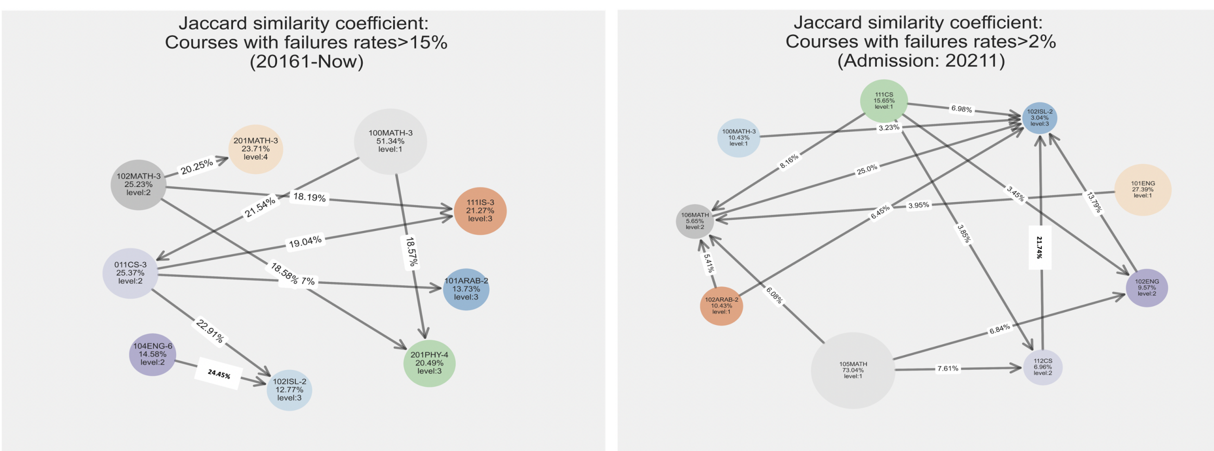

Figure 4: Correlation networks between the failed courses (edge direction denotes next level course) calculated by the Jaccard similarity index where similar students who failed in courses are the set.

B. FEATURES USED

We use admission scores of the students, gender and all first-level required courses scores as features. We also mentioned all features we used in Table 1. We find these features influence the students' academic success. We attempt to predict and improve student performance by utilization of using these mentioned features.

C. T-SNE ALGORITHM

t-SNE is a non-linear dimensionality reduction used for exploring high-dimensional data. It stands for t-distributed Stochastic Neighbor Embedding [20]. We use t-SNE algorithm with various factors such as GAT and AAT for analyzing and exploring the relationship between these factors and GPA of students.

D. MACHINE LEARNING ALGORITHMS

We train and test our classification models using five machine learning algorithms. These include a recently published classification algorithm such as Xgboost [21], and others common such as Logistic Recognition [22], Support Vector Machine (SVM) [23], K nearst neighbor (KNN) [24] and Random Forest (RF) [25]. Xgboost is short for (eXtreme Gradient Boosting). Gradient boost defines an objective function that contains two parts: training loss and regularization [21]. We compare these classification algorithms in their ability to detect performance students show that supervised machine learning methods using our novel features perform much better than using just traditional features.

V. EXPERIMENTS

A. ASSESSMENT DESIGN

We perform experiments with varying numbers of features (admission scores, admission scores and gender, and admission scores with all first-level courses scores) features. The intention behind such experiments is to show the significance of our proposed features in achieving better accuracy, and that it is not by chance. This further establishes the fact that our features are crucial in deriving high accuracy detection results.

We build our model for predicting the student performance. First, we train the model with appropriate samples. To prepare training samples we extract student records from the dataset. We extract admission scores of the students, gender and all first-level required courses scores features as discussed earlier. We follow the same steps for prediction of unknown samples.

B. EXPERIMENTING WITH ADMISSION SCORES FEATURES

In this experiment, we used only examples of admission scores for training and testing. Our primary goal is to show student performance accuracy. We used six classifiers such as Logistic Recognition, Random forest, KNN, SVM, GNB,and XGB. However, we showed the best performed classifiers such as Random forest, SVM, and GNB. In Figure 6 and Table 2, the accuracy of admission score features results beat the model of admission scores and gender features combination.

C. EXPERIMENTING WITH ADMISSION SCORES AND GENDER FEATURES

In this experiment, we combined two different types of feature which are admission scores and gender features to observe if the gender features impact and produce better accuracy performance using only admission scores as features. Unfortunately, we noticed that the accuracy performance is reduced in all used classifiers as shown in Figure 7 and Table 3.

D. EXPERIMENTING WITH ADMISSION SCORES AND ALL FIRST-LEVEL ENGLISH COURSES SCORES FEATURES

We also conducted an experiment to explore the accuracy performance of a combination of admission scores and all first-level English courses features model as shown in Figure 8 and Table 4. Its results are better than the results of model using only admission scores or combination of admission scores and gender features. We conclude that a model uses first-level English courses scores or first-level courses scores in general is better than a model uses student's characteristics information.

E. EXPERIMENTING WITH ADMISSION SCORES AND ALL FIRST-LEVEL MATH SCORES FEATURES

We also conducted an experiment for a model uses admission scores and all first-level Math courses features as shown in Figure 9 and Table 5. We noticed its accuracy performance beats all other combination features models. However, it is defeated by the model that uses combination of admission scores and all first-level computer science scores features and also combination of all various features.

F. EXPERIMENTING WITH ADMISSION SCORES AND ALL FIRST-LEVEL CS SCORES FEATURES

We also conducted an experiment for a model uses admission scores and all first-level CS courses features as shown in Figure 10 and Table 6. All accuracy performance results beat all other combination features models. However, it is defeated by the model that uses combination of all various features.

G. EXPERIMENTING WITH ADMISSION SCORES FEATURES AND ALL FIRST-LEVEL COURSES SCORES FEATURES

Here we combine two different types of feature which are admission scores and all first-level courses scores features to assess the importance and report the results on the dataset in Figure 11 and Table 7. We observe that a admission scores and all first-level courses scores features combination beats all other combinations of features

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

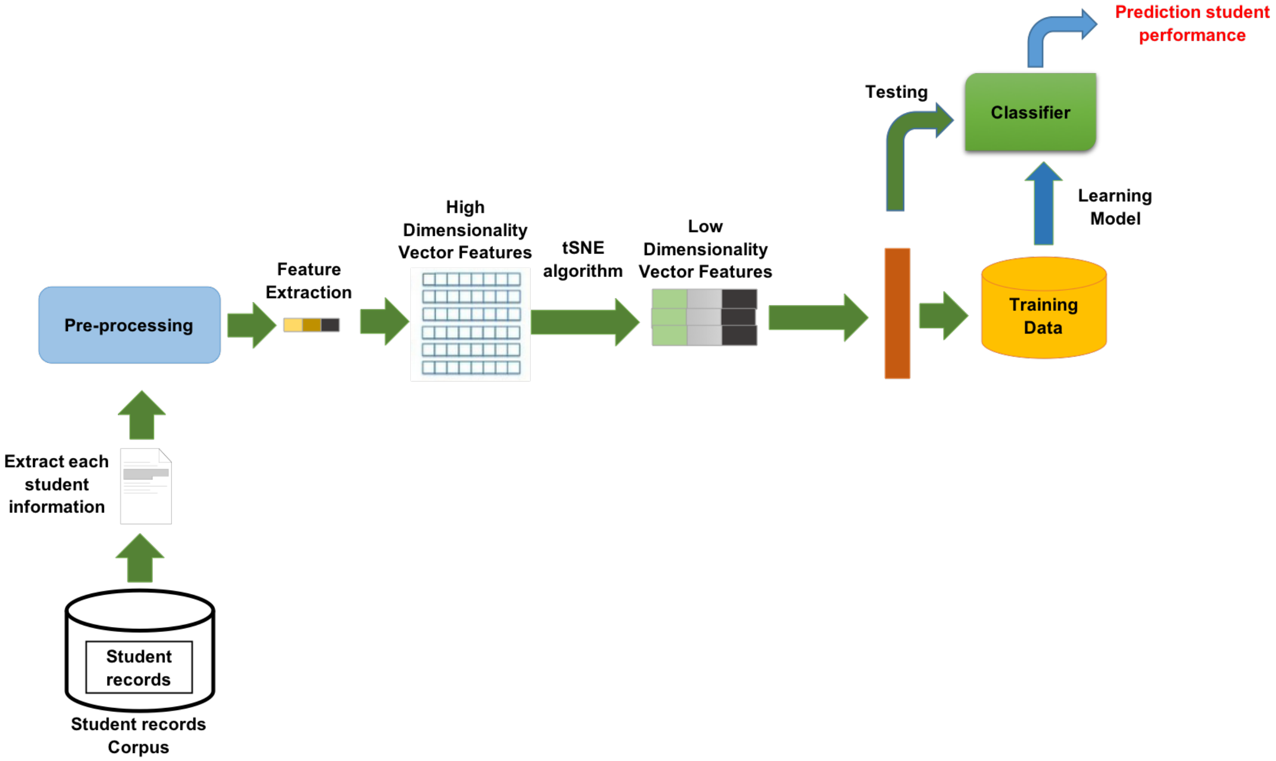

Figure 5: Workflow of the proposed student performance prediction framework.

types. It produces better results than admission scores and gender features combination for all sizes of data. A admission scores and all first-level courses scores features combination works better in comparison to a admission scores and gender features combination, because gender features only have two values either male or female. It effects on the accuracy performance of the classifiers.

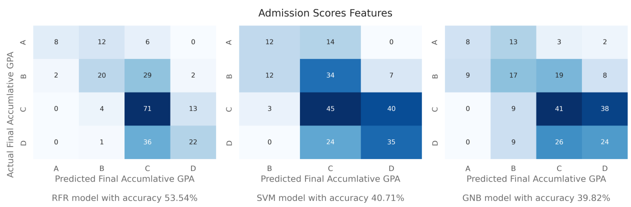

Figure 6: Admission scores features.

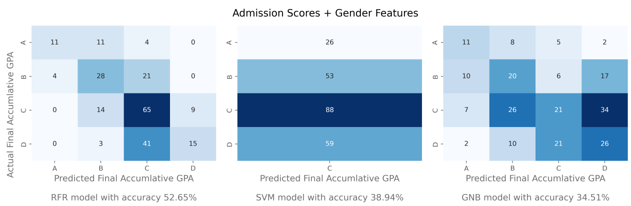

Figure 7: Admission scores and gender features.

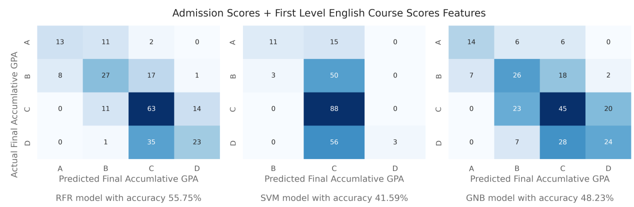

Figure 8: Admission scores and first level English courses scores features.

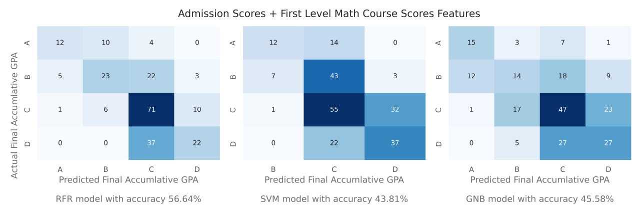

Figure 9: Admission scores and first level Math courses scores features.

H. EXPERIMENTING WITH DIMESIONALITY REDUCTION BY T-SNE

Because it is difficult to visualize data with more than two Dimensions, an essential task that involves the analysis and interpretation of high-dimensional data sets. In order to improve visualization, the t-SNE algorithm [20] is used for dimensionality reduction. This algorithm

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

Table 2: Results of models performance on **admission scores features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score
RFR	A	80.00%	30.77%	44.44%	81.82%	34.62%	48.65%	90.91%	38.46%	54.05%	93.75%	57.69%	71.43%
	B	54.05%	37.74%	44.44%	55.56%	37.74%	44.94%	64.29%	50.94%	56.84%	76.32%	54.72%	63.74%
	C	50.00%	80.68%	61.74%	50.35%	81.82%	62.34%	55.88%	86.36%	67.86%	63.57%	93.18%	75.58%
	D	59.46%	37.29%	45.83%	61.11%	37.29%	46.32%	72.97%	45.76%	56.25%	88.37%	64.41%	74.51%
	Accuracy	53.54%			54.42%			61.95%			72.57%		
SVM	A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	11.54%	20.69%	100.00%	34.62%	51.43%
	B	44.44%	22.64%	30.00%	48.28%	26.42%	34.15%	62.07%	33.96%	43.90%	86.21%	47.17%	60.98%
	C	38.46%	51.14%	43.90%	39.13%	51.14%	44.33%	41.07%	52.27%	46.00%	47.62%	56.82%	51.81%
	D	42.68%	59.32%	49.65%	42.68%	59.32%	49.65%	42.68%	59.32%	49.65%	46.99%	66.10%	54.93%
	Accuracy	40.71%			41.59%			45.13%			54.42%		
GNB	A	47.06%	30.77%	37.21%	47.06%	30.77%	37.21%	56.25%	34.62%	42.86%	81.25%	50.00%	61.90%
	B	35.42%	32.08%	33.66%	35.42%	32.08%	33.66%	42.55%	37.74%	40.00%	58.00%	54.72%	56.31%
	C	46.07%	46.59%	46.33%	46.07%	46.59%	46.33%	52.58%	57.95%	55.14%	62.50%	68.18%	65.22%
	D	33.33%	40.68%	36.64%	33.33%	40.68%	36.64%	37.88%	42.37%	40.00%	46.88%	50.85%	48.78%
	Accuracy	39.82%			39.82%			46.46%			58.41%		

Table 3: Results of models performance on **admission scores and gender features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score
RFR	A	73.33%	42.31%	53.66%	75.00%	46.15%	57.14%	82.35%	53.85%	65.12%	94.74%	69.23%	80.00%
	B	50.00%	52.83%	51.38%	51.85%	52.83%	52.34%	58.93%	62.26%	60.55%	72.22%	73.58%	72.90%
	C	49.62%	73.86%	59.36%	50.38%	76.14%	60.63%	56.80%	80.68%	66.67%	65.00%	88.64%	75.00%
	D	62.50%	25.42%	36.14%	65.22%	25.42%	36.59%	82.14%	38.98%	52.87%	93.94%	52.54%	67.39%
	Accuracy	52.65%			53.98%			62.39%			73.45%		
SVM	A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	B	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	C	38.94%	100.00%	56.05%	39.11%	100.00%	56.23%	44.67%	100.00%	61.75%	48.89%	100.00%	65.67%
	D	0.00%	0.00%	0.00%	100.00%	1.69%	3.33%	100.00%	49.15%	65.91%	100.00%	77.97%	87.62%
	Accuracy	38.94%			39.38%			51.77%			59.29%		
GNB	A	36.67%	42.31%	39.29%	36.67%	42.31%	39.29%	37.93%	42.31%	40.00%	48.28%	53.85%	50.91%
	B	31.25%	37.74%	34.19%	31.75%	37.74%	34.48%	37.29%	41.51%	39.29%	46.55%	50.94%	48.65%
	C	39.62%	23.86%	29.79%	41.82%	26.14%	32.17%	53.45%	35.23%	42.47%	67.19%	48.86%	56.58%
	D	32.91%	44.07%	37.68%	33.33%	44.07%	37.96%	37.50%	50.85%	43.17%	45.33%	57.63%	50.75%
	Accuracy	34.51%			35.40%			41.59%			52.21%		

Table 4: Results of models performance on **admission scores and first level English courses scores features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score	precision	recall	f1 score
RFR	A	61.90%	50.00%	55.32%	61.90%	50.00%	55.32%	70.83%	65.38%	68.00%	88.46%	88.46%	88.46%
	B	54.00%	50.94%	52.43%	57.14%	52.83%	54.90%	65.91%	54.72%	59.79%	88.64%	73.58%	80.41%
	C	53.85%	71.59%	61.46%	55.56%	73.86%	63.41%	60.00%	81.82%	69.23%	76.64%	93.18%	84.10%
	D	60.53%	38.98%	47.42%	61.54%	40.68%	48.98%	73.68%	47.46%	57.73%	91.84%	76.27%	83.33%
	Accuracy	55.75%			57.52%			64.6%			83.63%		
SVM	A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	3.85%	7.41%	100.00%	23.08%	37.50%
	B	21.43%	5.66%	8.96%	21.43%	5.66%	8.96%	23.08%	5.66%	9.09%	61.54%	15.09%	24.24%
	C	42.11%	100.00%	59.26%	43.78%	100.00%	60.90%	49.16%	100.00%	65.92%	57.14%	100.00%	72.73%
	D	100.00%	5.08%	9.68%	100.00%	18.64%	31.43%	100.00%	55.93%	71.74%	100.00%	89.83%	94.64%
	Accuracy	41.59%			45.13%			55.31%			68.58%		
GNB	A	66.67%	53.85%	59.57%	70.00%	53.85%	60.87%	77.78%	53.85%	63.64%	84.21%	61.54%	71.11%
	B	41.94%	49.06%	45.22%	42.86%	50.94%	46.55%	46.15%	56.60%	50.85%	54.55%	67.92%	60.50%
	C	46.39%	51.14%	48.65%	47.37%	51.14%	49.18%	52.13%	55.68%	53.85%	63.64%	63.64%	63.64%
	D	52.17%	40.68%	45.71%	54.17%	44.07%	48.60%	61.22%	50.85%	55.56%	71.70%	64.41%	67.86%
	Accuracy	48.23%			49.56%			54.42%			64.6%		

Table 5: Results of models performance on **admission scores and first level Math courses scores features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score
RFR	A	66.67%	46.15%	54.55%	66.67%	46.15%	54.55%	71.43%	57.69%	63.83%	77.27%	65.38%	70.83%
	B	58.97%	43.40%	50.00%	60.00%	45.28%	51.61%	67.50%	50.94%	58.06%	79.07%	64.15%	70.83%
	C	52.99%	80.68%	63.96%	53.38%	80.68%	64.25%	60.63%	87.50%	71.63%	70.43%	92.05%	79.80%
	D	62.86%	37.29%	46.81%	62.86%	37.29%	46.81%	81.58%	52.54%	63.92%	89.13%	69.49%	78.10%
	Accuracy	56.64%			57.08%			66.37%			76.55%		
SVM	A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	7.69%	14.29%
	B	35.00%	13.21%	19.18%	35.00%	13.21%	19.18%	50.00%	24.53%	32.91%	70.27%	49.06%	57.78%
	C	41.04%	62.50%	49.55%	41.04%	62.50%	49.55%	42.97%	62.50%	50.93%	49.57%	64.77%	56.16%
	D	51.39%	62.71%	56.49%	51.39%	62.71%	56.49%	51.39%	62.71%	56.49%	54.17%	66.10%	59.54%
	Accuracy	43.81%			43.81%			46.46%			54.87%		
GNB	A	53.57%	57.69%	55.56%	53.57%	57.69%	55.56%	53.57%	57.69%	55.56%	61.54%	61.54%	61.54%
	B	35.90%	26.42%	30.43%	36.84%	26.42%	30.77%	43.90%	33.96%	38.30%	53.49%	43.40%	47.92%
	C	47.47%	53.41%	50.27%	48.51%	55.68%	51.85%	53.47%	61.36%	57.14%	63.16%	68.18%	65.57%
	D	45.00%	45.76%	45.38%	45.76%	45.76%	45.76%	50.00%	47.46%	48.70%	61.29%	64.41%	62.81%
	Accuracy	45.58%			46.46%			50.88%			60.62%		

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3250702

Table 6: Results of models performance on **admission scores and first level CS courses scores features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score
RFR	A	70.00%	28.00%	40.00%	72.73%	32.00%	44.44%	75.00%	36.00%	48.65%	94.44%	68.00%	79.07%
	B	50.85%	55.56%	53.10%	52.63%	55.56%	54.05%	59.32%	64.81%	61.95%	75.86%	81.48%	78.57%
	C	55.46%	75.00%	63.77%	55.83%	76.14%	64.42%	66.97%	82.95%	74.11%	78.79%	88.64%	83.42%
	D	65.79%	42.37%	51.55%	65.79%	42.37%	51.55%	80.43%	62.71%	70.48%	88.24%	76.27%	81.82%
	Accuracy		56.64%			57.52%			68.14%			81.42%	
SVM	A	59.09%	52.00%	55.32%	59.09%	52.00%	55.32%	65.38%	68.00%	66.67%	71.43%	80.00%	75.47%
	B	53.06%	48.15%	50.49%	53.06%	48.15%	50.49%	57.78%	48.15%	52.53%	63.64%	51.85%	57.14%
	C	52.17%	81.82%	63.72%	52.55%	81.82%	64.00%	59.02%	81.82%	68.57%	68.57%	81.82%	74.61%
	D	70.59%	20.34%	31.58%	72.22%	22.03%	33.77%	84.85%	47.46%	60.87%	89.80%	74.58%	81.48%
	Accuracy		54.42%			54.87%			63.27%			72.57%	
GNB	A	61.54%	32.00%	42.11%	61.54%	32.00%	42.11%	61.54%	32.00%	42.11%	78.57%	44.00%	56.41%
	B	42.86%	44.44%	43.64%	42.86%	44.44%	43.64%	49.06%	48.15%	48.60%	55.77%	53.70%	54.72%
	C	52.08%	56.82%	54.35%	52.63%	56.82%	54.64%	59.14%	62.50%	60.77%	67.03%	69.32%	68.16%
	D	45.90%	47.46%	46.67%	46.77%	49.15%	47.93%	50.75%	57.63%	53.97%	59.42%	69.49%	64.06%
	Accuracy		48.67%			49.12%		54.42%				62.83%	

Table 7: Results of models performance on **admission scores and first level courses (Eng, Math, CS) scores features** on a sample of 226 students with grade support: A's=26, B's=53, C's=88, D's=59)

		Normal			Rounding 0.10±			Rounding 0.30±			Rounding 0.50±		
Model	GPA	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score	precsion	recall	f1_score
RFR	A	78.26%	69.23%	73.47%	79.17%	73.08%	76.00%	80.00%	76.92%	78.43%	92.31%	92.31%	92.31%
	B	67.39%	58.49%	62.63%	68.89%	58.49%	63.27%	78.57%	62.26%	69.47%	91.11%	77.36%	83.67%
	C	61.26%	77.27%	68.34%	61.26%	77.27%	68.34%	69.64%	88.64%	78.00%	81.82%	92.05%	86.63%
	D	71.74%	55.93%	62.86%	71.74%	55.93%	62.86%	85.11%	67.80%	75.47%	91.07%	86.44%	88.70%
	Accuracy		66.37%			66.81%			75.66%			87.17%	
SVM	A	76.00%	73.08%	74.51%	76.92%	76.92%	76.92%	77.78%	80.77%	79.25%	77.78%	80.77%	79.25%
	B	77.27%	31.48%	44.74%	80.95%	31.48%	45.33%	85.00%	31.48%	45.95%	86.96%	37.04%	51.95%
	C	51.32%	88.64%	65.00%	52.00%	88.64%	65.55%	55.17%	90.91%	68.67%	59.42%	93.18%	72.57%
	D	70.37%	32.76%	44.71%	72.41%	36.21%	48.28%	82.35%	48.28%	60.87%	89.47%	58.62%	70.83%
	Accuracy		58.85%			60.18%			64.6%			69.47%	
GNB	A	66.67%	53.85%	59.57%	66.67%	53.85%	59.57%	75.00%	57.69%	65.22%	87.50%	80.77%	84.00%
	B	48.78%	37.74%	42.55%	48.78%	37.74%	42.55%	61.90%	49.06%	54.74%	80.00%	60.38%	68.82%
	C	53.77%	64.77%	58.76%	54.29%	64.77%	59.07%	63.64%	71.59%	67.38%	74.73%	77.27%	75.98%
	D	50.00%	49.15%	49.57%	50.85%	50.85%	50.85%	58.46%	64.41%	61.29%	66.20%	79.66%	72.31%
	Accuracy		53.1%			53.54%			62.83%			74.34%	

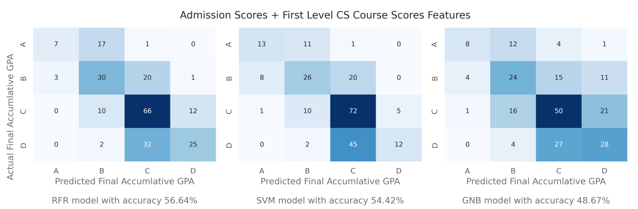

Figure 10: Admission scores and first level Computer Science courses (CS) scores features.

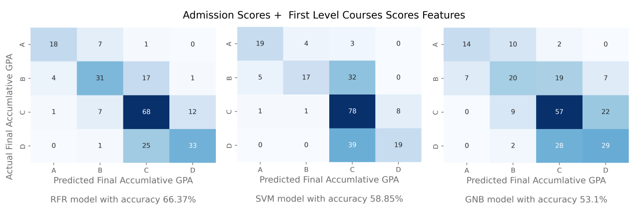

Figure 11: Admission scores and first level courses (Eng, Math, CS) scores features.

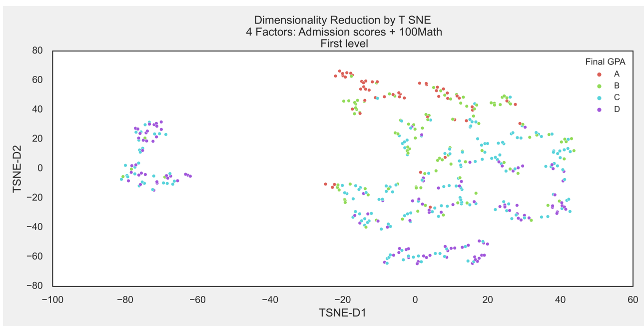

Figure 12: Admission scores and 100 Math using dimensionality reduction by tSNE

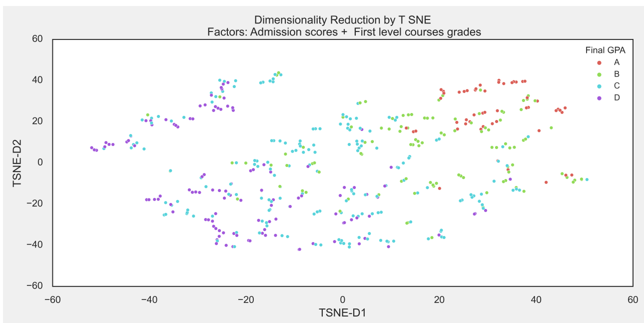

Figure 13: Admission scores and First level courses grades using dimensionality reduction by tSNE

is an effective nonlinear dimensionality reduction technique for visualizing data sets containing hundreds or even thousands of dimensions in 2D and 3D maps. We use 2D maps in this work. We use t-SNE to a lower-dimensional space to make the data easier for analyzing and visualizing. t-SNE reduced our data into 2D therefore, we can see the relationship between the classes.

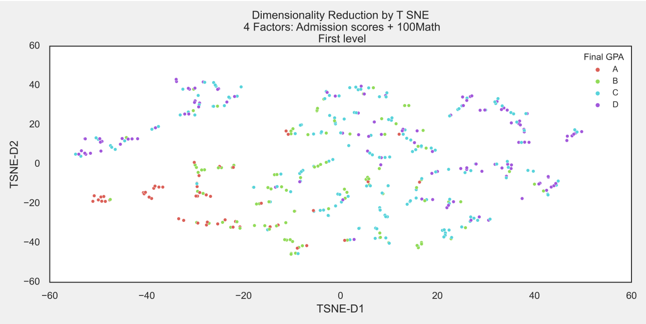

Figure 14: Admission scores and English using dimensionality reduction by tSNE

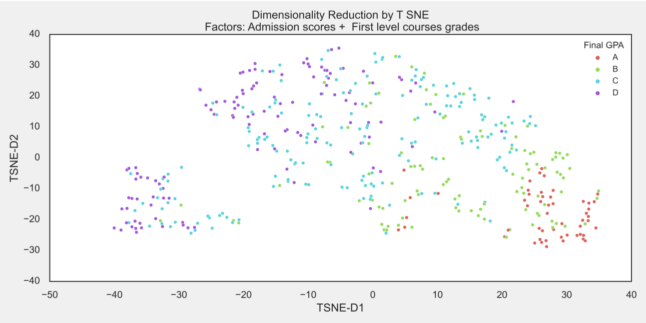

Figure 15: Admission scores and All first level courses grades using dimensionality reduction by tSNE

I. PERFORMANCE MODELS AFTER ROUNDING GRADES

We perform six different kinds of experiments with varying combination of features (admission score, admission scores+gender, admission scores+first level Eng courses scores, admission scores+first level Math courses scores, admission scores+first level CS courses scores, and admission scores+first level courses (Eng,Math, CS scores) features as shown in tables 2, 3, 4, 5, 6, and 7. The intention behind such experiments is to show the significance of our features after rounding grades of students with different scales in achieving better accuracy. We specifically round the GPA to the closest decimal point on different ranges including 0.10, 0.30, 0.50 before measuring the accuracy of the model based on the GPA label. This method simplifies the assessment of different results of the model performance and avoids difficulties when applying errors-based metrics that might be hard to interpret.

We compare the performance of our method with different common machine learning classifiers. Tables 2 and 3 show comparison of our results with different classifiers based on recall, precision and F1-score on the a sample of 226 students with grade support: A's=26, B's=53, C's=88 and D's=59. It is evident that our approach performs better with increasing the threshold of Rounding which is to compute the absolute difference between a grade and following grade after or before. Also, the performance of our approach improves substantially as we combine more features to predict student performance. The reason of the threshold causes increasing of the accuracy of the model is, as we see in Figures 12, 13, 14 and 15, all students' grades in the same category are very close to the grades in the next category.

J. EVALUATION METRICS

The evaluation of machine learning classifiers is critical when studying the learning models and their performance. To evaluate the performance of the classifier models, we have used similar evaluation measures that are adopted in most of the previous research experiments. It covers the prediction accuracy and F1-score under varying conditions of input parameters. Most of the time, we use classification accuracy to measure the performance of machine learning models, and we have also used confusion matrices to compare the prediction accuracy and failures.

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

$$Precision(P) = \frac{TP}{TP + FP} \tag{2}$$

$$Recall(R) = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

In the above equations, TP, TN, FN, FN, and FP are true positives, true negatives, false positives, and false negatives. We use F1-score as the primary performance indicator to evaluate all the classifier models used in our experiments. F1 score is a single metric that combines both precision and recall. The precision or True Positive Rate (TPR) is a way to look at the accuracy of positive predictions performed by a classifier and can be defined as follows: precision $= TP/(TP + FP)$ where the True Positives (TP) is the number of true positives, i.e., correct prediction of a positive sample, and the False Positives (FP), i.e., the wrong prediction. But precision is used with another parameter called recall. Recall is defined by $TP/(TP + FN)$.

We have also used a confusion matrix table to study the performance of classifiers. The confusion matrix is a table with rows and columns that report false positives, false negatives, true positives, and true negatives. This allows a more detailed analysis than the mere proportion of correct classifications, i.e., prediction accuracy.

VI. THREAT VALIDITY

VII. CONCLUSIONS

The student performance is a vital issue. It is difficult to deal with this issue. This paper presented an analysis of the results data mining research to develop models of studentsâ performance prediction. Our paper showed the use of machine learning algorithms to be better understand efficiency of the algorithms with data dimensionality reduction by T-SNE. It uses four factors such as admission scores and first level courses, academic achievement test (AAT) and general aptitude test (GAT).In the future, we would like to use deep learning architectures to construct the prediction and improve performance. It can be combined non-academic features with academic features.

References

[1] F. Giannakas, C. Troussas, I. Voyiatzis, and C. Sgouropoulou, "A deep learning classification framework for early prediction of team-based academic performance," *Applied Soft Computing*, vol. 106, p. 107355, 2021.

[2] H. Hamsa, S. Indiradevi, and J. J. Kizhakkethottam, "Student academic performance prediction model using decision tree and fuzzy genetic algorithm," *Procedia Technology*, vol. 25, pp. 326–332, 2016.

[3] B. K. Francis and S. S. Babu, "Predicting academic performance of students using a hybrid data mining approach," *Journal of medical systems*, vol. 43, no. 6, pp. 1–15, 2019.

[4] C. Romero and S. Ventura, "Educational data mining: A survey from 1995 to 2005," *Expert systems with applications*, vol. 33, no. 1, pp. 135–146, 2007.

[5] S. Helal, J. Li, L. Liu, E. Ebrahimie, S. Dawson, D. J. Murray, and Q. Long, "Predicting academic performance by considering student heterogeneity," *Knowledge-Based Systems*, vol. 161, pp. 134–146, 2018.

[6] X. Xu, J. Wang, H. Peng, and R. Wu, "Prediction of academic performance associated with internet usage behaviors using machine learning algorithms," *Computers in Human Behavior*, vol. 98, pp. 166–173, 2019.

[7] T. T. Dien, S. H. Luu, N. Thanh-Hai, and N. Thai-Nghe, "Deep learning with data transformation and factor analysis for student performance prediction."

[8] D. T. Ha, P. T. T. Loan, C. N. Giap, and N. T. L. Huong, "An empirical study for student academic performance prediction using machine learning techniques," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 18, no. 3, 2020.

[9] J. L. Rastrollo-Guerrero, J. A. Gomez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting studentsâ performance by means of machine learning: A review," *Applied sciences*, vol. 10, no. 3, p. 1042, 2020.

[10] M. Agaoglu, "Predicting instructor performance using data mining techniques in higher education," *IEEE Access*, vol. 4, pp. 2379–2387, 2016.

[11] A. Gonzalez-Nucamendi, J. Noguez, L. Neri, V. Robledo-Rella, R. M. G. García-Castelán, and D. Escobar-Castillejos, "The prediction of academic performance using engineering studentâs profiles," *Computers & Electrical Engineering*, vol. 93, p. 107288, 2021.

[12] A. Alshanqiti and A. Namoun, "Predicting student performance and its influential factors using hybrid regression and multi-label classification," *IEEE Access*, vol. 8, pp. 203 827–203 844, 2020.

[13] L. S. Maurya, M. S. Hussain, and S. Singh, "Developing classifiers through machine learning algorithms for student placement prediction based on academic performance," *Applied Artificial Intelligence*, vol. 35, no. 6, pp. 403–420, 2021.

[14] N. Aslam, I. Khan, L. Alamri, and R. Almuslim, "An improved early studentâs academic performance prediction using deep learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 12, pp. 108–122, 2021.

[15] D. M. Ahmed, A. M. Abdulazeez, D. Q. Zeebaree, and F. Y. Ahmed, "Predicting university's students performance based on machine learning techniques," in *2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS)*. IEEE, 2021, pp. 276–281.

[16] X. Liu and L. Niu, "A student performance predication approach based on multi-agent system and deep learning," in *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*. IEEE, 2021, pp. 681–688.

[17] E. De Leon Evangelista and B. D. Sy, "An approach for improved studentsâ performance prediction using homogeneous and heterogeneous ensemble methods," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 5, p. 5226, 2022.

[18] T. L. Quy, T. H. Nguyen, G. Friege, and E. Ntoutsi, "Evaluation of group fairness measures in student performance prediction problems," *arXiv preprint arXiv:2208.10625*, 2022.

[19] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, pp. 1–19, 2022.

[20] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[21] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[22] S. Russell and P. Norvig, "Artificial intelligence: a modern approach," 2002.

[23] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *the Journal of machine Learning research*, vol. 9, pp. 1871–1874, 2008.

[24] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.

[25] M. Pal, "Random forest classifier for remote sensing classification," *International journal of remote sensing*, vol. 26, no. 1, pp. 217–222, 2005.

ESSA ALHAZMI is an assistance professor in computer science program at Jazan University, Saudi Arabia. He is currently leading the vice deanship for research and development in the college of computer science and information technology at Jazan University. Dr. Alhazmi focuses on quantifying and modeling large-scale data using artificial intelligence algorithms. His research interests include different areas in data science, social computing, and networks science. Dr. Alhazmi has contributed and participated in academia and the research community by publishing and reviewing papers in different conferences and journals.

ABDULLAH SHENEAMER received the BSc degree in Computer Science in 2008 from King Abdulaziz University, Saudi Arabia; the MSc degree in Computer Science in 2012, and the PhD degree in Computer Science in 2017, both from University of Colorado at Colorado Springs, USA. He is an Associate Professor of Computer Science. He used to work as a former vice-dean of Faculty of Computer Science and Information Technology at the Jazan University, KSA. His research interests include data mining, machine learning, software engineering, and malware analysis. His current work focuses on software clone and refactoring software clone, malware detection, and code obfuscation detection using machine learning approaches. He published several papers in reputed international journals and conferences. He had reviewed several papers in reputed journals such as IEEE Access, IEEE Transaction on Parallel and Distributed Systems and Information Sciences at Elsevier, and Frontiers of Computer Science. He is a senior meta reviewer of IEEE International Conference on Machine Learning and Applications.