ixed-Ability in EFL Crowded Classes:
(Problems and Solutions)

Dr. Yasir Mohammed Haj Hamad

Assistant Professor-Karary University-College of Computer Science and Information Technology

المستخلص:

أصبح التنبؤ بأداء الطلاب شائعا في التنقيب عن البيانات التعليمية يتم استخدامه لتحسين أداء الطلاب وكذلك تحسين جودة المؤسسة في هذه الورقة يتم عمل تحليل للتنبؤ بتقديرات الطلاب في السنة النهائية بناء على المعدل التراكمي للطلاب للسنوات الأولى من الدراسة. تم اخذ البيانات من كلية الهندسة لعدد 1841 طالب مع أربعة خصائص،تم استخدام خمسة خوارزميات (Naïve Bayes, Random Forest, SVM, J48, logistic regression) لتحليل البيانات باستخدام برنامج ويكا .

أظهره النتائج ان خوارزمية logistic regression احسن أداء من باقي الخوارزميات التي قد حققت دقة أداء عالي(logistic regression90%,SVM89%,Naïve Bayes88.1%, J4887.2%, Random Forest86.9%) هذا يجعل من عملية التنقيب في البيانات فرصة لتحديد الطلاب الذين يتخرجون بنتائج سئية او لا يتخرجون على الاطلاق بحيث يمكن التدخل المبكر.

يمكن مستقبلا الاستفادة من هذا التحليل باستخدام بيانات بأحجام مختلفة وجامعات مختلفة وتطبيق تقنيات المعالجة المسبقة ومقارنته مع تقنيات لتنقيب البيانات مختلفة

Abstract:

Student performance prediction has become very popular in Educational Data Mining. It is used to improve the performance of students and also improves the quality of the institution In this paper predictive analysis was carried out to determine the class of grades of students in their final year using their GPA for the first three years of study. The data set of students has been taken from the Faculty of Engineering with 1841 instanceand 4 attributes.

Five data mining algorithms (Naïve Bayes, Random Forest, SVM, J48, logistic regression) were considered, by using WEKA tools the results show that logistic regression has the best performance to other classifiers ,the accuracywas achievedby Fivealgorithms (logistic regression90% ,SVM89% ,Naïve Bayes88.1% ,J4887.2% ,Random Forest86.9%) this. data mining creates an opportunity for identifying students that may graduatewith poor results or may not graduate at all, so that early intervention may be deployed. . For future work, this analysis can be further taken forward by using data sets from different size and universities and applying data pre-processing techniques and will be compared and analyzed with other data mining techniques.

Keywords: Education, Classification, Prediction, Random Forest, Educational Data Mining, Prediction, Classification Algorithms.

1. INTRODUCTION

Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students, and the settings which they learn in.

Whether educational data is taken from students' use of interactive learning environments, computer-supported collaborative learning, or administrative data from schools and universities, it often has multiple levels of meaningful hierarchy, which often need to be determined by properties of the data itself, rather than in advance. Issues of time, sequence, and context also play important roles in the study of educational data. (1).

In the first and second year of an engineering program, students are exposed to knowledge on sciences and basic introduction to engineering as a continuum of their secondary school education, and as an introduction to general engineering. In the third year, the curriculum is more focused on the core discipline of each engineering student, that is, electrical engineering, civil engineering, and so forth. By the end of the third year, engineering students are already grounded in the basics of their profession. The academic performance of engineering students from their first year to the third year is very vital in terms of acquisition of foundational knowledge, and its impact on their final graduation Cumulative Grade Point Average (CGPA). It is often said that beyond the third year it is very challenging for a student to move from the current class of grade (first class - 1st, second class upper division – 2|1, second class lower division – 2|2, and third class – 3rd) to a higher one due to the nature of academic courses at fourth year and fifth year which are more robust and touch core foundation of engineering disciplines.

Classification is an important task in data mining where student's performance can be predicted using a particular algorithm. Various classification algorithms can be used for the prediction like Naïve Bayes, Logistic Regression, Random Forest, J48, CART, Multi-Layer Perceptron etc. In these paper five algorithms, "Naive Bayes", "Logistic Regression", "Support Vector Machine", "J48" and "Random Forest" have been compared based on the classification accuracy and other comparison metrics. To predict the final degree of students using GPA in first and second and three year of an engineering program.

This paper will be used to reduce the low performance of graduate students by predicting student's performance prior to their arrival to the fourth year and student who are in need of special attention.

This paper organize with five section, section one Introduction section two background about EDM data mining and classifier three describes the related work section four experiment & results explain the used experiment methodology and describes the experiment tool and we used data set, algorithm evaluation, evaluation the results and comparison between the results of algorithms finally conclusion and recondition .

2. DATA MINING PREDICTION USING CLASSIFIERS

2.1 Classification

In the field of Data Mining Classification is a main and most important technique. Class and category of a value is identified using classification based on the previously categorized values. Some of the importantclassification techniques are discussed.

2.2 Classification Algorithms used for Prediction

Naive Bayes Algorithm: Naive Bayes Algorithm is considered as very efficient and easy algorithm.The classification rate is considerably very high and most of the cases it predicts accurate results. This algorithmproduces good results only when the data set is very large (2).

Random Forest Algorithm: To improve the predictive accuracy, Random Forest Algorith uses average values of the model. It is a meta-estimator, so it fits a number of decision trees on various sub samples of datasets. It also controls over fittings. The original sample size is always matches the sub-sample size but thesamples are drawn with replacement (3).

Neural Network Algorithm: Neural Network Algorithm endeavors to recognize hidden relationships in a set of data through many processes like the way the human brain operates. A Neural Network Algorithm is aseries of algorithms and it produces the biggest result without requiring redesigning the output criteria when itadapt to changing input (4).

J48 Algorithm: The ID3 algorithm's features are extended into J48 Algorithm. J48 has the following additional features such as, it accounts the missing values, decision tree pruning, continuous attribute valueranges, derivation of rules, etc. it is an open source java implementation of the C4.5 Algorithm in WEKA Data Mining tool (5).

C4.5 Algorithm: It is also a decision tree algorithm. Some of the features of C4.5 algorithm is as follows, it can be easily interpreted and very easy to implement. It accepts both continuous and discrete values.Some of the limitations are, when a small deviation in data can lead a completely different decision tree. Itcannot work with small dataset (2).

I. CLASSIFICATION ACCURACY

Accuracy is defined as the proportion of correct classification from overall number of cases and it depends on confusion matrix. Table 2 shows the confusion matrix that illustrates the number of correct and incorrect predictions made by the classification model compared to the actual value.

1. Correctly classified instance:

The correctly classified instance show the percentage of test instance that were correctly and in correctly classified the percentage of correctly classified instances is often called accuracy or sample accuracy

Kappa statistics:

Kappa is chance –corrected measure of agreement between the classification and true classes.

Confusion matrix:

A confusion matrix,some times called classification matrix is used to assess the prediction accuracy of model . it measure whether amodel is confused or not, that is whether the model is making mistakes in it predictions . The confusion matrix can be obtained from asset of different scales to compare classifications , including accuracy, which is widely used

The classifiers are evaluated by a confusion matrix which is a combination of four outcomes. In binary classification, the output is either positive or negative. The four different classifications are:

True positives (TP)-accurate positive prediction

False positives (FP)-wrong positives prediction

True negatives (TN)-accurate negative prediction

False negatives (FN)-wrong negative prediction

The effectiveness metrics for classifier used in the research are:-

Precision (p):-

$$\text{Precision}=\frac{TP}{TP+FP}\frac{TP}{TP+FP}$$

Number of true positives classifications divided by the sum of true positives and false positive classifications.

- Recall(R):

$$\text{Recall} = \frac{TP}{TP + FN}$$

i.e number of true positives classifications divided by the sum of true positive and false negative classifications .

F1-SCORE-

F1-score is the harmonic mean of precision and recall

$$\mathrm{F1-score} = \frac{2 * \mathrm{P} * \mathrm{R}}{(\mathrm{P} + \mathrm{R})}$$

Accuracy -

Accuracy is measured by dividing the number of correctly classified instances by the total number of instances.

$$accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}$$

$$\mathrm{Error}\ rate = \frac{F_P + F_N}{T_P + F_P + T_N + F_N}$$

Mean Absolute Error (MAE):-

MAE measures the average magnitude of errors in asset of prediction . it is the summation of the differences between predicted and observation divided by the total number of test samples.

$$\mathrm{MA}E = \frac{1}{\mathrm{n}}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

Root Mean Square Error (RMSE):-

It is the square root of the summation of the squared differences between predicted and actual observations, divided by the number of total test samples.

$$\mathrm{RM}SE = \sqrt{\frac{1}{\mathrm{n}}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

ROC curve:

It is another way to evaluate the performance of the classification (12) where FP values are represented on the y axis and TP values on the y axis

$$\mathrm{T_{PR}} = \frac{T_P}{T_P + F_N}$$

$$F_{PR} = \frac{F_P}{T_N + F_P}$$

Area under curve (AUC):

Another utility called (area under the curve) helps analyze the overall performance of the classification and the ideal classification has AUC.

The ROC is a good visualization tool to identifying the performance of classifier, we times need a numerical value for comparison purpose.

3. LITERATURE REVIEW

Many research studies have been done in educational data mining to predict the students' performance

In (7), the final CGPA of students was predictedusing multiple linear regression and correlation to analyse the yearly GPA, and various inferential statistics were developed. The study determined the correlationbetween the first-year result and the final-year result of the student. With the aidof a regression plot, the students' GPA for the five years of study was fitted using multiple linear regressions in order to explain how the GPA for each year contributedto the variations in the final CGPA of the students at graduation.

In (8) features such as student attendance, average scores, relevant course data, the level of student participation in class etc. were deployed in a data mining model for predicting the performance of 908 students.

In (9) a decision tree modelwas applied to predict the probability of failure of 1,547 students such that relevant knowledge can be acquired that will enable the management team to be able to deploy adequate and early intervention. In the study, the student grades were classified into five categories, and these are: excellent, very good, good, acceptable and fail. Ten input features that include the student's department, high school grades, level of participation in class, attendance, midterm scores, lab reports, homework grades, seminar score, completion of assignments and the overall grades were applied in the decision tree model developed

In (10) by using decision tree classifiers, the likelihood of a student to drop out of an institution was predicted through educational data mining.

In (11), association, classification, clustering and outlier detection data mining techniques were applied to analyse 3,314 graduate student performance records over a fifteen-year period. The dataset was analysed using Rule Induction, Naïve Bayesian classifier, K-Means clustering algorithm followed by density-based and distance-based outlier detection methods. 18 attributes of the student dataset were considered, and only 6 attributes: matriculation GPA, gender, specialty of the students, the city of the student, the grade and the type of secondary school attended were selected for the data mining analysis. The remaining 12 attributes were dropped due to their large variances and because some of the attributes are personal information that did not provide useful knowledge.

In (12) The unsupervised clustering analysis performed, identified four unique clusters in the dataset using k-means algorithm. Data mining method was applied by to evaluate student data towards identifying the key attributes that influence the academic performance of students. This provides an opportunity for improving the quality of higher education.

In (13), data mining technique was Applied to analyse student data at a Bulgarian university. The student dataset that was analysed, contained the personal and pre-admission attributes of each student. The Decision Tree Classifiers (J48), k-Nearest Neighbour, Bayesian, Naïve Bayes classifiers,the OneR, and the JRip Rule learners were applied to extract knowledge from the student dataset, and accuracy of 52e67% was achieved. The result showed that the number of courses failed in the first academic year and the admission score of the student are two major features among the very influential features in the classification analysis.

In (14) the authors used WEKA data mining software for the prediction of final student mark based on parameters in two different datasets. Each dataset contains information about different students from one college course in the past fourth semesters. The IBK shows the best accuracy among other classifiers

In (15) the authors represents a study that will be helped to the students and the teachers to improve the result of the students who are at the risk of failure. Information's like Attendance, Seminar and assignment marks were collected from the student's previous database, to predict the performance at the end of the semester. The authors used Naïve Bayes classification algorithm that shows a highest accuracy compared to other classification algorithms.

The researchersin (16) conducted a comparative research to test multiple decision tree algorithms on an educationaldataset to classify the educational performance of students. The study mainly focuses on selecting the best decision tree algorithm from among mostly used decision tree algorithms, and provides a benchmark to each one of them and found out that the Classification and Regression Tree method worked better on the tested dataset, which was selected based on the produced accuracy and precision using 10-fold cross validations

Researchers in (17) provided an overview on the data mining techniques that have been used to predict students' performance and also it focused on how the prediction algorithm can be used to identify the most important attributes in a student's data. Under the classification techniques, Neural Network and Decision Tree are the two methods highly used by the researchers for predicting students' performance.

In (18), predictive analysis was carried out to determine the extent to which the fifth year and final Cumulative Grade Point Average (CGPA) of engineering students in a Nigerian University can be determined using the program of study, the year of entry and the Grade Point Average (GPA) for the first three years of study as inputs into a Konstanz Information Miner (KNIME) based data mining model. Six data mining algorithms were considered, and a maximum accuracy of89.15% was achieved. The result was verified using both linear and pure quadratic regression models, and R2 values of 0.955 and 0.957 were recorded for both cases. This creates an opportunity for identifying students that may graduate with poor results or may not graduate at all, so that early intervention may be deployed.

In (19) analyze and evaluate the university students' perfore mance by applying different data mining classification techniques by using WEKA tool. The highest accuracy of classifier algorithms depends on the size and nature of the data. Five classifiers are used NaiveBayes, Bayesian Network, ID3, J48 and Neural Network Different performance measures are used to compare the results between these classifiers. The results show that Bayesian Network classifier has the highest accuracy among the other classifiers.

In (20) used J48, PART, Random Forest and Bayes classifiers to predict students" end semester grades on a data set of 300 students from different colleges and found that Random Forest classification algorithm gives the best results based on accuracy and classifier errors .

In (21)used Bayesian Network, J48, Naive Bayes, ID3, J48 and Neural Network classifiers to analyse and evaluate students" performance grades and found that Bayesian Network classifier has the highest accuracy among all classifiers. They concluded that the performance of the students of a university can be best classified using Bayesian Network classification methods. The dependency among random variables is depicted by using directed acyclic graphs, where the nodes in the graph represent the random variables. The dependency of random variables is depicted when a connection exists between a node and an arc .

In (22)compared the results of decision tree classifier and Naïve Bayes algorithm and found that „decision tree" gives better results than „Naïve Bayes" algorithm. "The benefit of „decision tree" is that it is easy to understand and interpret. The decision tree gives good performance with both numerical and categorical variables. Thus, it is one of the very powerful and widely used classifiers. WEKA uses J48 classifier to implement the C4.5 decision tree (10).

In (23)worked on Random Forest classifier and found that Random forest classifier reduce bias, variance and overfitting. Hence, it is very accurate as well as robust. It merges various decision trees together to give a better prediction model (12).

In (24)worked on the recommendation agents that watches the student activities and suggests some actions that will be beneficial for the students (13). After studying the various researches, that have been done for doing predictive analysis through different educational data mining techniques, the authors found the following six classification algorithms: „Naïve Bayes", „Logistic regression", „Multilayer Perceptron", „Support Vector Machine", „J48" and „Random Forest" to be the most promising classifiers to build a model for predicting whether a students will have a reappear in a course or not.

4. Experiments and Results

In this section, a relative report on the technique of extracting data from the classification and predicting the Students performancewill be applied based on the Grade Point Average, of first three year of student, and then predict student performance based on first and second and third year before they reach the four year and choice the best classifier dependent on Accuracy, Error rate, F - measure, exactness and review. Precision.

4.1 Evaluation Metrics

1. Precision

Precision is a ratio of true positive tuples and all positive tuples in a dataset .

2. Recall

Recall is a ratio of true positive tuples against positive and negative tuples

3. F-Measure

F - Measure is also called as F - Score. F - Measure is a mean of precision and recall. F- Measure value varies from 0 to 1. If the value of F-Measure is higher, then it is said to be a better classifier.

4. Accuracy

The classifiers accuracy is an important metric for evaluation. It is a ratio of positive tuples and negative tuples against all the tuples.

5. Error Rate

The error rate is an essential measure for evaluation. Lower error rate is said to be a better classifier. Error rate determines the error between the prediction and actual

4.2 Data set:

The data set name is full data .CSV this consist of the following 4 feature.(first year GPA, second year GPA ,third year GPA,-class of degree) From University of Science & Technology Faculty of Engineering with 1841 instance see figure 4.1Implementation algorithms of classification by using WEKA tools.

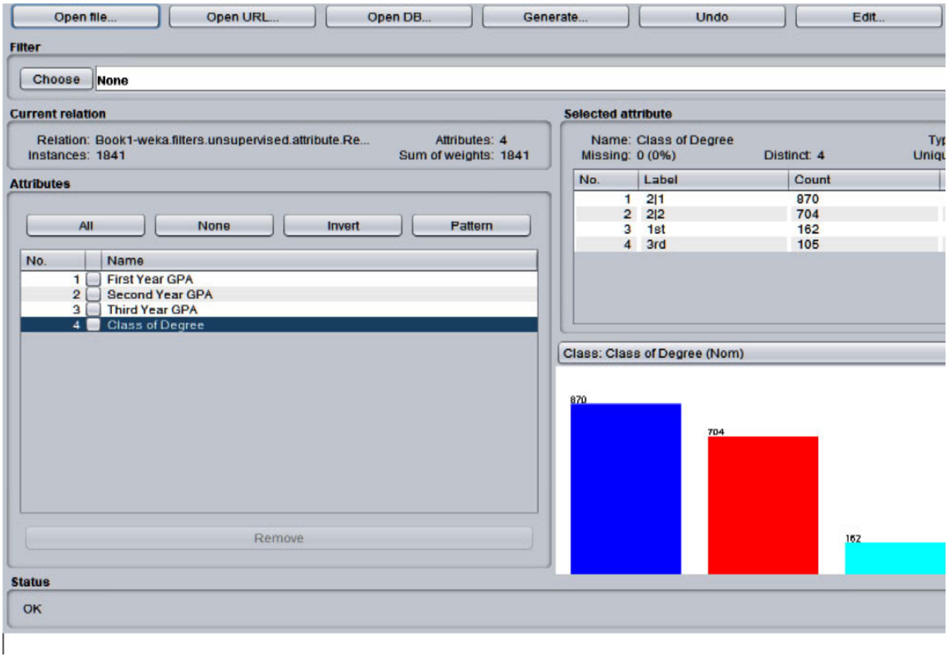

Figure.1.data set

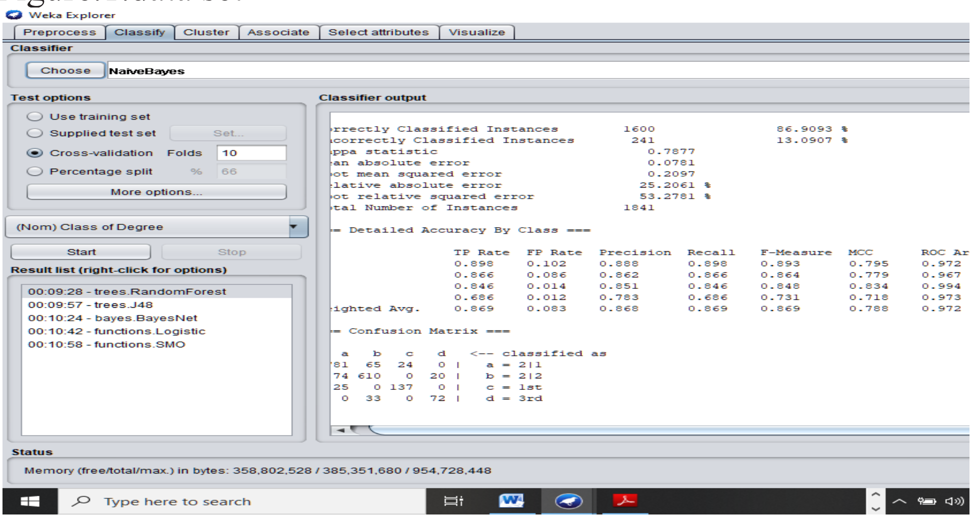

Figure.2.Result of classification model using randomforst algorithm

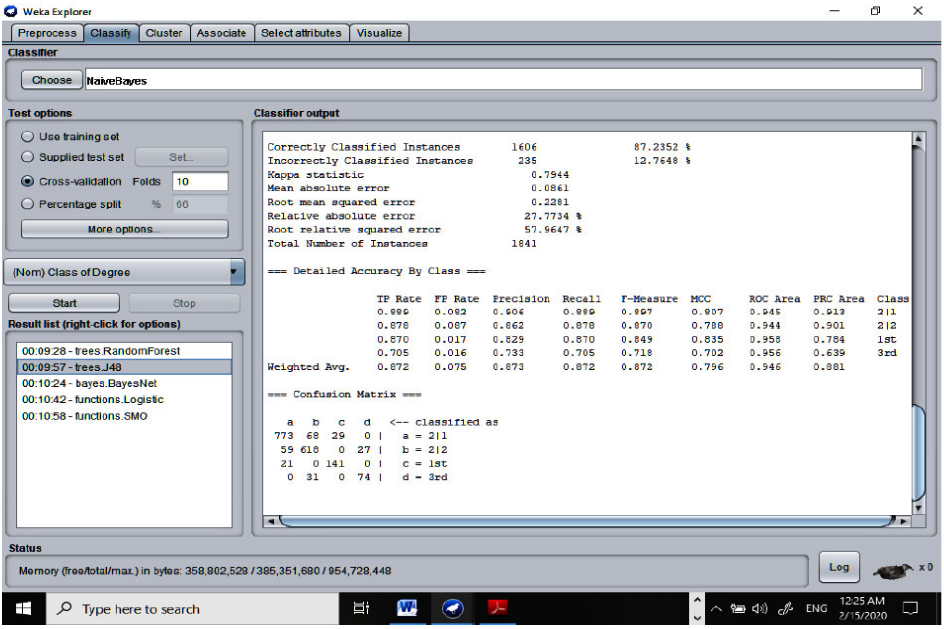

Figure.3.Result of classification model tree j48 algorithm

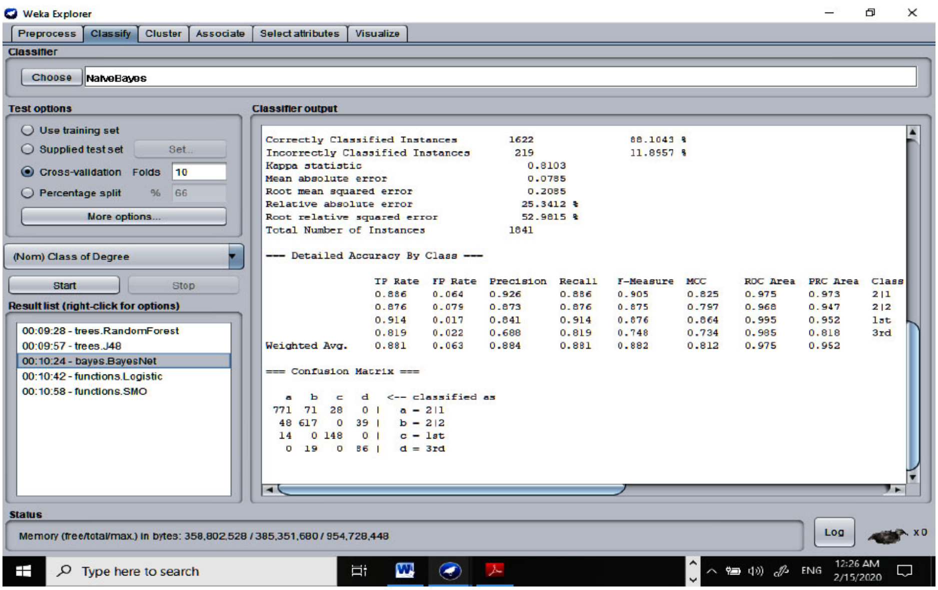

Figure.4.Result of classification model using Naïve Bayes algorithm

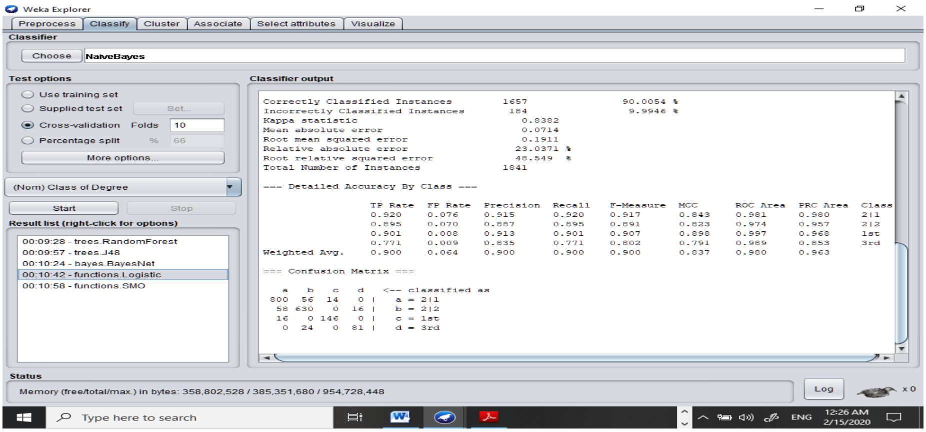

Figure.5.classification model using logistic algorithm.

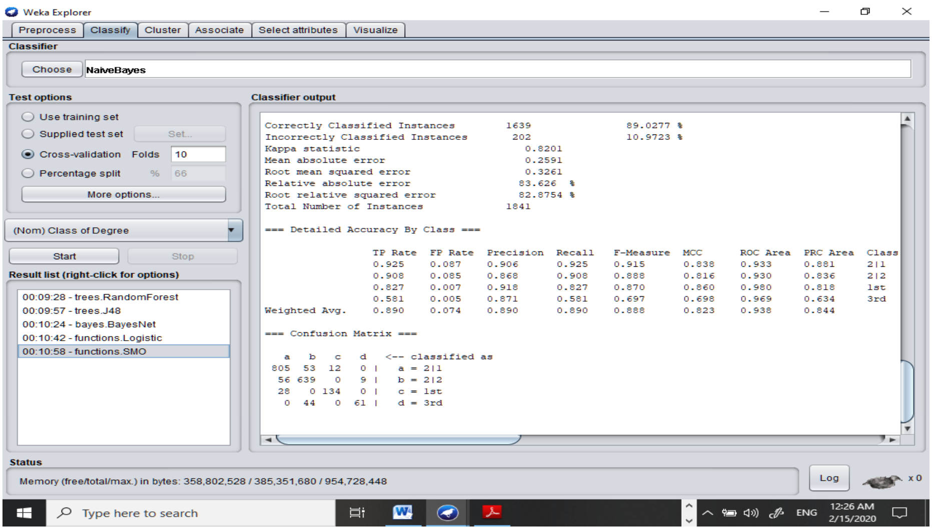

Figure.6.classification model using SVM algorithm.

RESULTS DISCUSSION

The experimental results discussion has done on selecting 1841 instance. Five selected classification algorithms were used; Randomforst, J48, Naive Bayes, SVM, J48 and logistic and each one has its own characteristics to classify the data set. Table 1 shows performance results of all classifiers by using WEKA, and Figure 7 shows the accuracy performance of classification techniquesand Table 2 shows Error measures in weakof All Classifiers

Table1Comparison for Accuracy of All Classifiers

Criteria	classifier				
	Ran-dom-forst	(J48)	Naïve Bays	SVM	logistic
Correctly classified instance	1600	1606	1622	1639	1657
Incorrectly classified instance	241	235	219	202	184
(%) Accuracy	% 86.9	% 87.2	% 88.1	% 89	% 90

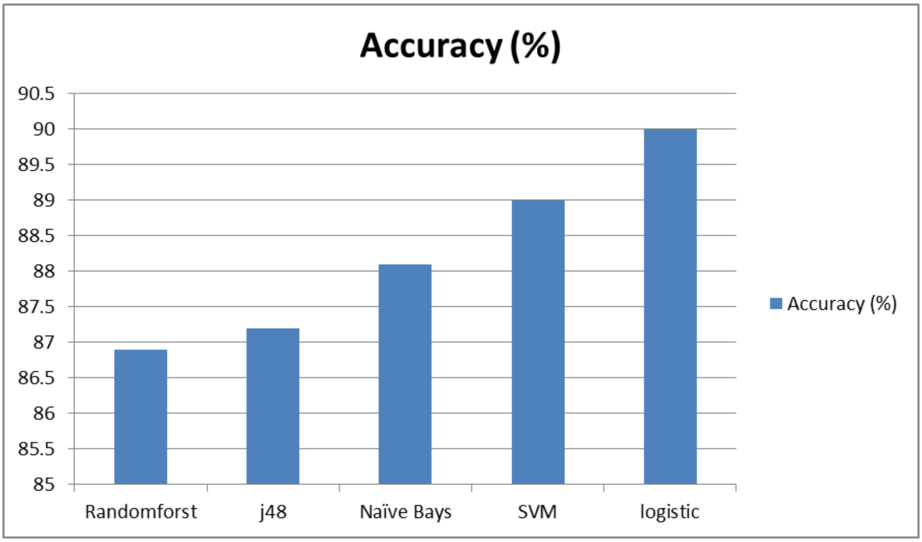

Figure.7.Classifiers Accuracy Performance

In table 1, the logistic regression classifier has more correctly classified instances than other classifiers, which is usually referred to the best accuracy model. The graphical representation in Figure 7 shows that the best classifier of students' performance based on their dataset is the logistic regression classifiers. In the result, logistic regression has an efficient classification among other classifiers.

Table 2 Error measures in weakof All Classifiers

Criteria	classifier				
	Random-forst	(J48)	Naïve Bays	SVM	logistic
Kappa statistic	0.7877	0.7944	0.8103	0.820	0.8382
Mean absolute error	0.0781	0.0861	0.0785	0.259	0.0714
Root mean squared error (RMSE)	0.2097	0.2281	0.2085	0.326	0.1911
Relative absolute error (RRSE)	% 25.2061	% 27.7734	25.3412 %	% 83.62	23.0371 %
Root relative squared error	% 53.2781	% 57.9647	52.9815 %	82.875%	48.549 %

In Table 2. Represents the error measures of all classifiers, it shows that logistic regression has a minimum error among other classifiers.

Table3 Weighted average of class label accuracy

classifier	TP Rate	FP Rate	Preci-sion	Recall	F-Mea-sure	MCC	ROC Area	PRC Area	Class
Random-forst	0.869	0.083	0.868	0.869	0.869	0.788	0.972	0.945	Weighted Age
(J48)	0.872	0.075	0.873	0.872	0.872	0.796	0.946	0.881	1|2
Naïve Bays	0.881	0.063	0.884	0.881	0.882	0.812	0.975	0.952	2|2
SVM	0.890	0.074	0.890	0.890	0.888	0.823	0.938	0.844	1st
logistic	0.900	0.064	0.900	0.900	0.900	0.837	0.980	0.963	3rd

Table 3 shows the performance accuracy of the five classifiers based on different classification metrics. These metrics are; (TP), (FP), Precision, Recall and F-measure measure are very important to determine the classifiers based on the accuracy. These metrics shows that logistic regression classifier performs better than other classifiers.

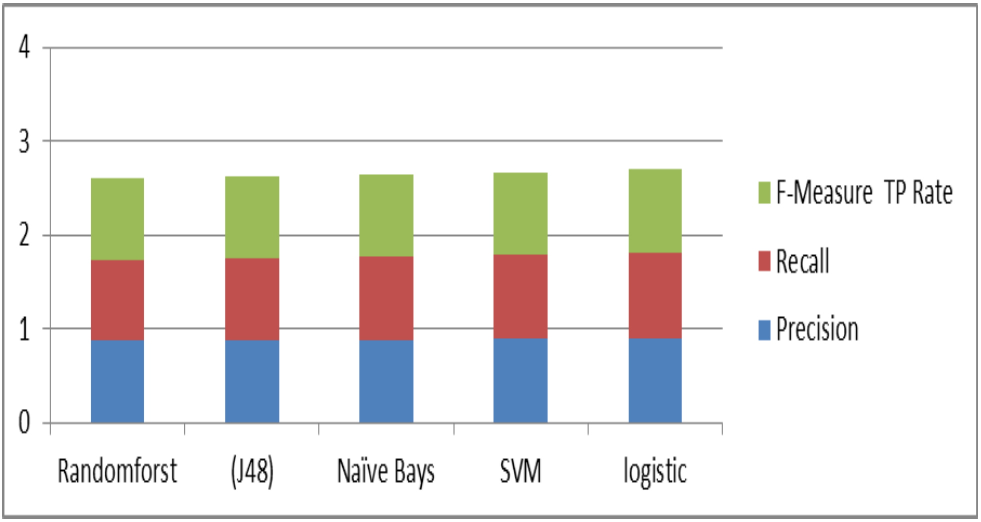

Figure.8.Classifiers Performance Metrics

In Figure 8 Precision, Recall and F-measures analyzed among all classifiers. It shows that the weighted average of logistic regression outperforms other classifiers.

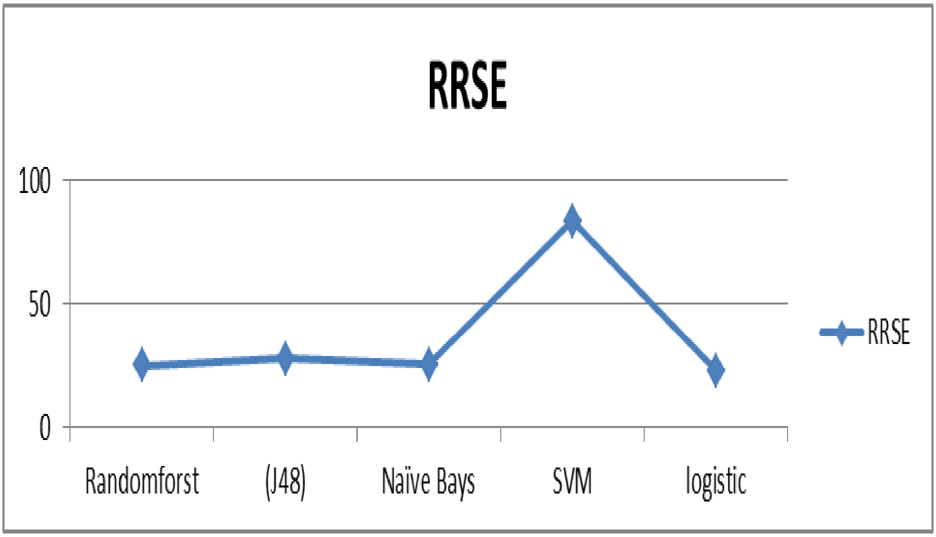

Figure.9.RMSE Metrics

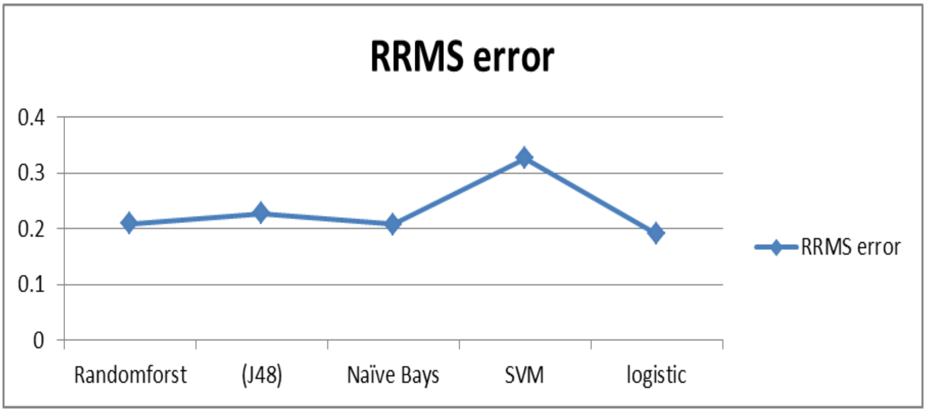

Figure.10. RMSE Metrics

In the result, logistic regression has an efficient classification among other classifiers.

Figure 9 and Figure 10 below show the analysis of all classifiers based on RMSE and RRSE. It shows that logistic regression has minimum error values among other classifiers.

In Figure 11 (TP), (FP metrics shows that logistic regression classifier performs better than other classifiers.

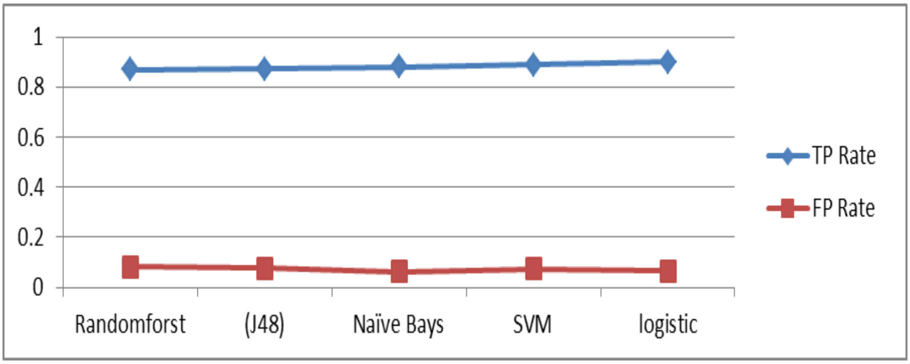

Figure.11. Performance Metrics TP/FP Rate

5. CONCLUSION

Student performance prediction has become very popular in Educational Data Mining. It is used toimprove the performance of students and also improves the quality of the institution. In EDM, ClassificationAlgorithms are used to predict the future results. Data set contains of 1841 instance and four attributes to predictive analysis to determine the class of grades of students in their final year using their GPA for the first three years of study using weka tools and Five classifiers are used Naïve Bayes, Random Forest, SVM, J48, logistic regression and the comparisons are made based on the accuracy among these classifiers and different error measures are used to determine the best classifier. Experiments results show that logistic regression has the best performance other classifiers. For future work, this analysis can be further taken forward by using data sets from different size and universities and also applying data pre-processing techniques and will be compared and analyzed with other data mining techniques.

and as an alternative to running five data mining algorithms separately as implemented in this study, the weka could be modified to incorporate the five data mining algorithms together in a model using a voting system such that the benefits of each algorithm can be combined.

Referance:

(1) http://educationaldatamining.org 26-6-2020.

(2) Sagardeep Roy, Anchal Garg "Analyzing Performance of Students by Using Data Mining Techniques- A Literature Survey" 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)GLA University, Mathura, Oct 26-28, 2017.

(3) https://www.analyticsindiamag.com/7-types-classification-algorithms 1-1-2021.

(4) https://www.investopedia.com/terms/n/neuralnetwork.asp 1-1-2021.

(5) Gaganjot Kaur, Amit Chhabra " Improved J48 Classification Algorithm for the Prediction of Diabetes" International Journal of Computer Applications (0975 – 8887) Volume 98 – No.22, July 2014.

(6) https://data-flair.training/blogs/classification-algorithms24-6-2019.

(7) Oguntunde, Pelumi, Okagbue, Hilary, Oguntunde, Mooneye Adedoyin,Opanuga, A., 2018. Analysis of the inter-relationship between students' first yearresults and their final graduating grades. Int. J. Adv. Appl. Sci. 5 (10), 1e6.

(8) Bucos, M., Dr_agulescu, B., 2018. Predicting student success using data generated intraditional educational environments. (Article). TEM J. 7 (3), 617e625.

(9) Ahmed, A.B.E.D., Elaraby, I.S., 2014. Data mining: a prediction for student's performanceusing classification method. World J. Comp. Appl. Technol. 2 (2),43e47.

(10) Yadav, S.K., Bharadwaj, B., Pal, S., 2012. Mining Education Data to Predict Student's Retention: A Comparative Study arXiv preprint arXiv:1203.2987.

(11) Tair, M.M.A., El-Halees, A.M., 2012. Mining educational data to improve students'

(1) performance: a case study. Int. J. Inf. Commun. Technol. Res. 2 (2), 140e146.

(12) Al-Radaideh, Q., Al-Shawakfa, E., Al-Najjar, M., 2006. Mining student data usingdecision trees. In: Paper Presented at the the 2006 International Arab Conference onInformation Technology. ACIT'2006.

(13) Kabakchieva, Dorina, 2013. Predicting student performance by using data mining methods for classification. Cybern. Inf. Technol. 13 (1).

(14) I. Milos, S. Petar, V. Mladen and A. Wejdan, Students' success prediction using Weka tool, INFOTEH-JAHORINA Vol. 15, March 2016. 684.

(15) P. Shruthi, B. Chaitra, Student Performance Prediction in Education Sector Using Data Mining, International Journal of Advanced Research in Computer Science and Software Engineering,

Volume 6, Issue 3, March 2016.

(16) S.K Yadav, B. Bharadwaj, and S. Pal, 2012. Data Mining Applications: A Comparative Study for Predicting Student's Performance. International Journal of Innovative Technology & Creative Engineering (ISSN: 2045-711), Vol. 1, No.12, December.

(17) A.Mohamed Shahiria,, W. Husaina , N. Abdul Rashida, "A Review on Predicting Student's Performance using Data Mining Techniques" Procedia Computer Science 72 (2015) 414 – 422, ELSEVIER.

(18) Aderibigbe Israel Adekitan , Odunayo Salau , The impact of engineering students' performance

(19) Hilal Almarabeh و I.J. Modern Education and Computer Science, 2017, 8, 9-15 Published Online August 2017 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijmecs.2017.08.02

(20) Hussain S, Dahan N.A, Ba-Alwi F.M, Ribata N., " Educational Data Mining and Analysis of Students Academic Performance Using WEKA", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 9, Issue 2, pp. 447-459, 2018

(21) Almarabeh, H., "Analysis of Students" Performance by Using Different Data Mining Classifiers" I.J. Modern Education and Computer Science, Vol. 9, Issue 8, pp. 9-15, 2017

(22) Patil, T. and S.S. Sherekar, "Performance Analysis of Naïve

Bayes and J4.8 Classification Algorithm for data classification", International Journal of Computer Science and Applications, Vol. 6, Issue 2, pp. 256-261, 2013

(23) Dekker, G.W., M. Pechenizkiy, and J.M. Vleeshouwers, "Predicting students drop out: A case study", EDM "09-Educational Data Mining 2009: 2nd International Conference on Educational Data Mining, pp. 41-50, 2009

(24) C. Romero and S. Ventura, " Educational data mining: A survey from 1995 to 2005", ELSEVIER, Expert Systems with Applications, Vol. 33, Issue 1, pp. 135–146, 2007 .