Students' Class Performance Prediction Using Machine Learning Classifiers

Adeel Ahmed[1], Kamlesh Kumar[2], Mansoor Ahmed Khuhro[3,*], Asif Ali Wagan[3], Imtiaz Ali Halepoto[4],
Rafaqat Hussain Arain[5]

[1]Department of Commerce and Data Science, University of Catania, Italy
[2]Department of Software Engineering, Sindh Madressatul Islam University, Karachi
[3]Department of Computer Science, Sindh Madressatul Islam University, Karachi
[4]Department of Software Engineering, QUEST, Nawabshah, Pakistan
[5]Department of Computer Science, Shah Abdul Latif University, Khairpur Mirs
[*]Corresponding author: makhuhro@smiu.edu.pk

Abstract

Nowadays, educational data mining is being employed as assessing tool for study and analysis of hidden patterns in academic databases which can be used to predict student's academic performance. This paper implements various machine learning classification techniques on students' academic records for results predication. For this purpose, data of MS(CS) students were collected from a public university of Pakistan through their assignments, quizzes, and sessional marks. The WEKA data mining tool has been used for performing all experiments namely, data pre-processing, classification, and visualization. For performance measure, classifier models were trained with 3- and 10-fold cross validation methods to evaluate classifiers' accuracy. The results show that bagging classifier combined with support vector machines outperform other classifiers in terms of accuracy, precision, recall, and F-measure score. The obtained outcomes confirm that our research provides significant contribution in prediction of students' academic performance which can ultimately be used to assists faculty members to focus low grades students in improving their academic records.

Keywords—Educational Data Mining; Preprocessing, Classification; Prediction, knowledge discovery

⬩

1 Introduction

Data mining is the process of extracting valuable, explicit, and nontrivial knowledge from large data depositories. However, its main purpose is to inquire large databases in order to find innovative and useful patterns that could assist in better decisions making for the future events that might otherwise remain unknown [1]. Nowadays, data mining is being used in vast areas such as in e-commerce, bioinformatics, recommendation system, outlier analysis, and other scientific applications. In this research, our focus is to use it in education which is termed as Educational Data Mining.

The Education Data Mining (EDM) is becoming 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 for future learning [2]. The researchers' work in education data mining is to analyze students' failures, attrition, and prediction of the performance of the students. The key area of EDM application is to improve the model of the students by identifying various characteristics and attributes which could play a seminal role in the model [3]. Currently, it is a hot area among the researchers due to its potential benefits in the education sector while developing a better model that can explore the hidden knowledge in the academic data [4]. For this purpose, many supervised and unsupervised machine learning techniques have been proposed in the past for EDM to develop an effective model for prediction students' performance [5]. Therefore, implementing data mining techniques in higher education institutes play a paramount role to improve the quality of education and learning experience of students. Moreover, students are important assets in higher education system and their timely graduation play a dominant role for the market. It is crucial for

Fig. 1: Research methodology

Fig. 2: The attributes selection for the mining process

2 Related Work

Educational data mining is a new evolving research in the field of educational analytics. Researchers are more focused on it to find some meaningful pattern which can lead to the accurate model for the prediction for students' performance and improve the quality of higher education in the institutes. In [8], authors proposed a method for analyzing students' performance using classification techniques. They aggregated the students' data and employed the ID3 and J48 algorithms on 239 instances. It was found that J48 gives maximum accuracy as compared to the ID3. It was concluded that class attendance has the highest gain value, and it is largely responsible for the success and failure of the students.

Similarly, another research was conducted by British university in Dubai for the prediction of distinguished students. Objective of this work was to improve students' research capabilities in order to solve the real-world problems. Experiments were performed on the dataset of the university students by implementing several data mining algorithms. The results showed that Support Vector Machine (SVM) with radial kernel provided most accurate prediction in students' performance [9]. In [10], supervised machine learning techniques were employed for analyzing the dropout ratio of students from the course. The classification was performed on past students' data in which results showed that artificial neural networks outperformed other methods.

In [11], authors presented classification techniques to predict performance of students using different algorithms namely J48, ID3, Naive Bayes, IB1, and OneR, implemented using WEKA explorer tool. The size of the data is 60 which is pre-processed and prepared in order to visualize the students' performance. The results showed that Naive Bayes algorithm is the most accurate classifier than others.

In [12], classification techniques were applied on the real data of the female students at the King Saud
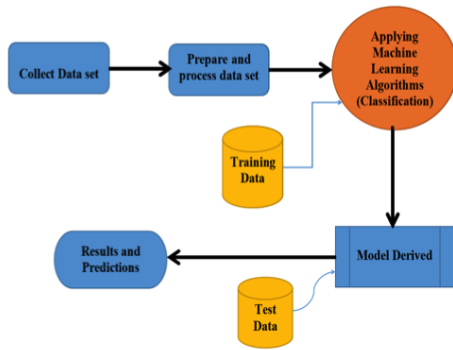
the institutes to detect the factors that adversely affect students' performance [6]. Many public and private universities collect a large amount of academic data, but do not utilize it in a manner that could provide valuable information to education decision-makers in order to plan, evaluate, and decide their educational programs. Therefore, there is a huge gap in the usage of EDM techniques in academics to improve the quality of education and research capabilities of the students [7].

The main aim of this research is to predict the overall class score of students before the final term examinations by using their assignments, quizzes and mid-term scores. For this purpose, the past results of the students were collected from a public-sector university in Pakistan. Based on this data, prediction model was developed using various machine learning techniques. We have used the data of the past batches of MS program in order to predict the performance of the students for various batches. We implemented machine learning classification algorithms by using WEKA tool. However, our main task is to ascertain the best classifier based on classification evaluation metrics such as accuracy, precision, recall and F-measure.

The rest of this paper is arranged as follows. Section 2 provides previous work which has been reported in the past in the field of Education Data Mining. Section 3 describes proposed methodology which discusses the dataset collection, pre-processing, and visualization. Section 4 discusses the implementation of the machine learning algorithms and provide results using well-known evaluation metrics namely, accuracy, precision, recall and F-measure in tabular format. Section 5 illustrates comparative evaluation among various used classifiers. Finally, Section 6 gives conclusion and future work.

Attribute	Description	Possible Values	Categorized Values
Gender	Gender of the students usually does not matter in our case	Male and Female	Male/Female
Batch Year	The data of two batches are used.	2018 and 2019	2019/2018
Program	The four programs are Computer Sciences, Media Sciences, Business Administration and Environmental Sciences.	MS(CS),MS(ES) MS(MS) and MS(BA)	MS(CS)/ MS(ES) / MS(MS)/ MS(BA
Sessional/ Mid Term Marks	It is the marks for the mid-term held in the course.	1 to 30	
Class Participation Marks	This is the marks given by the teacher on the observation of the class participation i.e. asking questions, giving suggestions and presenting any research-oriented topic	1 to 10	
Is Research Oriented	It describes whether a student is taking the program by thesis or course.	Yes and No	Yes/No
Student Type	It describes whether the student is part-time or full-time	Part-Time and Full-Time	Part-Time/ Full-Time
Financial Condition	It describes the students' financial conditions.	Good, Bad and Average.	Good/ Bad/Average
Attendance Marks	It is the attendance marks of the students in the class.	1 to 10	
Final Exam Marks	It is the marks of the final exams taken by the students.	1 to 50	
Total	It is the total marks taken during the course in the semester.	1 to 100	=SUM(Mid-term, Class-participation, attendance Marks, Final Marks)
Class Attribute	This class attribute is created according to the grading system of the university.	A+, A, B+, B, C, D and F Where F stands for the fail.	=IF(Total>=80,"A+", IF(Total>=70,"A", IF(Total>=65,"B+", IF(Total>=60,"B", IF(Total>=50,"C", IF(Total>=40,"D","F"))))))

TABLE 1: List of attributes and values

University to predict the student performances in a course. The authors implemented Naive Bayes, JRip, and Rule-based algorithms to predict the performance of final exams and total points obtained in the course. The model achieved 91% accuracy to predict the failing students prior to final exams.

In another paper [13], the authors employed different classification techniques for students' performance predication such as ID3, C4.5, and Naive Bayes. After comparison, results showed that decision tree classifiers provide higher accuracy prediction than other method. In [14], the decision tree algorithms were used to predict students' performance for passing or failing in the final exams based on its assessment data. The C4.5 and ID3 algorithms were employed

and compared. The results showed that C4.5 algorithm performed well on predicting the pass or fail ratios of the students. The results were given to the instructors in order to improve the performance of the students who were predicted to fail.

The authors at Michigan State University (MSU) applied various tree-based and non-tree-based classification algorithms to predict the final grades of the students. The data was collected through their web-based portal. The results reveled that the accuracy of the prediction of final exam grades increased when the classifier was combined with genetic algorithm [15].

In [16], authors presented comparative analysis of different data mining methods and algorithms in order to classify students through their Moodle usage data

and obtained final marks in the exams for decision making. In [17], Yadav and Pal implemented different decision tree classification algorithms for predicting students' grades who were most likely to fail in their final exams. In their work, they achieved 62.2%, 62.2% and 67.77% overall prediction accuracy for ID3, CART and C4.5 decision tree algorithms, respectively.

Dorina Kabakchieva [18] implemented data mining classification techniques as a research project for the Bulgarian University. The data about the students' personal information and their performance was collected from databases of the university which included 20 parameters of 10330 students. The results obtained through different classifiers varied between 52-67% accuracy.

Shana and Venkatachalam [19] investigated students' accomplishment by using 182 student records and 20 attributes. They used Pearson's Coefficient and F-Test techniques to predict students' performance. In this work, classification algorithms such as Decision-tree induction, REP tree, Simple CART, Naive Bayes were implemented. Results showed that Naive Bayes classification obtained highest accuracy.

In [20], Tanner and Toivanen analyzed 15000 students from an online touch-typing course using K-nearest neighbor method (KNN). This research aims to predict the students which are at high risk of failure in their early stage of the course and then provide an alert to the teacher to take improvement measures. The results indicate that KNN method works well in early test scores predication.

The literature reviewed reveals that in EDM research, mostly classification techniques have been employed in predicting the performance of the students. In our work, we have used various classification for the predication of students' performance on university databases which consists of different courses. The detailed research methodology is explained in the following section.

3 Proposed Methodology

In this section, our proposed methodology has been discussed in which algorithm steps are described as shown in Figure 1. Initially, dataset of students is collected from a public university of Pakistan which contains attributes such as batch year, program, research student, student type, financial condition, class, class participation, midterm, sessional, attendance, and final exam marks. After that, data pre-processing is performed in order to select specific attributes for model building. In the next step, different classifiers are trained on the dataset to predict the students'

performance and their obtained results are compared to assess the classifier's accuracy. The details are provided in following paragraph.

3.1 Data Collection

In this section, the data of graduated students is collected from a public University which contains 952 tuples of data from four different departments, namely Environmental Sciences, Computer Sciences, Media Sciences, and Business Administration. Out of 952 tuples, 13 tuples were omitted after detection of anomalies, and finally we chosen 939 tuples for data mining process. However, each instance contains 20 attributes out of which 12 most representative attributes were selected for building the model as shown in the Figure 2.

3.2 Data Pre-processing

Data pre-processing is an important step for filtering inconsistent, incomplete and noisy data. We collected the dataset from various tables of the university database by using the structured query language (SQL) and implemented the following steps in order to prepare our data for the analytics.

1) Firstly, we eliminated any discrepancy in the naming conventions according to the ID attributes of the table and replace the ids with the values for these attributes.

2) The duplicate records were removed from the tables.

3) the incomplete data of the students who were not shown and absent since the course started.

Finally, we resolved the contradiction in data. The data pre-processing is performed through converting the Excel file into the Comma Separated Values (CSV). After that, WEKA tool manager was implemented to convert CSV format into Attribute-Relation File Format (ARFF). An ARFF file consists of an ASCII text file that represents a record of instances yielding a set of attributes [21]. However, during pre-processing, we have created two extra attributes for the total marks obtained by the students at the end of the semester in the selected course which includes the final exam marks as well. The description of the attributes is shown in Table 1.

3.3 Data Visualization

After data pre-processing, the created ARFF file was loaded into the WEKA tool for data analysis which represents the attribute and its distributions. Figure 3 represents the data related to students per class
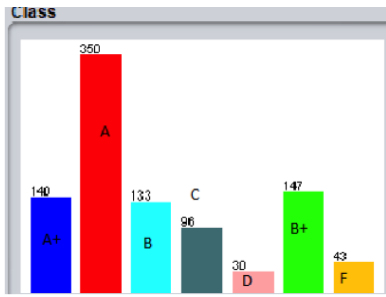
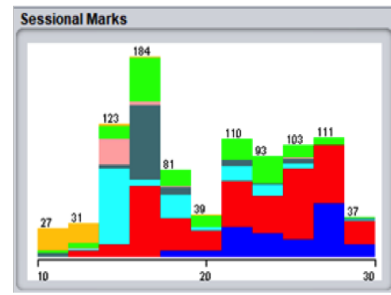Fig. 3: Number of students and class attributes

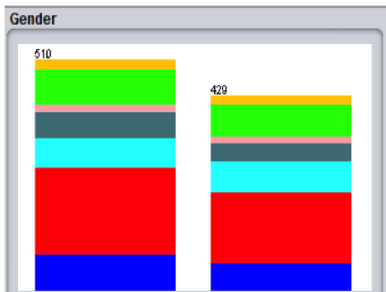Fig. 7: Sessional marks and class attributes

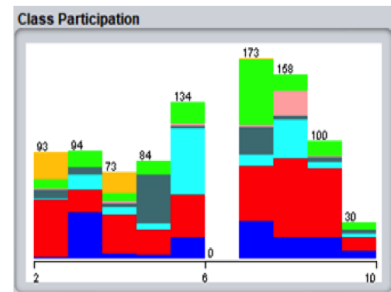

Fig. 4: Gender and class attributes
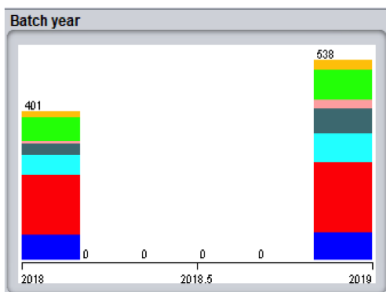

Fig. 8: Class Participation

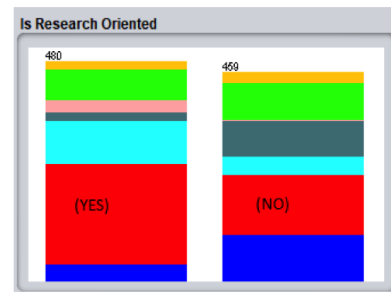Fig. 5: Batch year and class attributes

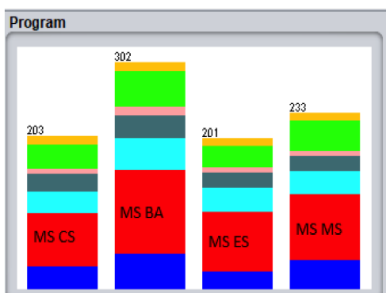Fig. 9: Research oriented and class attributes
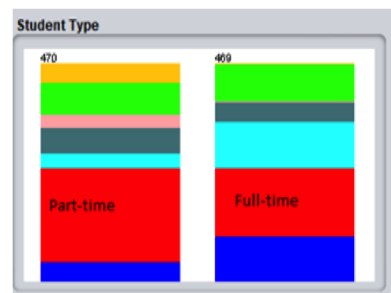

Fig. 6: Program and class attributes

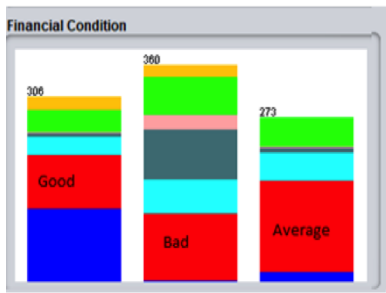Fig. 10: Student type and class attributes

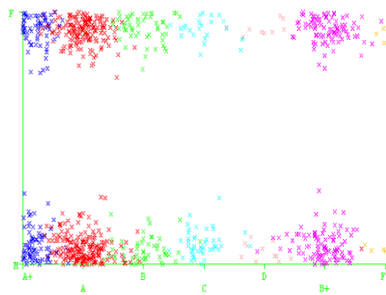Fig. 11: Students financial class attributes

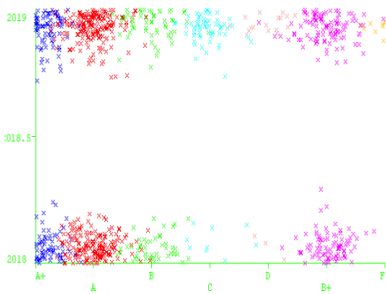Fig. 12: Grades and genders plot
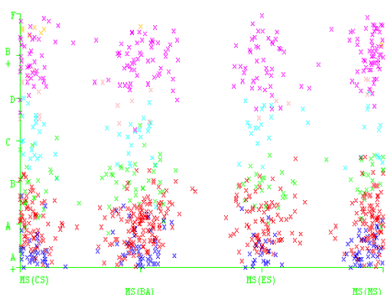

Fig. 13: Academic year plot with grades

Fig. 14: Programs and grades plot diagram

attributes. The number of students who passed the exams with "A+" "A", "B", "C" "D", "B+", and "F" are 140, 350, 133, 96, 30, 147 and 43, respectively. Further, Figure 4 shows the number of instances having gender male (510) and and female (429) students against class attribute. Similarly Figure 5 depicts batch year and class attributes of students. The number of students in the batch years 2018 and 2019 are 401 and 538 respectively.

Figure 6 represents the number of students in each program. Here, color in the bars displays the distribution of the class attribute of the data. The number of students in MS (CS), MS (BA), MS (ES), and MS (MS) are 203, 302, 201, and 233 respectively. However, Figure 7 provides the distribution of sessional marks having minimum value 10, and maximum value 20. The mean of the marks is 20.24 and standard deviation is 5.143.

Figure 8 shows class participation marks which ranges from 2 to 10 with mean value of 6.02 and standard deviation value of 2.28. Figure 9 shows that there are 480 research-oriented students, while 459 students are without a research degree program. Figure 10 represents the type of student in which 470 are part-time and rest are 469 full time students. Similarly, Figure 11 represents the financial condition of the students. The students having good financial condition and bad financial are 306 and 360 respectively, and average financial condition is 273. In addition, there are some other graphical representations which provide attributes relationship between the grade, gender attribute, batch year, grade, and program type, grade.

Figure 12 shows gender against the grades in which it indicates numbers of female students are less than the male students having below average grade B. In Figure 13, academic years are plotted against the grades where it shows that in year 2018 the number of average students are less than the number of average students for the year 2019. It predicts that the students in the year 2018 performed well as compared to year 2019. Figure 14 provides the comparison of the grades with the programs. It reveals that the number of below average students is smaller in Computer Science program and Media Studies as compared to other programs.

After the data visualization in WEKA, we applied different classification techniques on the data to predict students' academic performance. The main purpose of this research is to select a model which classifies most of the test data instances correctly compared to other implemented classifiers which can predict students' grades on future instances [22].

a	b	c	d	e	f	g	Classified as
25	17	0	0	0	2	0	a = A+
3	102	8	1	0	7	0	b = A
0	9	35	0	0	6	0	c = B
1	3	5	19	0	1	0	d = C
0	2	1	0	7	0	0	e = D
0	26	7	0	0	14	0	f = B+
0	1	1	0	0	0	16	g = F

TABLE 2: Confusion matrix of Naive Bayes

TF Rate	TP Rate	Precision	Recall	F Measure	Class
0.568	0.015	0.862	0.568	0.685	A+
0.843	0.293	0.638	0.843	0.726	A
0.700	0.082	0.614	0.700	0.654	B
0.655	0.003	0.950	0.655	0.776	C
0.700	0.000	1.000	0.700	0.824	D
0.298	0.059	0.467	0.298	0.364	B+
0.889	0.000	1.000	0.889	0.941	F
0.683	0.135	0.700	0.683	0.675	W.Av

TABLE 3: Precision, recall and F-measure of Byes net

a	b	c	d	e	f	g	Classified as
28	16	0	0	0	0	0	a = A+
5	104	5	1	0	6	0	b = A
0	7	33	0	0	10	0	c = B
1	4	4	20	0	0	0	d = C
0	2	0	0	7	1	0	e = D
0	17	7	0	0	23	0	f = B+
0	1	0	0	2	0	15	g = F

TABLE 4: Confusion matrix of A2DE

Precision	Recall	F-Measure	Class
0.824	0.636	0.718	A+
0.689	0.860	0.765	A
0.673	0.660	0.667	B
0.952	0.690	0.800	C
0.778	0.700	0.737	D
0.575	0.489	0.529	B+
1.000	0.833	0.909	F
0.733	0.721	0.719	Weighted Average

TABLE 5: Precision, recall and F-measure of A2DE

4 Results & Discussion

In this section, details of the implemented classifiers have been discussed for prediction of student's class performance. For this purpose, dataset is divided into training and test set. However, we have employed 3-Cross and 10-Cross fold validation method for performance analysis. As shown in Table 12, Naive Bayes algorithm and BayesNet algorithm achieved 68.33% and 65.28% on 3-cross and 10-Cross fold validation respectively. Table 2 represents the confusion matrix which tells that model correctly predicted 16 students who were failed and mis-classified 2 instance. For classes A+, A and B, it accuractely classified the number of students. However, the true positive rate of all classes is better except for the class of B+ values, because the number of correctly classified instances is less than the incorrectly classified instances.

Table 3 shows precision, recall and F-measures and weighted averages for each class. The average precision and recall of the Bayes Net algorithm is 0.7 and 0.683 with the F-measure of 0.675. In addition, The Naive Bayes algorithm provides 67.71% accuracy with the 3-Coss Fold and 65.17% accuracy with the 10-cross fold technique.

The variant of Naive Bayes, namely Averaged and Dependence Estimator (A2DE) algorithms provided 69.96% and 66% accuracy with 10-Cross fold and 3-Cross Fold respectively. The confusion matrix of A2DE, as shown in Table 4 reveals that 3-Cross fold validation predicts the classes much better than the Bayes Net algorithm. The students predicted to be in

a	b	c	d	e	f	g	Classified as
39	9	0	0	0	0	0	a = A+
3	114	0	0	0	5	0	b = A
0	3	38	2	0	6	1	c = B
0	4	5	20	0	0	0	d = C
0	3	0	0	7	0	0	e = D
0	16	3	0	0	28	0	f = B+
0	0	0	1	0	1	16	g = F

TABLE 6: Confusion matrix of LibSVM

Precision	Recall	F-Measure	Class
0.921	0.795	0.854	A+
0.765	0.942	0.844	A
0.826	0.760	0.792	B
0.870	0.690	0.769	C
1.000	0.700	0.824	D
0.718	0.596	0.651	B+
0.941	0.889	0.914	F
0.816	0.809	0.805	Weighted Average

TABLE 7: Precision, recall and F-measure of LibSVM

a	b	c	d	e	f	g	Classified as
35	9	0	0	0	0	0	a = A+
4	114	0	0	0	3	0	b = A
0	1	39	2	0	7	1	c = B
0	3	6	20	0	0	0	d = C
0	3	0	0	7	0	0	e = D
0	16	3	0	0	28	0	f = B+
0	0	0	0	0	1	17	g = F

TABLE 8: Confusion matrix of logistic R

Precision	Recall	F-Measure	Class
0.897	0.795	0.843	A+
0.781	0.942	0.854	A
0.813	0.782	0.796	B
0.909	0.690	0.784	C
1.000	0.700	0.824	D
0.718	0.596	0.651	B+
0.944	0.944	0.944	F
0.820	0.815	0.811	Weighted Average

TABLE 9: Recall and F-measure of logistic R Precision

a	b	c	d	e	f	g	Classified as
37	7	0	0	0	0	0	a = A+
5	108	0	0	0	8	0	b = A
0	5	31	6	1	6	1	c = B
0	2	4	21	0	2	0	d = C
0	3	0	0	7	0	0	e = D
0	20	9	1	0	17	0	f = B+
0	0	0	1	0	0	17	g = F

TABLE 10: Confusion matrix of bagging classifier

Class B+ are more accurate than other Naive Bayes techniques. Table 5 shows improved average precision and recall of 0.733 and 0.721 respectively. The average F-measure generated by A2DE model is 0.719 which is also better than Naive Bayes algorithm.

Further, LibSVM which is a variant of support vector machine classification algorithm provides accuracy of 80.87% and 77.95% on 3-Cross fold and 10-Cross fold validation as shown in Table 12. Table 6 depicts the confusion matrix values of LibSVM. The obtained results in Table 7 show that SVM algorithm gives much better results compared to Naïve Bayes algorithm. It provides an average precision value of 0.816 and recall value of 0.809. The average F-Measure value of LibSVM model is 0.805 which is optimal as shown in Table 7. If we examine logistic R algorithm in Table 12, it provides accuracy of the 74.12% and 74.60% on the 10-Cross fold and 3-Cross fold techniques respectively. For Logistic R Table 8 represents its confusion matrix and Table 9 shows average precision, recall and F-measure.

Precision	Recall	F-Measure	Class
0.881	0.841	0.860	A+
0.745	0.893	0.812	A
0.705	0.620	0.660	B
0.724	0.724	0.724	C
0.875	0.700	0.778	D
0.515	0.362	0.425	B+
0.944	0.944	0.944	F
0.737	0.746	0.736	Weighted Average

TABLE 11: Precision, recall and F-measure of bagging classifier

Lazy local KNN algorithm with 10-Cross fold, and 3-Cross folds validation provides accuracy of 71.56% and 75.23% as shown in Table 12. Similarly some other algorithms are also compared and presented in Table 12, where JRip algorithm provides the accuracy of 78.05% and 74.76% with the 10-Cross fold and the 3-Cross fold validation respectively. Further, three tree-based algorithms are also evaluated and compared such as J48, logical model tree and random forest. The J48 algorithm achieved accuracy of up to 78.05% and 76.47% using 3-cross fold and 10-cross fold validation respectively. Furthermore, Logical Model Tree (LMT) algorithm achieve accuracy of 76.80% and 78.38% using the 3-Cross fold and 10-Cross fold validation techniques respectively. Finally, Random Forest algorithm is applied on accuracy of 77.63% and 78.99% respectively over 3-Cross fold and 10-Cross fold validation respectively. The average precision, recall and F-measure values of the tested algorithms have been shown in tabular results.

5 Comparative Evaluation

This section provides comparative analysis among different implemented classifiers. The bagging classifier, a meta-estimator, when combined with the classifier of the LibSVM improves the accuracy as compared to Lib SVM algorithm. The combined algorithm predicts students' performance accurately and achieves 80.87% and 77.95% accuracy with 3-cross fold and 10-cross fold as shown in Table 12. Figure 15 depicts ROC curve which represents that accuracy for the prediction of the failure instances compared to the other classifiers. The area under the curve is 0.972 for the class attribute value F, while rest of the instances are correctly predicted. It shows that the prediction of the students having class B+ attribute has higher precision, recall and F-measure compared to LibSVM algorithm. Figure 16 shows the ROC curve of the attribute class B+ which is approximately 71% better than the other tested models. The average precision, recall and F-measure values of bagging classifier are quite higher, which confirms the efficacy of this model while predicting the students' failure and success in final examination.

In this research, total 11 classification algorithms were implemented for investigating the best classifier for the prediction of the student's performances in the final exams. Table 12 provides results of the implemented classifiers. It can see from tabular outcomes that bagging with support vector machine algorithm outperformed other tested algorithms for 3-cross and 10-croos fold validation techniques. It provided highest
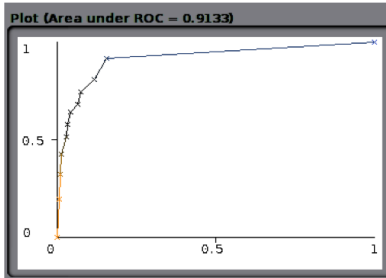
Fig. 15: ROC curve for failure prediction

Fig. 16: ROC curve for B+ prediction

accuracy of 81.50% with average precision, average recall, and F-measure value of 0.820, 0.815, and 0.811 respectively. From these numerical results, it can be concluded that bagging algorithm with SVM works well while predicting students' performance ratios in terms of success and failure.

6   Conclusion

In this paper, we investigated the university students' academic records to predict the performances of the students. For this purpose, dataset of students was created using their quizzes, assignments, mid-term and final term exams. WEKA was used to evaluate the results. Our main purpose in this research was to find out the most effective classifier which yields best accuracy. After comparative analysis, the obtained results revealed that bagging algorithm, when combined with SVM classifiers, gives optimal outcomes in terms of accuracy, precision, recall and F-measure score on all class attributes. This research was carried out for MS program students; however, in future we plan to extend this model on more academic programs to validate its efficacy. Furthermore, it is highly recommended that one can work on the selection of more significant attributes from academic records which might affect students' class performance from academic records.

References

[1] Larose, D.T. and Larose, C.D. "Discovering knowledge in data: an introduction to data mining." John Wiley & Sons, vol. 4, 2014.

[2] Mohamad, S.K. and Tasir, Z. "Educational data mining: A review." Procedia-Social and Behavioral Sciences, vol. 97, pp.320-324, 2013.

[3] Sukhija K, Jindal M, Aggarwal N. "The recent state of educational data mining: A survey and future visions." In IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), pp. 354-359, 2015.

[4] Dutt A, Ismail MA, Herawan T. "A Systematic Review on Educational Data Mining." IEEE Access, vol. 5, pp. 15991–16005, 2017.

[5] Almasri, A., Alkhawaldeh, R.S. and Çelebi, E. "Clustering-Based EMT Model for Predicting Student Performance." Arabian Journal for Science and Engineering, vol. 45, pp.10067-10078, 2020.

[6] Yanto, I.T.R., Sutoyo, E., Rahman, A., Hidayat, R., Ramli, A.A. and Fudzee, M.F.M. "Classification of Student Academic Performance using Fuzzy Soft Set. In 2020 International Conference on Smart Technology and Applications (ICoSTA), pp. 1-6, 2020.

[7] Siraj, F. and Ali, M. "Mining enrollment data using descriptive and predictive approaches." InTech—Open Access Company, pp. 53-72, 2011.

[8] Sunday, K., Ocheja, P., Hussain, S., Oyelere, S., Samson, B. and Agbo, F. "Analyzing student performance in programming education using classification techniques." International Journal of Emerging Technologies in Learning (iJET), vol. 15, no. 2, pp.127-144, 2020.

[9] Zohair, L.M.A. "Prediction of Student's performance by modelling small dataset size." International Journal of Educational Technology in Higher Education, vol. 16, no. 1, pp.1-18, 2019.

[10] Tomasevic, N., Gvozdenovic, N. and Vranes, S. "An overview and comparison of supervised data mining techniques for student exam performance prediction." Computers & education, vol. 143, p.103676, 2020.

[11] Roy, S. and Garg, A. "October. Predicting academic performance of student using classification techniques." In 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), pp. 568-572, 2017.

[12] Al-Barrak, M.A. and Al-Razgan, M.S. "Predicting students'performance through classification: a case study." Journal of Theoretical & Applied Information Technology, vol. 75, no. 2, 2015

[13] Abdulsalam, S.O., Saheed, Y.K., Hambali, M.A., Salau-Ibrahim, T.T. and Akinbowale, N.B. "Students' Performance Analysis Using Decision Tree Algorithms." Anale. Seria Informatică, vol. 15, no. 1, pp.55-62, 2017.

[14] S. Anupama Kumar and Dr. Vijayalakshmi M.N. "Efficiency of decision trees in predicting student's academic performance." First International Conference on Computer Science and Engineering and Applications, pp. 335–343, 2011.

[15] Minaei-Bidgoli, B., Kashy, D.A., Kortemeyer, G. and Punch, W.F. "Predicting student performance: an application of data mining methods with an educational web-based system." In 33rd Annual Frontiers in Education, Vol. 1, pp. T2A-13, 2003.

[16] Romero, C., Ventura, S., Espejo, P.G. and Hervás, C. "Data mining algorithms to classify students." In Educational data mining, 2008.

[17] Yadav, S.K. and Pal, S. "Data mining: A prediction for performance improvement of engineering students using classification." arXiv preprint arXiv:1203.3832, 2012.

[18] Kabakchieva, D. "Predicting student performance by using

S. No	Algorithm	3-Cross Fold in %	10-Cross Fold in %	Precision (Avg.) 3-Cross Fold	Recall (Avg.) 3-Cross Fold	F-Measure (Avg.) 3-Cross Fold
1	A2DE classifier	72.10	69.96	0.733	0.721	0.719
2	Bayes Net	68.33	65.28	0.7	0.68	0.67
3	Naive Bayes	67.71	65.17	0.681	0.677	0.662
4	LibSVM	80.87	77.95	0.816	0.809	0.805
5	KNN	75.23	71.56	0.750	0.749	0.747
6	Bagging with SVM	81.50	77.95	0.820	0.815	0.811
7	Logistic R	74.60	74.12	0.737	0.746	0.736
8	JRip	78.05	74.76	0.784	0.781	0.781
9	J48	76.46	78.05	0.788	0.781	0.783
10	LMT	76.80	78.38	0.770	0.768	0.768
11	Random Forest	78.99	77.63	0.799	0.790	0.790

TABLE 12: Comparative evaluation of implemented machine learning classifiers in terms of precision, recall, and F-measure

data mining methods for classification." Cybernetics and information technologies, vol. 13, no. 1, pp.61-72, 2013.

[19] Shana, J. and Venkatachalam, T. "Identifying key performance indicators and predicting the result from student data." International Journal of Computer Applications, vol. 25, no. 9, pp.45-48, 2011.

[20] Tanner, T. and Toivonen, H. "Predicting and preventing student failure–using the k-nearest neighbour method to predict student performance in an online course environment." International Journal of Learning Technology, vol. 5, no. 4, pp.356-377, 2010.

[21] Attribute-Relation File Format (ARFF). [cited 10 Jan 2020]. Available: https://www.cs.waikato.ac.nz/ml/weka/arff.html

[22] Romero, C., Ventura, S., Pechenizkiy, M. and Baker, R.S. eds. "Handbook of educational data mining." CRC press, 2010.