Comparasion of Feature Selection Techniques in different Classifiers for predicting Students' Academic Performance

Nilar Aye
University of Computer Studies, Yangon
nilaraye@ucsy.edu.mm

***Abstract*—Teaching evaluation is one of the most vital needs that have to be carried out to determine the quality of education in every academic institution. Higher education is mostly assessed by using the students' grades achieved in the examination. A new emerging research area which provides educational organizations to predict the performance of their students is Educational Data Mining (EDM). In EDM, Feature Selection (FS) is the important task applied in getting better the quality of prediction model for educational data sets. FS algorithm removes the irrelevant or redundant features from the educational repositories and influences the accuracy of classifiers used in various EDM practices. In order to implement the prediction model with better accuracy, it is necessary to apply the good quality of educational data set with the most relevant features. Therefore, the decision depends on suchlike good quality data set can be able to enhance the educational quality by predicting the students' performance. So, it is indispensable to choose carefully which feature selection algorithm is more appropriate with the classifier. For this purpose, this system is implemented to make comparisons between five feature selection techniques and their impact on the fifteen classifiers on the Students' Performance Data set from University of Jordan using WEKA tool. The results of present comparative study effectively support the educational area to recommend the best combination of feature selection method and classification algorithm.**

Keywords— Educational Data Mining (EDM), Feature Selection (FS), classifiers

I. INTRODUCTION

Education is one of the most fundamental factors for the development of a nation. Cultivating outstanding members of the organization depends on the quality of education which is the most required ingredient [9]. Educational institutions provide their students with more competitive education because the number of institutions is increasing speedily [6]. The principal aim of educational institutions is to furnish the quality of education for their students and to capture better the qualification of administrative decisions. One of the key practices to get success in doing the highest quality in the educational institution is by discovering valuable knowledge from the data of education to analyze the most relevant behaviors that may influence the students' performance [11].

Students' performance is a basic factor which is necessary for higher learning institutions. Excellent records of academic achievements make universities high quality [3]. Nowadays, it has been developing many techniques proposed to evaluate the students' performance. Among the proposed techniques, the useful technique which is applied to analyze students' performance is Data Mining [3]. It is a form of discovery of knowledge that can be used for resolving problems in a specific domain [13].

As an educational database is developed by collecting of numerous records related to the students' information, various data mining tools are developed and applied upon this educational database to retrieve required students' information and to detect the invisible relationship [5]. Educational Data Mining (EDM) is one of the practices of Data Mining techniques on the data of education. This technique examines raw data of students received from educational systems and then transforms these data into effective decision for respective educators that could potentially have an important impact on the research area related to the education and application [10]. There are different kinds of well-known data mining tasks related to the educational data mining e.g. classification, clustering, outlier detection, association rules, prediction etc. [5]. Classification is the most familiar and most efficient data mining techniques used to implement a prediction model.

In this paper, the system generates the prediction model that is going not only to be profitable to the teachers, parents, faculties and administrators for reminding the students during their academic year but also for giving advice to the students to facilitate reflection and self-regulation during their study. Therefore, the presented prediction model needs to have the highest performance which greatly relies on the choice of selection of most relevant features from the list of attributes in the students' performance data set. This highest performance can be obtained by means of utilizing various kinds of feature selection techniques on the students' performance data set to generate the most suitable features. The proposed system evaluates the accuracy of the fifteen classifiers loaded with the selected features generated by the applied feature selection algorithms.

II. RELATED WORK

Data mining and technologies related to it have been making actual transformations to the Educational field. In this section, a great research studies on the various Feature Selection algorithms applied in the major Educational Data Mining works are explained and the accuracy results of performance in different combinations of feature selections along with several kinds of classifiers are analyzed.

N. Rachburee and W. Punlumjeak [9] conducted the study of comparative results between Greedy Forward selection, Information Gain Ratio, Chi-Square and mRMR Feature Selections. This study is experimented on the database which contains the first year students' information of University of Technology, Thailand. This database has 15 features. Their system compared the prediction accuracy results obtained by using Artificial Neural Network (ANN), Naïve Bayes, Decision Tree and k-NN and finally they concluded that Greedy Forward selection is able to generate higher results for prediction by deploying Artificial Neural Network (ANN). When this accuracy result is compared with the result of Naïve Bayes, Decision Tree and k-NN, it is better than others.

Ramaswami M., & Bhaskaran R, 2010 [6] have designed a framework to develop a predictive model. They used the predictive variables with the highest influence that were received by utilizing feature selection. The data set needed for this system contains 772 records related to the students and it has 7-class variables. They proved that the present model received more satisfactory accuracy than other predictive models.

Maryam Zaaffar et al. [1] attempted to select the suitable attributes for classification and carried out an analysis of Feature Selection Algorithm's performance. In this system, Six Feature Selection Algorithms were applied. According to the system's results, Principal Component has achieved better results by applying it random forest classifier.

Wattana punlumjeak et al. [1] carried out an experiment which applied feature selection to predict the students' performance. The generated model classified the performance of students on Microsoft Azure platform with a large student data set. They received 90.60% of accuracy using mutual information. They discovered that it was the best result as it combined with Neural Network classifier.

R. Sumitha [8] has presented an approach to make a prediction of students' outcome for a data set with 300 students. In order to design a prediction model, the system used 250 students as training data and 50 students as test data. Attributes selected algorithm along with classifiers proved that the rate of prediction changes from 80% to 98%. Naive Bayes, Multilayer Perception (MLP), Sequential Minimal Optimization, REP tree and J48 were applied. Among them, J48 received 97% of accuracy.

III. DATAMINING DEFINITION AND TECHNIQUES

Data Mining is a research field that deals in extracting valuable information from databases, data warehouses or other information repositories. It is also called Knowledge Discovery in Databases (KDD) [10]. Data Mining is able to highlight meaningful information and provides decision making. Data mining is comprised in various kinds of techniques such as Association Rule Mining, Clustering, Classification, Prediction, etc. In order to create models that are developed to analyze data, classification and prediction techniques are used. And then, these models are used to assess other data [6].

IV. CLASSIFICATION ALGORITHMS

Classification is one of the Data Mining techniques that generate a set of models (or functions) which analyze the features of data and differentiate the respective classes or concepts of the data. And then, generated model is proposed to predict the class of objects which have unknown class label [11].

Current system is designed to get highly performance of classifiers. Different feature selection techniques are deployed on students' performance data set to generate the most suitable features before applying any classifier and it presents how the performance of classification model highly depends on the most relevant features in the data set. Various classifiers can be applied to predict the students' performance. This paper applied fifteen classification algorithms that are: Bayesian Network (BN), Naïve Bayes (NB), NaiveBayesUpdateable (NBU), Multilayer Perception **(**MLP), Simple Logistic (SL), Sequential Minimal Optimization (SMO), Decision Table (DT), OneR, JRip, PART, REPTree (RepT), Random Forest (RF), J48, ZeroR and RandomTree (RT).

V. FEATURE SELECTION METHODS

There may be many irrelevant attributes in data to be mined. So, removing these irrelevant attributes in data is a critical task that must be performed. Most of the mining algorithms are not able to perform adequately on the database which has a lot of features or attributes.

Therefore, feature selection techniques are required to be used before applying any kind of mining algorithm. The main objectives of feature selections are to be more accurate, to increase the accuracy of the applied model and to facilitate more rapidly and more cost-effective models [13].

Unnecessary features can contain noisy data affecting the classifiers' accuracy negatively [7]. In pre-processing step, Feature Selection algorithms assist to choose the subset of suitable features on the data. This selected subset features provide data mining techniques with the better accuracy for prediction and lower the complexity of computation [9]. Feature selection algorithms are categorized depend on different feature evaluation measures. They are filter, wrapper and hybrid models. Feature selection algorithms select the most relevant features of the educational data set. The selected features can be deployed on various classifiers to evaluate the prediction accuracy.

In this presented paper, the research process is on five important Feature Selection algorithms such as CfsSubsetEval, GainRatioAttributeEval, Principal Components, InfoGainAttributeEval and Classifier-AttributeEval. The above mentioned five feature selection techniques are combined with fifteen classifiers to develop the fast and more accurate model to predict the student's performance.

VI. SCOPE AND SAMPLE OF THE STUDY

Students' Performance data set, xAPI-Edu-Data, used here is collected from a Learning Management System Kalboard-360 of University of Jordan, and 16 attributes and 480 student records are stored in that data set [8]. The attributes, their descriptions and possible values are explained in Table I.

VII. METHODOLOGY OF THE STUDY

The proposed system compares the performance of the fifteen classifiers which are combined along with five feature selection techniques on Students' Performance data set from University of Jordan using WEKA tool.

WEKA (Waikato Environment for Knowledge Analysis) is free source software programmed in Java language. It provides different machine learning techniques and data mining tasks to do Data Preprocessing, Classification, Regression, Clustering and Association Rules [2]. WEKA is composed of various machine learning algorithms to solve real world data mining problems. It is able to run on almost any platform. These algorithms can also be employed directly to a data set [6]. In WEKA, there are feature selection methods that consist of two categories: the way by which feature subsets are evaluated known as Feature evaluator and the way by which space of feature subsets are searched known as Search method [4].

TABLE I. LIST OF ATTRIBUTES' DESCRIPTION AND THEIR POSSIBLE VALUES

No.	Features	Description	Possible Values
1	Gender	Sex of students	M, F
2	Nationality	Student nationality	{KW, Lebanon, Egypt, SaudiArabia, USA, Jordan, Venzuela, Iran, Tunis, Syria, Morocco, Iraq, Palestine, Lybia}
3	Place of Birth	Birth place for the student	{KW, Lebanon, Egypt, SaudiArabia, USA, Jordan, Venzuela, Iran, Tunis, Syria, Morocco, Iraq, Palestine, Lybia}
4	StageID (Educational Stages)	Stage student belongs	{Lowerlevel, Middleschool, HighSchool}
5	GradeID	Student' grade	{G-01, G-02, G-03, G-04, G-05, G-06, G-07, G-08, G-09, G-10, G-11, G-12}
6	SectionID	Section name of Student	{A, B, C}
7	Topic	Couse Topic	{Math, English, IT, Arabic, Science, Quran, French, Spanish, History, Biology, Chemistry, Geology}
8	Semester	Semester Name	{First, Second}
9	Relation	Parent responsible for student	{Father, Mother}
10	Raised Hand	Raised Hand on class	Numbers of time
11	VisITedResources	Visited resources	Numbers of time
12	AnnouncementsView	Viewing announcements	Numbers of time
13	Discussion	Student Participation on learning process	Numbers of time
14	ParentAnsweringSurvey	Parent Participation on learning process	{Yes, No}
15	ParentschoolSatisfaction	Parent satisfaction level from school	{Good, Bad}
16	StudentAbsenceDays	Absence days of student	{Above-7, Under-7}
17	Result	Result from mark calculation	{High, Middle, Low}

VIII. MEASURE PERFORMANCE

The various feature selection techniques impact on the performance of the different classifiers. The percentage of tuples allocated in a right class can be used to calculate the classification accuracy. A confusion matrix is a useful tool to determine the performances of each classifier. In Table II, confusion matrix values are identified for each measurement [8]. Equation (1) to (4) refers to Table II and the proposed system uses these equations to evaluate the system's performance.

TABLE II. BASIC CONFUSION MATRIX

		Detected	
		Positive	**Negative**
Actual	**Positive**	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

*a) **Recall**:* Recall is the ratio of total number of true positive to the total number of positives. So,

$$\text{Recall} = \frac{TP}{TP + FN} \quad (1)$$

*b) **Precision**:* Precision is the ratio of total number of true positive to the total number of positive classifications. So,

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

*c) **F-measure**:* F-measure is a harmonic mean of precision and recall. So,

$$F - \text{measure} = 2 * \frac{P * R}{P + R} \quad (3)$$

*d) **Accuracy**:* Accuracy is the ratio of the sum of correct classifications to the total count of cases [12]. So,

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (4)$$

IX. EXPERIMENTAL WORK

Firstly, the researchers collected the information of students related to their performances and then this information was designed in an excel sheet to be able to load in WEKA. And then, this data set was saved as .csv file format to carry out the pre-processing steps like cleaning and integration process on the gathered data. After that, the system converted the cleaned and integrated students' performance data set into .ARFF formats. The proposed system was implemented according to the processes shown in Fig 1.

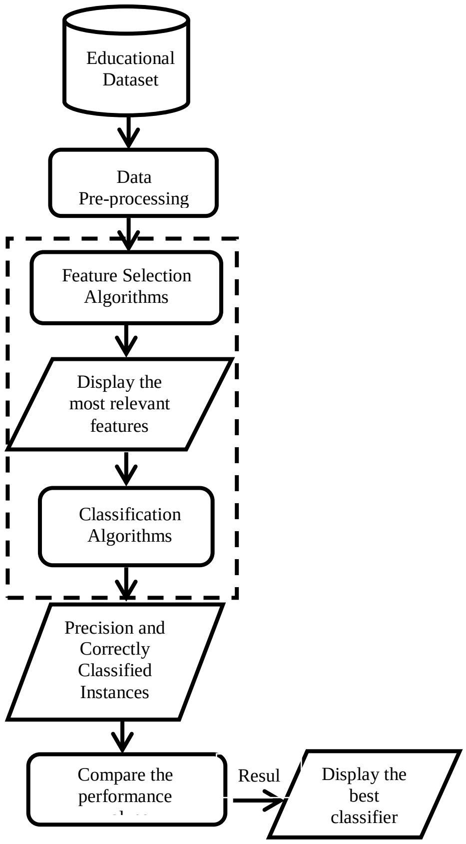

Fig. 1. Proposed Model for the prediction of students' performance

In the first step, the students' performance .ARFF data set was loaded as input into WEKA explorer. In the second step, before the data was mined by different classifiers, five feature selection techniques such as CfsSubsetEval, GainRatioAttributeEval, Principal Components, InfoGainAttributeEval and ClassifierAttributeEval were applied on the input data set to select the feature subsets which contained the most relevant features in order to increase the result of mining accuracy. The Best First search was chosen for the CfsSubsetEval and the Ranker Search method was chosen for GainRatioAttributeEval, Principal Components, InfoGainAttributeEval and ClassifierAttributeEval. After applying feature selection techniques on the input .ARFF format data set, the system generated the most relevant feature subsets illustrated in Table III.

TABLE III. THE LIST OF SELECTED ATTRIBUTE SUBSETS GENERATED BY FIVE FEATURE SELECTION TECHNIQUES

Feature Selection Techniques	Selected Attributes
CfsSubsetEval	1, 9, 10, 11, 12, 14, 16
GainRatioAttributeEval	16, 10, 11, 12, 14, 9, 15
Principal Components	1, 2, 3, 4, 5, 6, 7
InfoGainAttributeEval	11, 16, 10, 12, 14, 2, 9
ClassifierAttributeEval.	16, 15, 6, 5, 4, 3, 2

In the third step, the most relevant features generated by each feature selection technique were chosen from the original data set and the irrelevant attributes were removed from it to apply the different classifiers. After removing each irrelevant attributes, fifteen classifiers were separately deployed on this data set with the most relevant features. All experiments were run by applying a ten-fold cross validation approach. In the fourth step, the system generated the performance values such as Precision, Recall and F-measure. Finally, the system compared the generated performance values and displayed the best classifier.

X. EXPERIMENTAL RESULTS AND DISCUSSION

This proposed system makes a new effort to evaluate the performance of five Feature Selection algorithms using the students' performance data set. The capabilities of these algorithms are measured through Precision, Recall, F-measure and prediction accuracy (correctly classified instances). The harmonic mean of precision and recall is F-measure [9].

The results of the comparative study of fifteen different Classifiers that are carried out against feature subsets generated by the five different feature selection procedures are reported in Table IV to VIII. These tables describe the performance values evaluated by each of the feature selection algorithm through fifteen classifiers. Moreover, each results table consists of four columns such as fs-classifier, Precision, Recall and F-measure values.

TABLE IV. PERFORMANCE MEASURE OF CFSSUBSETEVAL ON DATA SET APPLYING FIFTEEN CLASSIFIERS

fs-classifier	Precision	Recall	F-measure
CSE-NB	0.68	0.679	0.671
CSE-BN	0.72	0.721	0.718
CSE-NBU	0.68	0.679	0.671
CSE-MLP	0.739	0.74	0.739
CSE-SL	0.746	0.746	0.746
CSE-SMO	0.74	0.742	0.741
CSE-DT	0.705	0.704	0.704
CSE-JRip	0.711	0.71	0.71
CSE-OneR	0.621	0.613	0.608
CSE-PART	0.724	0.725	0.724
CSE-DS	0.496	0.521	0.485
CSE-J48	0.746	0.746	0.746
CSE-RF	0.762	0.763	0.762
CSE-RT	0.685	0.685	0.685
CSE-RepT	0.706	0.706	0.706

Table IV indicates the experimental results of the system's measurements for fifteen classifiers with CfsSubsetEval feature selection algorithm using students' performance data set.

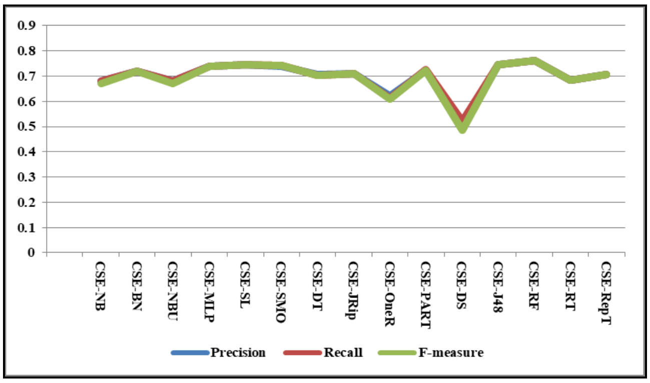

Fig. 2. Graphical illustration of the Performance of CfsSubsetEval

From the above Fig. 2, it is a graphical illustration of the results tested with CfsSubsetEval feature selection algorithm. According to the results illustrated in Table IV and Fig. 2, they describe that the Decision Stump (DS) classifier performed on the input data set with the minimum performance with CfsSubsetEval; Moreover, Random Forest (RF) method generated comparatively better measurements than other classifiers with the same Feature Selection technique and same data set.

The values in Table V and Fig. 3 illustrate the comparison results of classifiers' accuracy values used on the data set of students' performance with feature selection technique, GainRatioAttributeEval. The results prove that the measurements of Precision, Recall and F-measure are relatively low after loading the data set with Decision Stump (DS) classifier. Both J48 and Random Forest (RF) get the same values of accuracy performance which are comparatively increased than other classifiers deploying GainRatioAttributeEval. It is analyzed that J48 classifier is able to perform the best with the highest performance values using among other feature selection techniques on the same dataset.

TABLE V. PERFORMANCE MEASURE OF GAINRATIOATTRIBUTEEVAL ON DATA SET APPLYING FIFTEEN CLASSIFIERS

fs-classifier	Precision	Recall	F-measure
GRA-NB	0.706	0.704	0.698
GRA-BN	0.74	0.74	0.738
GRA-NBU	0.706	0.704	0.698
GRA-MLP	0.733	0.733	0.733
GRA-SL	0.752	0.752	0.751
GRA-SMO	0.76	0.76	0.76
GRA-DT	0.76	0.76	0.76
GRA-JRip	0.736	0.735	0.735
GRA-OneR	0.632	0.625	0.622
GRA-PART	0.735	0.735	0.735
GRA-DS	0.493	0.525	0.47
GRA-J48	0.764	0.765	0.764
GRA-RF	0.764	0.765	0.764
GRA-RT	0.686	0.688	0.686
GRA-RepT	0.723	0.723	0.723

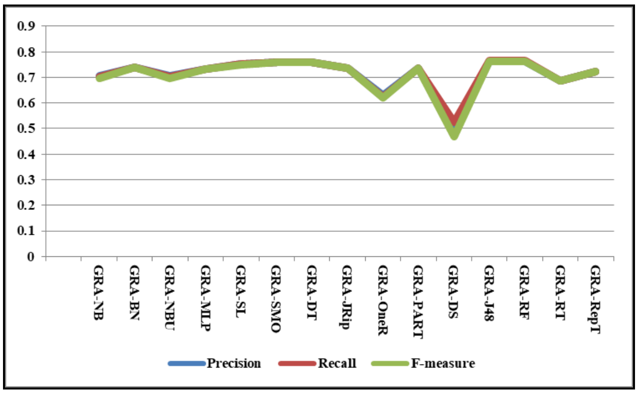

Fig. 3. Graphical illustration of the Performance of GainRatioAttributeEval

Table VI and Fig. 4 exhibit the detail results tested with InfoGainAttributeEval method. These generated accuracy values describe that Random Forest (RF) classifier performed comparatively better in performance by applying InfoGainAttributeEval, however, Decision Stump (DS) depicts poor performance with InfoGainAttributeEval using the same data set. Random Forest (RF) classifier generated the highest level of accuracy through the InfoGainAttributeEval using among other feature selections.

TABLE VI. PERFORMANCE MEASURE OF INFOGAINATTRIBUTEEVAL ON DATA SET APPLYING FIFTEEN CLASSIFIERS

fs-classifier	Precision	Recall	F-measure
InfoGA-NB	0.689	0.688	0.68
InfoGA-BN	0.745	0.746	0.745
InfoGA-NBU	0.689	0.688	0.68
InfoGA-MLP	0.737	0.738	0.736
InfoGA-SL	0.752	0.752	0.752
InfoGA-SMO	0.74	0.74	0.739
InfoGA-DT	0.706	0.704	0.705
InfoGA-JRip	0.746	0.744	0.744
InfoGA-OneR	0.632	0.625	0.622
InfoGA-PART	0.758	0.758	0.758
InfoGA-DS	0.493	0.525	0.47
InfoGA-J48	0.735	0.733	0.734
InfoGA-RF	0.769	0.769	0.768
InfoGA-RT	0.667	0.667	0.667
InfoGA-RepT	0.695	0.694	0.694

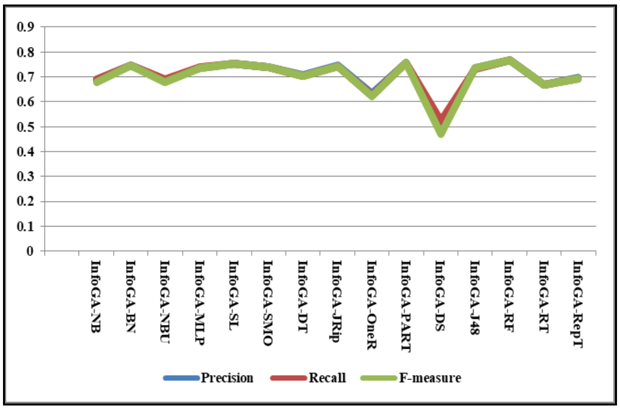

Fig. 4. Graphical illustration of the Performance of InfoGainAttributeEval

Table VII consists of the tested results of selected fifteen classifiers using ClassifierAttributeEval. Fig. 5 is the pictorial illustration of the performance of ClassifierAttributeEval. The values in Table VII and Fig. 5 represent that Sequential Minimal Optimization (SMO) classifier performed comparatively better, while the Decision Stump (DS) classifier got the poor performance. After running the WEKA tool using the Decision Stump (DS) classifier combined with different feature subset selections, the generated accuracy performance values of Decision Stump (DS) classifier along with GainRatioAttributeEval, InfoGainAttributeEval and ClassifierAttributeEval techniques are the same but these measurements are lower than the measurements of it with CfsSubsetEval feature selection technique.

TABLE VII. PERFORMANCE MEASURE OF CLASSIFIERATTRIBUTEEVAL ON DATA SET APPLYING FIFTEEN CLASSIFIERS

fs-classifier	Precision	Recall	F-measure
CFA-NB	0.603	0.6	0.6
CFA-BN	0.597	0.594	0.593
CFA-NBU	0.603	0.6	0.6
CFA-MLP	0.607	0.608	0.607
CFA-SL	0.636	0.638	0.637
CFA-SMO	0.661	0.66	0.658
CFA-DT	0.634	0.633	0.633
CFA-JRip	0.631	0.627	0.614
CFA-OneR	0.493	0.525	0.47
CFA-PART	0.622	0.627	0.622
CFA-DS	0.493	0.525	0.47
CFA-J48	0.65	0.646	0.638
CFA-RF	0.628	0.629	0.629
CFA-RT	0.616	0.617	0.615
CFA-RepT	0.626	0.629	0.627

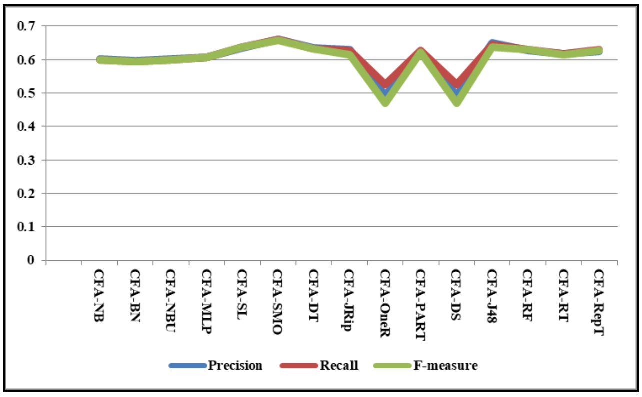

Fig. 5. Graphical illustration of the Performance of ClassifierAttributeEval

When the system is experimented with the selected feature subset generated by Principal Component feature selection algorithm to predict the students' performance, it reports the results of performance values using different classifiers as illustrated in Table VIII and Fig. 6. It is inspected through the results' analysis that the performance values generated by Random Tree (RT) are the best results with Principal Component, however, the Decision Stump (DS) classifier clarifies no value of Precision and F-measure with Principal Component feature selection. Moreover, it is analysed that Recall value of classifier Decision Stump (DS) is the lowest not only among other classifiers but also using other feature selection techniques on the same data set.

Although the performance values of Random Tree (RT) are the highest values using Principal Component feature selection algorithm, these accuracy values are the lowest values among of other feature selection techniques.

TABLE VIII. PERFORMANCE MEASURE OF PRINCIPAL COMPONENT ON DATA SET APPLYING FIFTEEN CLASSIFIERS

fs-classifier	Precision	Recall	F-measure
PRC-NB	0.46	0.454	0.455
PRC-BN	0.453	0.448	0.448
PRC-NBU	0.46	0.454	0.455
PRC-MLP	0.522	0.521	0.52
PRC-SL	0.461	0.463	0.455
PRC-SMO	0.466	0.469	0.448
PRC-DT	0.444	0.446	0.434
PRC-JRip	0.472	0.458	0.382
PRC-OneR	0.476	0.465	0.384
PRC-PART	0.502	0.506	0.499
PRC--DS	0	0.421	0
PRC-J48	0.498	0.498	0.494
PRC-RF	0.519	0.521	0.519
PRC-RT	0.532	0.531	0.526
PRC-RepT	0.503	0.488	0.454

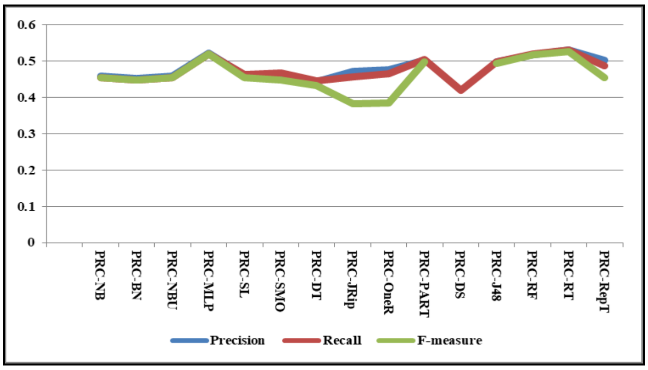

Fig. 6. Graphical illustration of the Performance of Principal Component feature selection

The following Fig. 7 highlights correctly classified instances obtained under fifteen classifiers through five feature selection algorithms by illustrating the graphical representation.

After analyzing the experimental results, it is found that GainRatioAttributeEval feature selection technique enhances the accuracy and the performance of the most classification techniques.

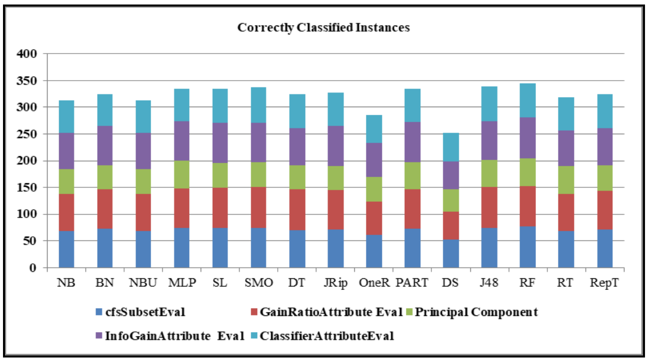

Fig. 7. Graphical illustration of Correctly Classified Instances

XI. CONCLUSION

Predicting academic performances of student is not only one of the key responsibilities of any educational institutions but also a very critical research that to be taken in the field of Educational Data Mining. This provides either teachers or administrators to make proper decisions to enhance the education standard of the students. Feature selection is one of the most prominent factors to perform the data mining tasks successfully through selecting the useful or relevant attributes in the data set. In this system, different classification techniques are combined with each feature selection method to compare the performance measures between feature selection methods and their impact on learning algorithms. Finally, experimental results applying WEKA tools prove that GainRatioAttributeEval feature selection technique provides the classifiers comparatively better results than other feature selection techniques. This experimental result supports the educational institutions to understand which feature selection algorithm influences the performance of a classifier.

REFERENCES

[1] A.Dinesh Kumar, R.Pandi Selvam, V.Palanisamy, "Prediction of Student Performance using Hybrid Classification", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-4, November 2019

[2] Adepu Rajesh, "Data Mining Classification of Teaching Evaluation uisng WEKA", International Joural of Latest Trends in Engineering and Technology, Vol.(9), Issue(1), pp.082-086, DOI:http://dx.doi.org/10.21172/1.91.12, e-ISSN:2278-621X

[3] Amirah Mohamed Shahiria, Wahidah Husain, Nur'aini Abdul Rashid, "A Review on Predicting Student's Performance using Data Mining Techniques", The Third Information Systems International Conference, Procedia Computer Science 72 (2015) 414 – 422

[4] Dr. Kadhim B.S. Al Janabi and Rusul Kadhim, "Data Reduction Techniques: A Comparative Study for Attribute Selection Methods", International Journal of Advanced Computer Science and Technology, ISSN 2249-3123 Volume 8, Number 1 (2018), pp. 1-13, © Research India Publications, http://www.ripublication.com

[5] Humera Shaziya, Raniah Zaheer, G. Kavitha, " Prediction of Students Performance in Semester Exams using a Naïve Bayes Classifer", International Journal of Innovative Research in Science, Engineering and Technology,(An ISO 3297: 2007 Certified Organization), Vol. 4, Issue 10, October 2015.

[6] Jai Ruby, Dr. K. David, "Predicting the Performance of Students in Higher Education Using Data Mining Classification Algorithms - A Case Study", International Journal for esearch in Applied Science & Engineering Technology (IJRASET), Volume 2, Issue XI, November 2014, ISSN: 2321-9653

[7] K. Pavya, Dr. B.Srinivasan, "Feature Selection Techniques in Data Mining: A Study", ISSN: 2455-2631 © June 2017 IJSDR | Volume 2, Issue 6

[8] M. Anoopkumar, A.M.J.Md. Zubair Rahman, "Model of Tuned J48 Classification and Analysis of Performance Prediction in Educational Data Mining", International Journal of Applied Engineering Research ISSN 0973-4562 Volumn 13, Number 20 (2018), pp.14717-14727

[9] Maryam Zaffar, K.S. Savita, Manzoor Ahmed Hashmani, Syed Sajjad Hussain Rizvi, "A Study of Feature Selection Algorihms for Predicting Students Academic Perfromance", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 5, 2018

[10] Parneet Kaur, Manpreet Singh, Gurpreet Singh Josan, " Classification and prediction based data mining algorithms to predict slow learners in education sector", 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

[11] S.J. Ahmed Abu Hammad, "Mining Educational To Analyze Student's Performance", Central European Researchers Journal, Vol.4 Issue 2

[12] Sadiq Hussain, Neama Abdulaziz Dahan, Fadl Mutaher Ba-Alwib, Najoua Ribata, "Educational Data Mining and Analysis of Students' Academic Performance Using WEKA", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 9, No. 2, February 2018, pp. 447~459, ISSN: 2502-4752, DOI: 10.11591/ijeecs.v9.i2. pp. 447-459

[13] Sunita Beniwal, Jitender Arora, "Classification and Feature Selection Techniques in Data Mining", International Journal of Engineering Research & Technology (IJERT), Vol. 1, Issue 6, August - 2012, ISSN: 2278-0181