A Comparative Analysis of Exam Timetable Using Data Mining Techniques

Bilal Sowan

Department of Computer Network Systems, Applied Science Private University, Amman, Jordan

Summary
Knowledge discovery and data mining is an emerging practice that is applied in a wide range domain fields, for the purpose of extracting implicit knowledge from a huge database. This knowledge helps in making a decision in particular fields. It is one of the important developing applications is the higher education field. This paper proposes a data mining model, which is based on different and well-known classification algorithms. The model is able to extract implicit knowledge from the higher education dataset, specifically the dataset concerns the student satisfaction of courses-exams timetable. The courses-exams timetable satisfaction is considered as one of the complex and main significant factors that effects the student passing of exams. The paper also studies the reasons behind the student exam passing. Therefore, the application of data mining model in this phenomenon improves the decision making process and considers an automated analytical tool that provides better and clearer knowledge. The importance of this knowledge is to provide a feedback for the higher educational institution management related to the courses-exams timetable satisfaction for further decision quality improvements. The proposed model is validated with several experiments for the purpose of comparing different classification algorithms to select the fitting one on the dataset that is used in this research. J48, REPTree, MLP, SVM, SVM, JRip, and Prism are applied to evaluate the performance of the proposed model. As a result, both, MLP and Prism, outperformed the other algorithms.

Key words:
Data mining, Classification algorithms, Knowledge, Higher education.

1. Introduction

Knowledge Discovery in Databases (KDD) and data mining are the well-known concepts that apply the techniques to extract knowledge from large databases [1]. KDD is the complete processes of extracting knowledge, where the data mining is one of the main processes of KDD [2, 3]. In general, KDD and data mining are used as an interchangeable terms [4, 5]. Data mining approaches have been evolved in order to extract knowledge and frequent pattern in different domains [6].

Nowadays, the knowledge is necessary as an important factor in the world economy. The knowledge extraction process play a vital role as a resultant to ensure the knowledge quality. As a result, the extraction process reflects the effect of the dependency and the decision that has been taken based on the extracted knowledge [7]. Rapid growing of data that produced by a wide range of fields, such as marketing, medical, software engineering [8], telecommunication and education require an automated methods and tools in order to help human in extracting knowledge from the growing data [9].

Data mining includes many tasks. One of the most important tasks is the classification task. In classification, the model is constructed in order to learn dataset for the purpose of predicting or classifying a future value. Generally, the data in classification should consist of different features (inputs) and the class label (output). Higher Education Institutes (HEI) is one of the attractive fields for applying knowledge extraction techniques. It is considered as data centers, thus, it is important to identify and validate the processes of the usefulness of the data to improve the educational procedures to increase the education quality. Education quality can be achieved by providing high quality knowledge to assist in transferring powerful and effective learning objectives between HEI stakeholders. As a result, it is required to apply data mining techniques in this field [7].

The application of data mining for extracting knowledge from educational data has a great attention, which reflects on educational outcomes. The outcomes help in improving learning process that implies the students' success. It is believed that one of the most important factors effecting students success is courses-exams timetable satisfaction and convenience. What is meant by satisfaction here is the fair distribution of exams all over the period assigned for exams. This fair distribution can reduce the submission of two exams on the same day or on consecutive days to the minimum. Leaving gaps between exams gives the student the opportunity to review the exam material before the exam and reduces the stress on the students of submitting exams consecutively. Basically, when the courses-exams timetable is convenient for students, this will be reflected on students in two ways. The first, the student will have the opportunity to gain better understanding of the material of a particular course as there is no stress on them. The second, the student will have enough time to be able to

Manuscript received January 5, 2017
Manuscript revised January 20, 2017

review the exam material fully. Accordingly, students can achieve better results in the exams. As a consequence, the education quality of the institution (University) will be increased in general.

This paper proposes a classification model based on different and well-known classification algorithms. The goal of this study is to predict the student satisfaction based on different factors such as student studying level, student general average (GPA), student previous semester average and some other factors detailed in Table 1. The model is able to extract implicit knowledge from the courses-exams timetable satisfaction dataset in the higher education field. The dataset is taken from a real life example of a higher education institute, a University in Jordan. The University adopts first, second, and final exams scheme. The dataset included the input factors (Table 1) and the student satisfaction of the first, second, and final exams timetables. After each exam period (the exam period is two weeks), a questionnaire is distributed on a random sample of the students and the required data is collected. Each random sample included around 50 of different studying levels, first, second, third, and fourth year students. The three random samples resulted 147 students in total. The questionnaire results were analyzed for the purpose of comparing J48, JRip, SVM, MLP, REPTree, and Prism classification algorithms. The algorithms are used to classify the student's satisfaction of courses-exams timetable. As a result, both, MLP and Prism, provided better results over the other algorithms.

The reset of the paper is organized as follows. Section 2, presents the literature review and related work. Section 3, demonstrates the proposed classification model. Section 4, shows the experimental results. Section 6, provides the conclusions and future work.

2. Literature Review and Related Work

Data mining has different application fields. Aburrous et al. [10] proposed a model that combines fuzzy logic and data mining approaches. The model is able to classify electronic banking phishing website. The results showed effectiveness of using fuzzy data mining to identify and classify the phishing website. One of the most important application fields of data mining is the educational field. The application of data mining tasks in educational system is known as the Educational Data Mining (EDM) [11]. The educational system is an important domain field, which includes many factors related to the educational process such as courses-exams timetable and courses scheduling. These factors play a significant role in the quality of educational system toward improving the educational level and seeking for achieving the needs of students, academic staff, and administrators. All stakeholders in the educational process that can affect and effected by this process, get benefits from applying data mining on educational data [12]. Natek and Zwilling [13] studied the importance of the application of data mining algorithms on a small data set in the domain of Higher Education Institutes (HEI). The paper concluded that using data mining is considered as an important tool to extract knowledge for helping in decision making process rather than depending on human experience for predicting the students' success rate in a particular course. This, in general, helps the students to achieve their goals of getting better results in their exams. Blagojević and Micić [14] presented an intelligent system for e-learning based on a method called PDCA (Plan, Do, Check, Act). The aim of the study was to predict the student behavior patterns in e-learning systems or (Learning Management Systems). The system improved the web-based intelligent report. Edin and Mirza [12] compared different data mining techniques to evaluate the student success. The paper studied the factors that influence the student passing grade. However, they did not include the factors that affect the passing or failing the students. Kaur et al. [11] compared various classification algorithms such as Multi-Layer Perception (MLP), J48, SMO, REPTree, and Naïve Bayes. The algorithms were applied for recognizing the slow learner students. The paper depicted the importance of data mining role in the field of educational field.

Al Deen et al. [15] evaluated the effectiveness of using the classification based association rules technique MMAC. The evaluation is applied on different datasets from UCI repository. The MMAC was compared with several classification algorithms such as C4.5, OneR, PART, RIPPER, CBA, and Naïve Bayes. The comparison was against the classification accuracy and the run time. As a result MMAC produced a good classification accuracy, whereas Naïve Bayes and OneR are quick in terms of run time for building classification model.

3. Proposed Classification Model

In this section, the dataset is described and the proposed model is constructed to evaluate the dataset. Furthermore, several performance evaluation measures are illustrated.

In this paper, the proposed model compares different classification algorithms to predict the student satisfaction of the courses-exams timetable. The proposed model identifies the reasons that effect the student exam passing. This model is constructed based on the previously discussed literature and an expert consultation who proposed the problem of student satisfaction with courses-exams timetable. Exams timetable was chosen to be studied in this paper because it is considered as one of the main factors that affect the student grade and exam passing.

The proposed model (shown in Fig. 1) is described below in details. The model consists of the following steps:

Step1: The higher education is selected as an application domain.

Step2: The student satisfaction of the courses-exams timetable dataset, which includes different factors in addition to the satisfaction result as summarized in (Table 1), is created through data collection using a questionnaire.

Step3: The dataset is cleaned by removing the missing values and the outliers in the preprocessing process. After that, dataset is discretized by dividing some features values into intervals.

Step4: The dataset is divided into training set for learning and testing set for evaluating the model.

Step5: The model is constructed based on the training set using different and well-known classification algorithms such as, J48, JRip, MLP, REPTree, SVM, and Prism.

Step6: The classification algorithms are tested and evaluated by calculating the accuracy and the error rate.

Step7: Finally, the results of the proposed model are evaluated.

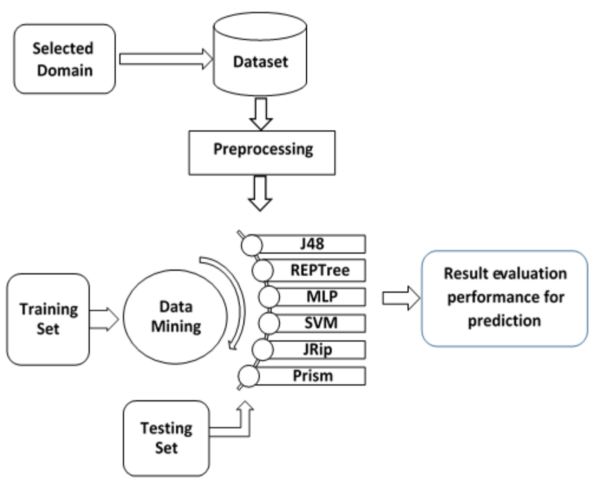

Fig. 1 Proposed classification model.

3.1 Dataset Description

A real world example was used to collect the dataset used in the proposed model. The data is collected using a questionnaire conducted at a University in Jordan. The questionnaire was randomly distributed on different student samples. The students submit three exams per semester, first, second, and final exams. The questionnaire was distributed three times, one after each exam period that is two weeks long. The dataset consists of 9 features including the class of the student satisfaction. A sample totaling of 147 questionnaires (students) for the three exams were collected. The description details of the student satisfaction dataset including input features and output class label shown in Table 1 are illustrated below:

- Student level of study is an ordinal category data type, which categorized into 4 levels. For example, 1 is for first year students, 2 second year students, and so forth.
- Student GPA is a continuous value data type. Its values range from 60 to 100.
- Student grade for the previous semester is a continuous value data type. Its values range from 60 to 100.
- Number of courses inside the faculty per semester. This feature differentiates between the courses belong to the faculty and those that belong to other faculties as university requirement courses. The value for this feature is a nominal category data type ranging from 1 to 6 courses.
- Number of courses outside the faculty per semester is the feature that describes number of courses that belongs to university requirement courses. The value for this feature is a nominal category data type ranging from 1 to 3 courses.
- Maximum number of exams per day for the first term exam. It represents the maximum total number of exams came together for a student in one day. The feature values is a nominal category data type includes 1 exam up to 3 exams. Maximum number of exams per day for the second term exam and the final term exam are similar features for the second tem exam and final term exam.
- Student satisfaction is a class label. It represents the general satisfaction of courses-exams timetable. Student satisfaction is distributed into four labels including the following:
  - Class label 1 represents poor satisfaction.
  - Class label 2 represents fair satisfaction.
  - Class label 3 represents good satisfaction.
  - Class label 4 represents excellent satisfaction.

As a preprocessing step, a discretization of some input features of the student satisfaction dataset is applied. The input features of the continuous values, including student GPA and student grade features, are discretized into 4 categories based on their importance. Category 4 indicates the highest GPA and the highest student grade (features number 2 and 3 in Table 1). Also, the missing values and incomplete records (questionnaires) are removed.

Table 1: Student satisfaction dataset description

Number	Attribute Name	Value
1	Student level of study	Ordinal (1, 2, 3, 4)
2	Student GPA (Grade Point Average)	Continuous (60-100)
3	Student grade (previous semester)	Continuous (60-100)
4	Number of courses inside the faculty	Nominal (1, 2, 3, 4, 5, 6)
5	Number of courses outside the faculty	Nominal (1, 2, 3)
6	Max number of exams per day (first exam)	Nominal (1, 2, 3)
7	Max number of exams per day (second exam)	Nominal (1, 2, 3)
8	Max number of exams per day (final exam)	Nominal (1, 2, 3)
9	Student satisfaction (class labels)	Ordinal (poor=1, fair=2, good=3, or excellent=4)

2.1 Performance Evaluation Measures

Several performance measures are usually used in the research area to evaluate classification problems. This research applies the well-known measures introduced in [16-19], such as accuracy and others (mentioned below), to evaluate the classification algorithms correctness. The classification accuracy is one of the most important measures that lead to evaluate the correctness. Confusion matrix in Table 2, Fig. 2, and the following Eq. (1) into Eq. (8) illustrate the full list of measures used in this research.

Table 2: Binary confusion matrix with two classes, YES and NO.

		Predicted classes		Total
		Yes	NO	
Actual classes	Yes	TP	FN	N
	NO	FP	TN	M
Total		N'	M'	

True Positive (TP): the number of cases (class labels) are correctly classified as positive by the classifier (classification algorithm) (number of cases where predicted YES, and the actual class label is YES).

$$\text{TP rate} = \frac{\text{TP}}{\text{N}} \tag{1}$$

True Negative (TN): the number of cases (class labels) are correctly classified as negative by the classifier (number of cases where predicted NO, and the actual class label is NO).

$$\text{TN rate} = \frac{\text{TN}}{\text{M}} \tag{2}$$

False Positive (FP): the number of cases (class labels) are incorrectly classified as positive by the classifier (number of cases where predicted YES, and the actual class label is NO).

$$\text{FP rate} = \frac{\text{FP}}{\text{M}} \tag{3}$$

False Negative (FN): the number of cases (class labels) are incorrectly classified as negative by the classifier (number of cases where predicted NO, and the actual class label is YES).

$$\text{FN rate} = \frac{\text{FN}}{\text{N}} \tag{4}$$

Based on these previous terms shown in the Eq. (1) into Eq. (4) the following Eq. (5) into Eq. (8) are:

Precision measures the data of class labels with the positive labels identified by the classifier (measures the predictive power of the classifier).

$$\text{Precision} = \frac{\text{TP}}{(\text{TP}+\text{FP})} \tag{5}$$

Sensitivity (Recall) measures how the effectiveness of the classifier to identify the positive labels.

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP}+\text{FN})} \tag{6}$$

The **accuracy** measures the overall effectiveness of the classifier.

$$\text{Accuracy} = \frac{(\text{TP}+\text{TN})}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})} \tag{7}$$

F-Measure is the harmonic mean of precision and recall. The relationship between precision and recall is inverse, which means an increase in precision corresponds with a decrease in recall.

$$\text{F} - \text{Measure} = \frac{(2 * \text{Precision} * \text{Sensitivity})}{(\text{Precision} + \text{Sensitivity})} \tag{8}$$

The F-measure ranges from 0 to 1. If the value is 0, then no relevant records are retrieved, and if it is 1, then all retrieved records are relevant and all relevant records are retrieved [20].

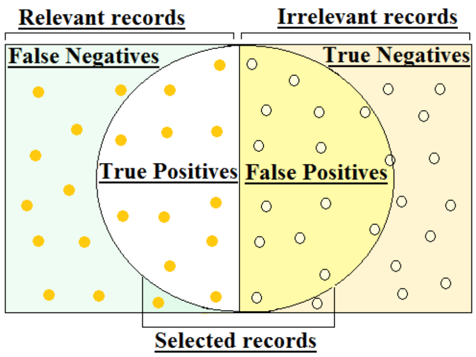

Fig. 2 Precision and recall.

4. Experimental Results

This section presents the results of evaluating the proposed model and the empirical studies. In this study, the predictive accuracy of various classification algorithms for predicting student satisfaction of courses-exams timetable are compared. J48, JRip, SVM, MLP, REPTree, and Prism classification algorithms are investigated in order to determine the most suitable one that is able to learn the student satisfaction dataset for building the proposed classification model.

The experiments on the dataset, which is consisted of 147 students (questionnaires), were conducted using 10 folds cross-validation method. This is considered as model evaluation process. In cross-validation method, the dataset is divided randomly into 10 blocks of equal size, i.e. the dataset is split to training and testing sets. Nine blocks are used for training to build the classification model, whereas one block is used for testing to calculate the error rates. This process is repeated 10 times. Each time one of the blocks, which is used for testing in the previous iteration, is replaced with another block for training. Then, the average error rate is calculated across all blocks.

WEKA data mining tool [21] was used to perform this study. The standard format is applied on the dataset in order to use WEKA tool. The statistical measures for each classifier (classification algorithm) are shown in Table 3 and Table 4. FP rate is an indicator used to measure the performance of predictive accuracy. Classification algorithm with the lowest FP rate has higher accuracy than others. Table 4 indicates that Prism has produced the lowest weighted average of FP rate (10.2%) for all class labels, followed by MLP (12.5%). SVM has the highest weighted average of FP rate (30.3%).

Table 3: Evaluation measures (precision, recall, and F-measure) comparison of each class label for all classification algorithms.

Classifier	Class	Precision	Recall	F-Measure
J48	1	0.75	0.6	0.667
	2	0.833	0.781	0.806
	3	0.804	0.932	0.863
	4	1	0.182	0.308
	(WA)[a]	0.82	0.808	0.789
JRip	1	0.8	0.8	0.8
	2	0.829	0.906	0.866
	3	0.839	0.886	0.862
	4	0.667	0.182	0.286
	(WA)[a]	0.82	0.829	0.813
MLP	1	0.824	0.933	0.875
	2	0.875	0.875	0.875
	3	0.88	0.92	0.9
	4	0.6	0.273	0.375
	(WA)[a]	0.852	0.863	0.852
SVM	1	0	0	0
	2	0.813	0.813	0.813
	3	0.754	0.977	0.851
	4	0	0	0
	(WA)[a]	0.633	0.767	0.691
REPTree	1	0.714	0.333	0.455
	2	0.839	0.813	0.825
	3	0.776	0.943	0.851
	4	1	0.091	0.167
	(WA)[a]	0.8	0.788	0.753
Prism	1	1	0.933	0.966
	2	0.763	0.935	0.841
	3	0.908	0.919	0.913
	4	1	0.364	0.533
	(WA)[a]	0.893	0.881	0.874

(WA)[a]:Weighted Average.

Table 4: Evaluation measures (TP rate, FP rate, and ROC area) comparison of each class label for all classification algorithms.

Classifier	Class	TP Rate	FP Rate	ROC Area
J48	1	0.6	0.023	0.911
	2	0.781	0.044	0.917
	3	0.932	0.345	0.843
	4	0.182	0	0.729
	(WA)[a]	0.808	0.22	0.858
JRip	1	0.8	0.023	0.925
	2	0.906	0.053	0.914
	3	0.886	0.259	0.805
	4	0.182	0.007	0.764
	(WA)[a]	0.829	0.17	0.838
MLP	1	0.933	0.023	0.971
	2	0.875	0.035	0.961
	3	0.92	0.19	0.904
	4	0.273	0.015	0.775
	(WA)[a]	0.863	0.125	0.914
SVM	1	0	0	0.5
	2	0.813	0.053	0.88
	3	0.977	0.483	0.747
	4	0	0	0.5
	(WA)[a]	0.767	0.303	0.732
REPTree	1	0.333	0.015	0.861
	2	0.813	0.044	0.912
	3	0.943	0.414	0.842
	4	0.091	0	0.575
	(WA)[a]	0.788	0.261	0.839
Prism	1	0.933	0	0.967
	2	0.935	0.08	0.914
	3	0.919	0.14	0.88
	4	0.364	0	0.682
	(WA)[a]	0.881	0.102	0.881

(WA)[a]: Weighted Average.

Table 5 and Fig. 3 represent the classification accuracy that is generated by all classification algorithms. It is revealed that MLP and Prism produced similar accuracy results. It is also found that both, MLP and Prism, outperformed J48, JRip, SVM, and REPTree algorithms. Table 5 and Fig. 4 depict the error rate for all the classification algorithms. Although the experiments illustrate that the minimum error rate is generated by Prism, it was not able to classify some instances. It is worth mentioning that the Prism and MLP classification algorithms almost produced the same error rate with a difference of existing unclassified instances in case of the Prim. In contrast, the highest error rate is produced by SVM classification algorithm, which also generated the lowest classification accuracy.

To find the classification algorithm that beat others based on the statistical significance 95%, paired t-test was used. Table 5 shows that the accuracy results of both MLP and Prism are better than the accuracy of J48, JRip, SVM, and REPTree. Furthermore, the results illustrate that there is no statistical significant difference between using “Prism or MLP” and “J48 or JRip”. In contrast, there was a statistical significant difference between using “Prism, MLP, J48, or JRip” and the other classification algorithms. A “*” next to each value in Table 5 indicates that a significance difference exists between it and those do not have a “*”.

Table 5: Classification accuracy and error rate comparison for all classification algorithms.

Classification Algorithm	Accuracy % (Correctly Classified Instances)	Error rate% (Incorrectly Classified Instances)
Prism	86.3%	11.6% (2%)[a]
MLP	86.3%	13.7%
JRip	82.9%	17.1%
J48	80.8%	19.2%
REPTree	78.8%*	21.2%
SVM	76.7%*	23.3%
(2%)[a]: Unclassified Instances.		

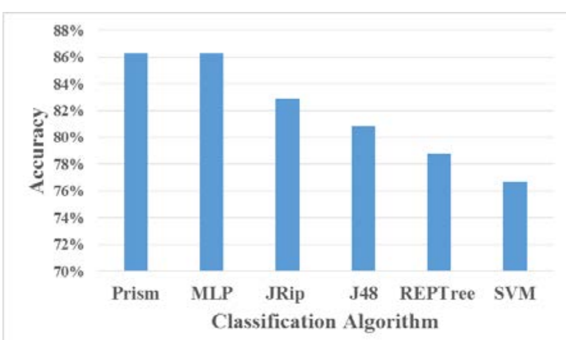

Fig. 3 Prediction accuracy for all classification algorithms.

Figs. (5-8) show the Receiver Operating Characteristic (ROC) of students’ satisfaction for all class labels. ROC is considered as the most powerful performance measure that reflect the predictive accuracy [18]. The ROC is the graphical illustration for the Area Under the Curve (AUC) that state the relation between the TP and the FP. In general, the closer ROC curve to the top left corner, the higher accuracy of the classification algorithm. The optimal value of AUC is 1. That means, the relationship between TP and FP rates indicates that TP rate=1 and FP rate=0. Results depict that MLP and Prism achieved the highest ROC values. This is clear from the trend of the MLP and Prism curves that are directed more toward the TP than toward the FP. Hence, the accuracy of MLP and Prism is considered the most satisfactory among all the classification algorithms.

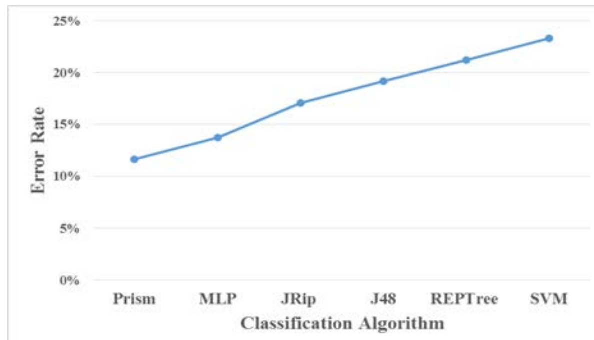

Fig. 4 The error rate for all classification algorithms.

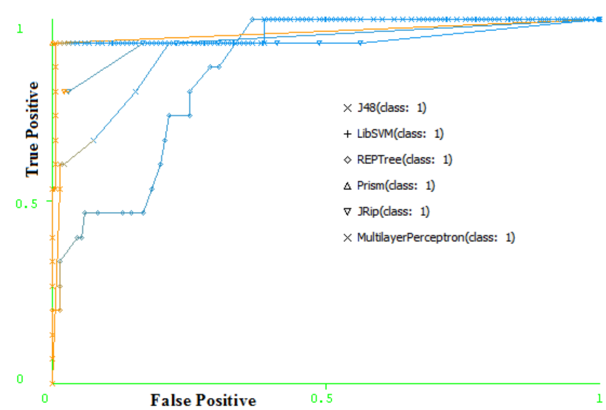

Fig. 5 ROC curve of class label 1 for all classification algorithms.

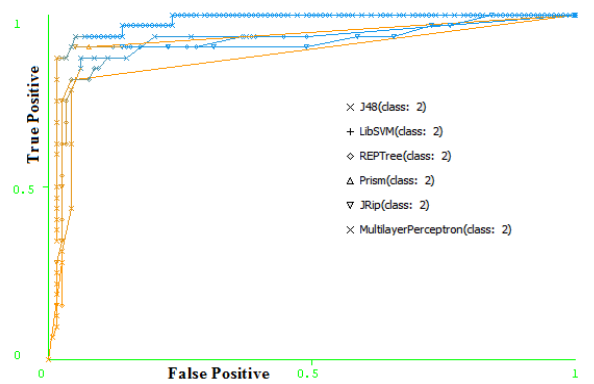

Fig. 6 ROC curve of class label 2 for all classification algorithms.

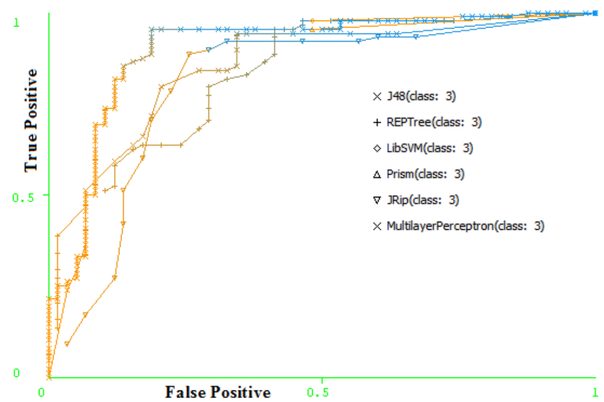

Fig. 7 ROC curve of class label 3 for all classification algorithms.

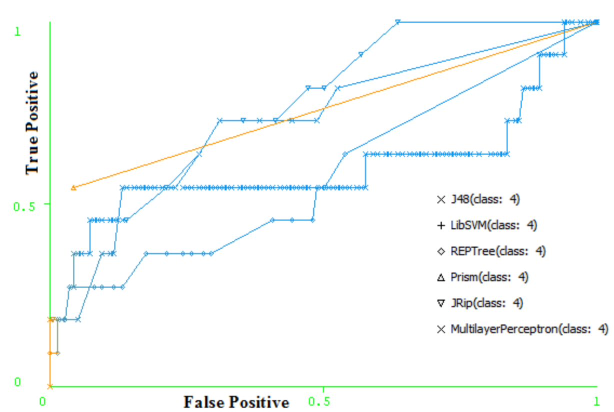

Fig. 8 ROC curve of class label 4 for all classification algorithms.

5. Conclusions and Future Work

This paper showed a proposed classification model. The model applied a variety of well-known classification algorithms on a higher education dataset. The classification algorithms include J48, JRip, MLP, SVM, REPTree and Prism. Results revealed that using both, MLP and Prism, achieved satisfactory results in terms of classification accuracy than using the other algorithms. The main contribution of this research is building a classification model based on real dataset collected in one of the higher education institutions (a University). The main goal of the proposed model is to identify the reasons that have effect on the student exams passing. The model examined the fair distribution of exams in three examination periods. The fair exam distribution, which provides gabs between exams, affects the ability of the student to achieve better knowledge in a course, gain better opportunity to review the exam material and subjected the student to less stress. Although a small data set was available, the model is considered as a significant tool that can be used to enhance the decision quality in the higher education depending on the features used in the model.

As a future work, the model can be extended to explore other factors such as the semester courses schedule. Such a model can help in building an expert system tool that is considered as a full automatic decision support system for higher education.

Acknowledgments

The author is grateful to the Applied Science Private University, Amman, Jordan, for the full financial support granted to this research.

Refrences

[1] Agrawal, R. and J.C. Shafer, Parallel mining of association rules: Design, implementation, and experience. 1996: IBM Thomas J. Watson Research Division.
[2] Fayyad, U.M., et al., Advances in knowledge discovery and data mining. Vol. 21. 1996: AAAI press Menlo Park.
[3] Piateski, G. and W. Frawley, Knowledge discovery in databases. 1991: MIT press.
[4] Piatetsky-Shapiro, G., Knowledge discovery in databases: 10 years after. ACM SIGKDD Explorations Newsletter, 2000. 1(2): p. 59-61.
[5] Kurgan, L.A. and P. Musilek, A survey of Knowledge Discovery and Data Mining process models. The Knowledge Engineering Review, 2006. 21(01): p. 1-24.
[6] Mariscal, G., O. Marban, and C. Fernandez, A survey of data mining and knowledge discovery process models and methodologies. The Knowledge Engineering Review, 2010. 25(02): p. 137-166.
[7] Chalaris, M., et al., Improving quality of educational processes providing new knowledge using data mining techniques. Procedia-Social and Behavioral Sciences, 2014. 147: p. 390-397.
[8] Mesquita, D.P., et al., Classification with reject option for software defect prediction. Applied Soft Computing, 2016. 49: p. 1085-1093.

[9] Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth, From data mining to knowledge discovery in databases. AI magazine, 1996. 17(3): p. 37.
[10] Aburrous, M., et al., Intelligent phishing detection system for e-banking using fuzzy data mining. Expert systems with applications, 2010. 37(12): p. 7913-7921.
[11] Kaur, P., M. Singh, and G.S. Josan, Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector. Procedia Computer Science, 2015. 57: p. 500-508.
[12] Osmanbegović, E. and M. Suljić, Data mining approach for predicting student performance. Economic Review, 2012. 10(1).
[13] Natek, S. and M. Zwilling, Student data mining solution–knowledge management system related to higher education institutions. Expert systems with applications, 2014. 41(14): p. 6400-6407.
[14] Blagojević, M. and Ž. Micić, A web-based intelligent report e-learning system using data mining techniques. Computers & Electrical Engineering, 2013. 39(2): p. 465-474.
[15] Al Deen, A., M. Nofal, and S. Bani-Ahmad, Classification based On Association-Rule Mining Techniques: A General Survey and Empirical Comparative Evaluation. Ubiquitous Computing and Communication Journal, 2011. 5(3): p. 9-17.
[16] Sokolova, M., N. Japkowicz, and S. Szpakowicz. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. in Australasian Joint Conference on Artificial Intelligence. 2006. Springer.
[17] Sokolova, M. and G. Lapalme, A systematic analysis of performance measures for classification tasks. Information Processing & Management, 2009. 45(4): p. 427-437.
[18] Fawcett, T., An introduction to ROC analysis. Pattern recognition letters, 2006. 27(8): p. 861-874.
[19] Keramati, A., et al., Improved churn prediction in telecommunication industry using data mining techniques. Applied Soft Computing, 2014. 24: p. 994-1012.
[20] Zhang, E. and Y. Zhang, F-measure, in Encyclopedia of Database Systems. 2009, Springer. p. 1147-1147.
[21] Hall, M., et al., The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 2009. 11(1): p. 10-18.

Bilal Sowan is currently an assistant professor at the Faculty of Information Technology, Applied Science Private University, Amman, Jordan. Dr. Sowan holds a Ph.D. degree in Computing from University of Bradford, UK. His research interests are in data mining and human computer interaction.