Student Early Grade Prediction Model Using Machine Learning Algorithms with Educational Data Mining

[1]Lohitha Chitrada, [2]S Prasad Babu Vagolu, [3] Prof. Vedavathi K

[1]PG Student, [2]Asst Professor, [3]Professor and HOD, Department of Computer Science, GITAM Institute of Science, GITAM, India.

ABSTRACT - Machine Learning is a part of data science in which we create the models and use them to predict. And it deals with many algorithms which are accurate.

It is at the top of the line to the most intriguing vocations with regards to information examination today. As information sources multiply alongside the figuring capacity to process them, going directly to the information is a standout amongst the most suitable approaches to rapidly pick up experiences and make expectations.

Machine Learning can be used for certain aspects like regression, classification, decision making, clustering, forecasting, deep learning, inductive logic programming, hereditary calculations, SVM inadequate word reference learning, and so on. Directed learning or arrangement is the AI assignment of deriving a capacity from a piece of named information.

Machine learning is used in artificial intelligence (AI).AI is the physical entity of Machine learning. In artificial intelligence the body takes the data from the sensor or any tools. AI uses Machine learning in order to check the given data from the receptors and acts according to it. AI body learns from its previous data all by themselves and works according to the events.

In Machine learning the process begins with the analysing of previous data provided by the users, sensors or direct experiences. In the process we analyse certain patterns which effect the output. After analysing we train the model by using these patterns. Then the model is ready to predict the output. Here the process has less priority than the data. The main motto is to train the machine in order to avoid human interaction.

Student Grade Prediction is a way of predicting a student grade based on his/her previous marks. This also makes the student know whether he/she is in a position to reach his/her expected marks or not. If this model shows that he/she needs to improve then that student can prepare more for that semester so that he/she can reach their expected score. This help students to know their capabilities and also their weaknesses so that they can make use of those capabilities and work even more on them to get great opportunities and also to make them know their weaknesses so that they can strive hard to overcome them and then achieve their expected scores.

Keywords: Machine Learning, Student Grade, Data Mining, Scipy, Numpy, Matplotlib, Random forest

I. INTRODUCTION

Extensive efforts have been made in order to predict student performance for different aims, like: detecting at risk students, assurance of student retention, course and resource allocations, and many others. This research aims to predict student performance to engage distinct students in researches and innovative projects that could improve universities reputation and ranking nationally and internationally. However, analyzing students records for start up to medium size institutes or schools, like the of small students' dataset size in educational domains.

Additionally, in most researches that were aimed to classify or predict, researchers used to spend much efforts just to extract the important indicators that could be more useful in British University in Dubai which have small size of students records, have never been explored in educational or learning analytics domain. Yet, that were

 DOI : 10.35291/2454-9150.2021.0265

investigated in other fields, like: health sciences and Chemists .So, this project aims to explore the utilization possibility constructing reasonable accurate predictive models. They will either use features ranking algorithms or will look at the selected features while training the dataset on different machine learning algorithms. Instead, and until recently, there have been no research efforts to investigate the ability of visualization or clustering techniques in identifying such indicators for small dataset, especially in the learning analytics domain .If such studies will be conducted, its outcomes might prove the feasibility of mitigating the hassle that is normally spent on features extraction or selection processes.

Prediction of student's performance became an urgent desire in most of educational entities and institutes. That is essential in order to help at-risk students and assure their retention, providing the excellent learning resources and experience, and improving the university's ranking and reputation. However, that might be difficult to be achieved for startup to mid-sized universities, especially those which are specialized in graduate and post graduate programs, and have small students' records for analysis. So, the main aim of this project is to prove the possibility of training and modelling a small dataset size and the feasibility of creating a prediction model with credible accuracy rate. This also makes the student know whether he/she is in a position to reach his/her expected marks or not. If this model shows that he/she needs to improve then that student can prepare more for that semester so that he/she can reach their expected score.

IN SIMPLE TERMS:

Student Grade Prediction is a way of predicting a student grade based on his/her previous marks.

This also makes the student know whether he/she is in a position to reach his/her expected marks or not. If this model shows that he/she needs to improve then that student can prepare more for that semester so that he/she can reach their expected score.

		Predicted			
		C1	C2	C3	Σ
Actual	C1	96.8 %	0.9 %	0.6 %	212
	C2	1.4 %	93.9 %	0.6 %	112
	C3	1.8 %	5.2 %	98.8 %	176
	Σ	217	115	168	500

Fig 1.1

II. LITERATURE REVIEW

Peoplediscussedtheforemostreasonsfortheincreasingdemand ofpython.x While though python is slower in runtime compared to the compiled language slike C,C++,it is preferred by developers within the field of information analytics, numerical computations and majority technical domains like AI, ML, Deep Learning etc.

We have explained about the python importance within the world of programming. We discussed the important libraries and packages that are available in python. Python provides many tools to make the exploration of scientific problems easier. One in every of its strength is NetworkX in conjunction with packages Scipy, Numpy, Matplotlib and their links to LINPACK, ODE integration tools and other tools written in FORTRAN and C allows analysis and implementation of algorithms for analyzing dynamics of network coupled oscillation .

Small world phenomena are been observed and applied in many types of technologies. A study of information within which the attributes behave and relationship exists. The project is all about data and its behavior. Prediction of the final grade can be defined as follows "Given a dataset containing attribute of 396 Portuguese students where using the features available from dataset and define classification algorithms to identify whether the student performs good in final grade exam, also to evaluate different machine learning models on the dataset."

This also makes the student know whether he/she is in a position to reach his/her expected marks or not. If this model shows that he/she needs to improve then that student can prepare more for that semester so that he/she can reach their expected score.

USAGE:

By using web Link one can access the page easily and can enter his/her details.

By using them the model will Predict the GRADE of the student based on his/her previous marks.

Model can even predict whether he/she will pass/fail in the next semester

A message will also be given based on his/her Grade Such as:

1. If the student grade is "A" then the message is "You Are Excellent".

2. If the student grade is "B" then the message is "You Are Good".

3. If the student grade is "C" then the message is "You Are OK".

4. If the student grade is "D" then the message is "You Need To Improve".

5.If the student grade is "F" then the message is "You Need To Work Hard".

By that message the student can prepare according to that.

CONTRIBUTING:

Student Grade prediction is a model which is designed using "Machine Learning" technology. Machine Learning means predicting the present based on past scenarios and

 DOI : 10.35291/2454-9150.2021.0265

predicting the future based on past and present scenarios.

In higher educational institutes, early grade prediction is an important area of interest as it allows instructors to improve students' performance in their courses by providing special attention at the early stages. Machine learning techniques can be utilized for students' grades prediction in different courses. However, the performance of these techniques is highly dependent on the quality of data that made the selection of model a challenging task. Therefore, in this paper, we evaluate different state-of-the-art machine learning techniques for university students grade prediction. Ultimately we find that Restricted Boltzmann Machines (RBM) can more accurately predict students' grades. The predicted grades by these techniques visualize uncertainty on student learning and can be used for confidence gains, student degree planning, personalized advising, and to enable instructors to identify potential students who might need assistance in relevant courses.

PROBLEM IDENTIFICATION

The problem statement can be defined as follows "Given a dataset containing attribute of 396 Portuguese students where using the features available from dataset and define classification algorithms to identify whether the student performs good in final grade exam, also to evaluate different machine learning models on the dataset."

OBJECTIVE

The main objective of this project is to use data mining methodologies to study students performance in the courses. Data mining provides many tasks that could be used to study the student performance. In this research, the classification task is used to evaluate students performance and as there are many approaches that are used for data classification, the decision tree method is used here. Information like Attendance, Class test, and Assignment marks was collected from the students management system, to predict the performance at the end of the semester.

The overall vision for the Performance Prediction System is that it will fulfill the following objectives:

• To create a user friendly web interface on which the system can be implemented.

• To be able to predict the student performance using the Naive Bayes and ID3 algorithms.

• To determine the more efficient data mining classifier among the two classifiers used.

• To be able to make the performance prediction methodology more efficient and accurate.

Prediction of the final grade of Portuguese high school students. The overall study is explained in four sections, including this introduction. The following section will talk about the used methodology.

III. PROPOSEDMETHODOLOGY

Since universities are prestigious places of higher education, students' retention in these universities is a matter of high concern. It has been found that most of the students' drop-out from the universities during their first year is due to lack of proper support in undergraduate courses. Due to this reason, the first year of the undergraduate student is referred as a "make or break" year. Without getting any support on the course domain and its complexity, it may de motivate a student and can be the cause to withdraw the course.

There is a great need to develop an appropriate solution to assist students retention at higher education institutions. Early grade prediction is one of the solutions that have a tendency to monitor students' progress in the degree courses at the University and will lead to improving the students' learning process based on predicted grades.

Using machine learning with Educational Data Mining can improve the learning process of students. Different models can be developed to predict students' grades in the enrolled courses, which provide valuable information to facilitate students' retention in those courses. This information can be used to early identify students at-risk based on which a system can 1 suggest the instructors to provide special attention to those students. This information can also help in predicting the students' grades in different courses to monitor their performance in a better way that can enhance the students' retention rate of the universities.

Using various packages such as cufflinks, seaborn&matplotlib to represent the data along with different attributes graphically or pictorially to analyze the dataset for predicting the Final Grade(G3).

IV. IMPLEMENTATION

The dataset collected belongs to a Portuguese high school students from the internet.

Inputs: Independent variables that influence the behavior of output variable.

['school', 'sex', 'age', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu',

'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime', 'studytime',

'failures', 'schoolsup', 'famsup', 'paid', 'activities', 'nursery',

'higher', 'internet', 'romantic', 'famrel', 'freetime', 'goout', 'Dalc',

'Walc', 'health', 'absences', 'G1', 'G2', 'G3'],

dtype='object')]

 DOI : 10.35291/2454-9150.2021.0265

Output: Exited

```
Index(['school', 'sex', 'age', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu',
       'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime', 'studytime',
       'failures', 'schoolsup', 'famsup', 'paid', 'activities', 'nursery',
       'higher', 'internet', 'romantic', 'famrel', 'freetime', 'goout', 'Dalc',
       'Walc', 'health', 'absences', 'G1', 'G2', 'G3'],
      dtype='object')
```

Describing the data:

	age	Medu	Fedu	traveltime	studytime	failures	famrel	freetime
count	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000	395.000000
mean	16.696203	2.749367	2.521519	1.448101	2.035443	0.334177	3.944304	3.235443
std	1.276043	1.094735	1.088201	0.697505	0.839240	0.743651	0.896659	0.998862
min	15.000000	0.000000	0.000000	1.000000	1.000000	0.000000	1.000000	1.000000
25%	16.000000	2.000000	2.000000	1.000000	1.000000	0.000000	4.000000	3.000000
50%	17.000000	3.000000	2.000000	1.000000	2.000000	0.000000	4.000000	3.000000
75%	18.000000	4.000000	3.000000	2.000000	2.000000	0.000000	5.000000	4.000000
max	22.000000	4.000000	4.000000	4.000000	4.000000	3.000000	5.000000	5.000000

Exploratory data analysis:

Correlation:

```
stud.corr()['G3'].sort_values()
failures     -0.360415
age          -0.161579
goout        -0.132791
romantic     -0.129970
traveltime   -0.117142
schoolsup    -0.082788
guardian     -0.070109
health       -0.061335
Pstatus      -0.058009
Dalc         -0.054660
Walc         -0.051939
school       -0.045017
famsup       -0.039157
freetime      0.011307
activities    0.016100
absences      0.034247
Fjob          0.042286
famrel        0.051363
nursery       0.051568
famsize       0.081407
studytime     0.097820
internet      0.098483
paid          0.101996
Mjob          0.102082
sex           0.103456
address       0.105756
reason        0.121994
Fedu          0.152457
higher        0.182465
Medu          0.217147
G1            0.801468
G2            0.904868
G3            1.000000
Name: G3, dtype: float64
```

Data visualization (HEAT MAP):

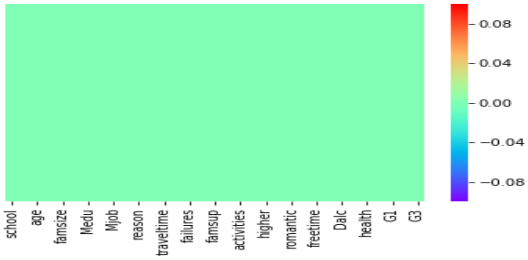
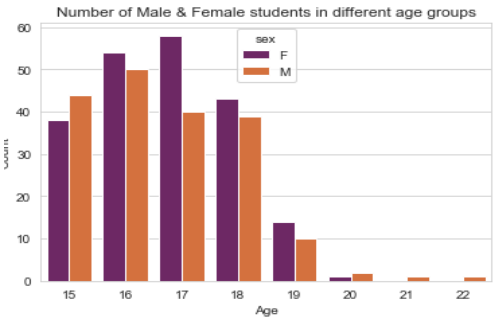


Fig 1.2 BAR PLOTs FOR DIFFERENT ATTRIBUTES:



- The student age seems to be ranging from 15-19, where gender distribution
- The age group above 19 may be outliers, year back students or droupouts.

Fig 1.3

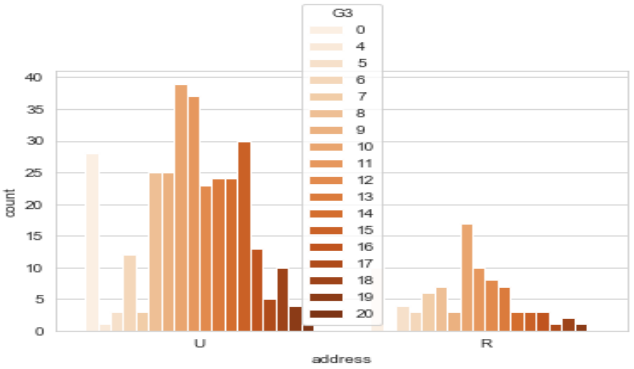


Fig 1.4

Boxplot of final grades

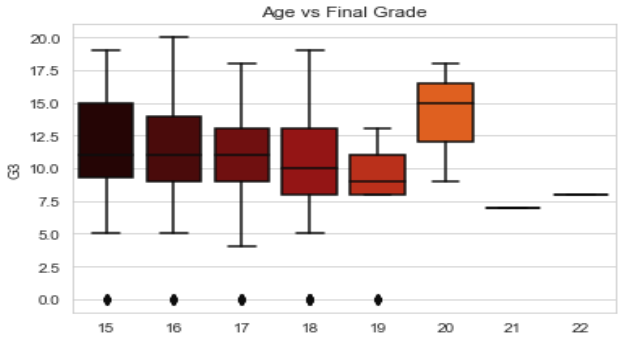


Fig 1.5

Testing and Training the data

In this model we use the testing size of 20% and the training size of 80%.

We split the data into x_train, y_train, x_test and y_test. The training data is used to train the model. The testing data is used for prediction by using the model which is trained from the training data. The training data is taken random from 365 rows. For every compilation in the model the data changes and the results may vary every time.

Here training and testing is done by The normal method.

```
x_train.head()
```

	G3	failures	Medu	higher	age	Fedu	goout	romantic	reason
16	14	0	4	1	16	4	3	0	3
66	12	0	4	1	15	4	3	1	3
211	13	0	4	1	17	4	5	1	1
7	6	0	4	1	17	4	4	0	1
19	10	0	4	1	16	3	3	0	1

Fig 1.6

V. ALGORITHMS USED

LINEAR REGRESSION -

Linear regressionmay be defined as the statistical model that analyzes the linear relationship between a dependent variable with given set of independent variables. Linear relationship between variables means that when the value of one or more independent variables will change (increase or decrease), the value of dependent variable will also change accordingly (increase or decrease)

 DOI : 10.35291/2454-9150.2021.0265

Elastic-Net REGRESSION-

The **Elastic-Net** is a regularized regression method that linearly combines both penalties i.e. L1 and L2 of the Lasso and Ridge regression methods. It is useful when there are multiple correlated features. The difference between Lass and Elastic-Net lies in the fact that Lasso is likely to pick one of these features at random while elastic-net is likely to pick both at once.

SVM -

SVM is fully known as Support Vector Machine. SVM is in the list of supervised machine learning algorithms which performs analysis through differentiating the data points which may be linearly or non-linearly separable. It uses the hyper plane to separate the data points. However, it is mostly employed in classification problems

RANDOM FOREST ALGORITHM–

Random forest algorithm is used in machine learning which is employed for both classification and regression analysis. It involves building several decision trees and combining the output to give the result of the model. Each individual tree in the forest gives a class prediction value and the class which gets the maximum number of votes is declared as the result of the classifier.

Extra Trees ALGORITHM–

Extra trees algorithm is another extension of bagged decision tree ensemble method. In this method, the random trees are constructed from the samples of the training dataset.

Gradient Boosted ALGORITHM–

In **gradient boosting**, many models are trained sequentially. Each new model gradually minimizes the loss function (y = ax + b + e, where 'e' is the error term) of the whole system using Gradient Descent method.

The learning method consecutively fits new models to give a more accurate estimate of the response variable. The main idea behind this algorithm is to construct new base learners which can be optimally correlated with negative gradient of the loss function, relevant to the whole ensemble.

In Python Sklearn library, we use Gradient Tree Boosting or GBRT which is a generalization of boosting to arbitrary differentiable loss functions. It can be utilized for both regression and classification problems.

Base Line

A **baseline** is a method that uses heuristics, simple summary statistics, randomness, or machine learning to create predictions for a dataset. You can use these predictions to measure the baseline's performance (e.g., accuracy)-- this metric will then become what you compare any other machine learning algorithm against.

```
# Evaluate several ml models by training on training set and testing on testing set
def evaluate(X_train, X_test, y_train, y_test):
    # Names of models
    model_name_list = ['Linear Regression', 'ElasticNet Regression',
                       'Random Forest', 'Extra Trees', 'SVM',
                       'Gradient Boosted', 'Baseline']
    X_train = X_train.drop('G3', axis='columns')
    X_test = X_test.drop('G3', axis='columns')

    # Instantiate the models
    model1 = LinearRegression()
    model2 = ElasticNet(alpha=1.0, l1_ratio=0.5)
    model3 = RandomForestRegressor(n_estimators=100)
    model4 = ExtraTreesRegressor(n_estimators=100)
    model5 = SVR(kernel='rbf', degree=3, C=1.0, gamma='auto')
    model6 = GradientBoostingRegressor(n_estimators=50)

    # Dataframe for results
    results = pd.DataFrame(columns=['mae', 'rmse'], index = model_name_list)

    # Train and predict with each model
    for i, model in enumerate([model1, model2, model3, model4, model5, model6]):
        model.fit(X_train, y_train)
        predictions = model.predict(X_test)

        # Metrics
        mae = np.mean(abs(predictions - y_test))
        rmse = np.sqrt(np.mean((predictions - y_test) ** 2))

        # Insert results into the dataframe
        model_name = model_name_list[i]
        results.loc[model_name, :] = [mae, rmse]

    # Median Value Baseline Metrics
    baseline = np.median(y_train)
    baseline_mae = np.mean(abs(baseline - y_test))
    baseline_rmse = np.sqrt(np.mean((baseline - y_test) ** 2))

    results.loc['Baseline', :] = [baseline_mae, baseline_rmse]

    return results
```

PREDICTOR:

We used median baseline to predict the new values given by the user. We take all normalized values as inputs and tell about churn.

```
Median Baseline  MAE: 3.7879
Median Baseline RMSE: 4.8252
```

VI. RESULTS AND DISCUSSIONS

Since the accuracy is the main key metric that the evaluation of machine learning models will be relying on, the baseline accuracy is calculated at first (also called 'no information' rate) for both datasets.

After cleaning the dataset we test and train the system to the dataset and then we used seven different algorithms(Linear Regression, SVM algorithm, Random forest algorithm, Elastic net Regression, Extra trees algorithm, Gradient boosted algorithm and Baseline algorithm) to choose the best algorithm suitable for this specific data.

In that function, the attributes in that dataset were grouped according to their similarity with the help of agglomerative traditional hierarchical clustering algorithm that is embedded within the heatmap function. In other words, since clustering is performed for rows and columns, then, the attributes and values that are similar to each other were grouped close to each other in one cluster.

	mae	rmse
Linear Regression	3.48512	4.4326
ElasticNet Regression	3.60805	4.57327
Random Forest	3.72601	4.61621
Extra Trees	3.7797	4.77882
SVM	3.54927	4.58147
Gradient Boosted	3.57244	4.50059
Baseline	3.78788	4.82523

Fig 1.7

 DOI : 10.35291/2454-9150.2021.0265

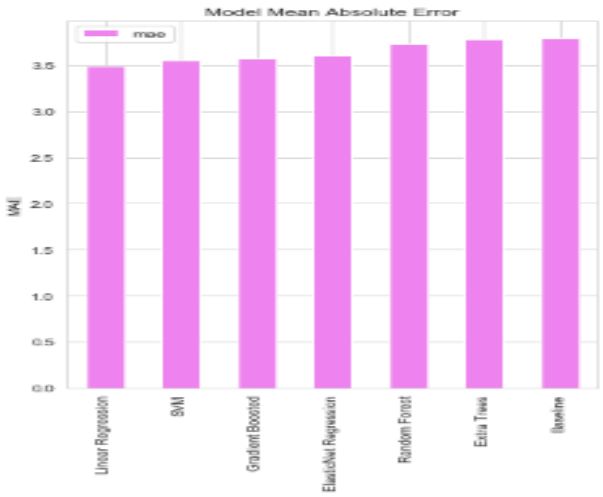


Fig 1.8

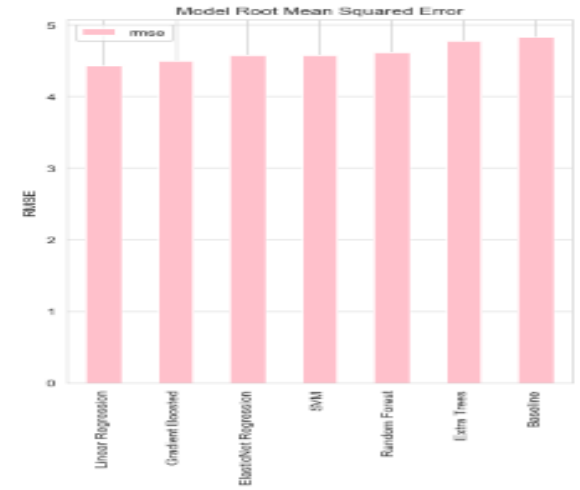


Fig 1.9

The extracted key indicators, which was extracted from the visualization analysis, were fed in the five chosen classification algorithms. But, it's worth mentioning that since the chosen classification algorithms have the capability to train two different attributes' types, i.e. nominal and numeric, both were tried and trained. Then, the accuracy results were evaluated to see which variable type can work efficiently with each classification algorithm in training the datasets of interest.

The predictions rates, in comparison with the baseline, are 76.3% and 69.7% for dissertation grade and all courses grade class, in respectively.

At a certain instance we can see the algorithm procedures as same but the randomness of the data may vary the accuracy. As we check the recall score for every algorithm we can see that recall score=100. If it is not 100 then the classification becomes wrong. The priority is high to the accuracy but, we have to consider the recall score and auc_roc scores. They play an important role in the prediction. The algorithm predicts but, it is predicted correctly or not tell by recall score.

As the data is biased we can see the accuracy is very low due to the data provided. The data has only 20% of exited so we can't train a model which works more accurately.

As we can see the auc_roc scores are between 50-60 which is the acceptance range.

This score helps us to accept the algorithm to predict. The recall score is more important than the accuracy score. Theprediction becomes more accurate when we train more data which is preprocessed. The data must consist of equal exited and non exited rows. The data is totally biased as we can see in the graphs above.

To remove the biased nature by allocating the data to all other ranges. The hidden patterns tell about the chain resembling the exited nature. As we train the data with train split 0.2 also effect the accuracy but, the selection of algorithm would be same.

For certain purpose we may encrypt the data then we may face some problems. The encrypted data is complex and varies with floating values then, the accuracy may differ every time but, final selection is same. As complex the data we cant handle or we can't decrease the comparison between the attributes. The data frame may have the values which we cant enter in the prediction box.

However, in as much as the model has a high accuracy, the final grades of students may go wrong. This could be improved by providing retraining the model with more data over time while in the meantime working with the model to save some that would improve by the students performance in exams.

VII. CONCLUSION &FUTURESCOPE

Predicting students' performance for post graduate study is important for any educational institutions. It is important especially, for those who are aiming to give students opportunities in doing something useful in their field of study, and those who are aiming to well manage the needed teaching resources for excellent learning experiences, like the British University. The British University in Dubai is a start-up research-based institute which aims to improve its reputation and ranking by selecting high performing students to engage them in solving real world issues. So, predicting distinguished students is an urgent desire. Additionally, knowing students' performance in each course beforehand is a main requirement in order to help at risk students by mitigating the challenges that they are facing in their learning journeys and helping them excel in the learning process.

The support vector machine classifier with radial kernel was the one which proved its efficiency (among the rest of classifiers) in predicting students' performance in all courses' grades, including their dissertation projects' grade. The main reason that may be attributed to that classifier's success is the model training method that its used, which relies only on a few data points or samples (those which are very close to the hyperplane) to build its classification model.

 DOI : 10.35291/2454-9150.2021.0265

Only students administration records to form the classification models, ignoring by that other variables that could affect students' learning outcomes, like: attendance, instructor course delivery, and many others. That was because the main focus in this project was to explore the feasibility of utilizing from small dataset size in predicting student performance and to shed the light on the importance of visualization and dendrogram in identifying valuablepredictors.

REFERENCES:

[1] Asif, R., Merceron, A., Ali, S. A., &Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers in Education, 171-194.

2]Brodersen, K. H., Ong, C. S., Stephan, K. E., &Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. In Proceedings - international conference on pattern recognition.

[3]Byrt, T., Bishop, J., & Carlin (1990). Prevalence adjusted bias. Journal of Clinical Epidemiology.

[4]Comendador, B. E. V., Rabago, L. W., &Tanguilig, B. T. (2016). An educational model based on knowledge discovery in databases (KDD) to predict learner's behavior using classification techniques. In 2016 IEEE Int. Conf. Signal process. Commun.Comput.

[5]Fleiss, J. L., & Paik, M. C. (2003). The measurement of interrater agreement, in statistical methods for rates and proportions.In Statistical methods for rates and proportions.

[6]*Litman, D. J., & Forbes-Riley, K. (2004). Predicting student emotions in computer-human tutoring dialogues.In Proceedings of the 42nd annual meeting on Association for Computational Linguistics - ACL '04.*

[7] Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of Statistical Software Ozden Gur Ali and UmutArıt urk. Dynamic churn prediction framework with more effective use of rare event data: The case of private banking. Expert Systems with Applications, 41(17):7889–7903, 2014.

[8] Amal M Almana, Mehmet SabihAksoy, and RasheedAlzahrani. A survey on data mining techniques in customer churn analysis for telecom industry. Journal of Engineering Research and Applications, 4(5):165–171, 2014.

[9] Tianqi Chen and Carlos Guestrin.Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acmsigkdd international conference on knowledge discovery and data mining, pages 785–794. ACM, 2016.

[10 Charles Chu, GuandongXu, James Brownlow, and Bin Fu. Deployment of churn prediction model in financial services industry. In Behavioral, Economic and Socio-cultural Computing (BESC), 2016 International Conference on, pages 1–

2. IEEE, 2016.

[11] KristofCoussement and Dirk Van den Poel. Integrating the voice of customers through call center emails into a decision support system for churn prediction. Information & Management, 45(3):164–174, 2008.

[12] Mohammed Abdul HaqueFarquad, Vadlamani Ravi, and S BapiRaju. Churn prediction using comprehensible support vector machine: An analytical crmap-plication. Applied Soft Computing, 19:31–40, 2014.

[13] Mohammed Hassouna, Ali Tarhini, Tariq Elyas, and Mohammad Saeed AbouTrab.Customer churn in mobile markets a comparison of techniques. arXivpreprint arXiv:1607.07792, 2016.

[14] Yiqing Huang, Fangzhou Zhu, Mingxuan Yuan, Ke Deng, Yanhua Li, Bing Ni, Wenyuan Dai, Qiang Yang, and JiaZeng. Telco churn prediction with big data. In Proceedings of the 2015 ACM SIGMOD international conference on management of data, pages 607–618. ACM, 2015.

[15] AdemKarahoca, OytunBilgen, and DilekKarahoca. Churn management of e-banking customers by fuzzy ahp. In Handbook of Research on Financial and Banking Crisis Prediction through Early Warning Systems, pages 155–172. IGI Global, 2016.

[16] Abbas Keramati, HajarGhaneei, and Seyed Mohammad Mirmohammadi. De-veloping a prediction model for customer churn from electronic banking services

using data mining. Financial Innovation, 2(1):10, 2016.

[17] Gordon S Linoff and Michael JA Berry. Data mining techniques: for marketing, sales, and customer relationship management. John Wiley & Sons, 2011.

..

 DOI : 10.35291/2454-9150.2021.0265