

The Most Efficient Classifiers for the Students' Academic Dataset

Ebtehal Ibrahim Al-Fairouz¹, Mohammed Abdullah Al-Hagery²

Department of Computer Science, College of Computer
Qassim University, Buraydah, Saudi Arabia

Abstract—Educational institutions contain a vast collection of data accumulated for years, so it is difficult to use this data to solve problems related to the progress of the educational process and also contribute to achieving quality. For this reason, the use of data mining techniques helps to extract hidden knowledge that helps in making the decisions necessary to develop education and achieve quality requirements. The data of this study obtained from the College of Business and Economics at Qassim University. Three of the classifiers were compared in this study Decision Tree, Random Forest and Naïve Bayes. The results showed that Random Forest outperforms other algorithms with 71.5% of Precision, 71.2% F1-score, and also it got 71.3% of Recall and Classification Accuracy (CA). This study helps reduce failure by providing an academic advisor to students who have weaknesses in achieving a high-Grade Point Average (GPA). It also helps in developing the educational process by discovering and overcoming weaknesses.

Keywords—Data mining; student performance; classification algorithms; evaluation

I. INTRODUCTION

Over the past few years, the world has seen rapid progress in technology. This progress led to the accumulation of information and data and its availability in all sectors systems, such as educational, health, social, and others. This data can be used to discover and analyze the obstacles and problems facing these sectors by employ data mining techniques [1], [2].

Data mining is considered as an interdisciplinary approach and an essential step in knowledge discovery [1]. It is used to extract useful and hidden information from large databases. Through the use of data mining tasks, it is also possible to answer questions that cannot be known through other techniques such as queries or reports[3]. The essential function of data mining is to use various algorithms such as classification, clustering, regression and association rules to discover hidden patterns that help in many important decisions [2]. The classification technique is a supervised learning task, the data in this method is classified into pre-defined classes. It is a frequently used method for creating models that are used to predict futuristic patterns. Examples of classification algorithms are Decision Tree, Naïve Bayes, Logistic Regression, K-nearest neighbor, Neural Networks, etc.[4]. The process of discovering knowledge involves several steps: collect data, clean data, pre-processing data, and then using data mining techniques.

This paper aims to use data mining techniques to examine student performance, the classification algorithms will be used

to classify the student's GPA through the use of historical data from the College of Business and Economics at Qassim University from 2014 to 2018.

Moreover, the importance of these results was situated in their practical application in educational institutions, as recommended, to use the best classification model on any academic data such as data of university students, institutes, schools, etc., for students' performance prediction. Furthermore, classify students in many aspects that assist the institution to enhance the educational process.

The remaining of the paper is divided into the following sections: the second section reviews the work related to this study, the third section the methodology of the study, the fourth section includes the results and the discussion, whereas the last section presents the conclusions.

II. RELATED WORKS

The research by Ramaphosa et al. [5] was about primary schools students from four cities in South Africa. The goal of their study was to recognize a predictive algorithm to detect learners' performance and make appropriate decisions for improvement. They analyzed the data using WEKA tool by employing classification algorithms namely Naive Bayes, BayersNet, J48 and JRip. They proved by their results that the J48 algorithm is the best model of prediction when compared to other algorithms by 99.13% classification accuracy. Ultimately, they reported that their study assists the schools in early discovering the academic performance of learners and enable stakeholders to improve the results of weak students.

According to the study by Abu Amrieh et al. [6], they found there is a relationship between the academic performance of the student and the behavior of the student (student interaction with the e-learning system). In their study, they used the dataset from Kalboard 360 which contains 500 records and 16 attributes. The student performance was predicted by applying classification algorithms which are Decision Tree, Artificial Neural Network and Naive Bayes by using WEKA tool. Besides, for improving classifier performance, they implemented ensemble methods namely Boosting, Random Forests and Bagging. Their results showed a robust relationship among academic performance and the behavior of the student, where the predictive model with behavioral attributes achieved higher accuracy than the predictive model without behavioural attributes. Furthermore, they observed an improvement in accuracy when they used ensemble techniques. Finally, they explained that this model

supports stakeholders in understanding students and identify weaknesses and develop their learning process in addition to reducing failure. Another study by Al-Noshan et al. concentrated on a set of important factors affecting the students' performance in the first year of the university [7], likewise Al-Rofiyee et al. [8]. Also, in [9], [10], [11] the authors compared the classifiers accuracy but using a medical dataset.

On the other hand, Rahman and Islam [12] applied four traditional classification algorithms which are K-NN algorithms, Naïve Bayes, Decision Tree and Artificial Neural Network algorithms. Besides, they used bagging ensemble method, boosting ensemble method and at last ensemble filtering technique, which helps extract hidden knowledge from student data that makes it easier for educational establishments to improve their quality of education. Their results indicated that the ensemble filtering technique obtained the best accuracy among all the algorithms.

Roy and Garg [13] applied the J48, Naïve Bayes, and MLP. The results showed that J48 obtained 73.92% accuracy which was the highest accuracy among the used algorithms. The objective of their study was to identify and predict the factors that affect student academic performance, where the performance of students can be affected by different attributes such as related to school, social and demographic.

The aim of the study by Guerra et al. [14] was to predict the performance of students in specific courses. In their study, they applied Decision Tree techniques on the dataset of IFMS in Brazil from 2012 to 2015 by utilizing WEKA tool. Their results showed that the J4.8 classification algorithm achieves the best results with cross-validation and pruning by 75.8%.

Ahmed and Elaraby [15] applied classification techniques to predict students' performance in the final assessment. They collected the dataset from the information system department from the year 2005 to 2010. The tool used in this work is WEKA by applying a Decision Tree algorithm (ID3). Through their results, they explained that their study help improves student performance as well as identify students who need the advice to guide them and make the appropriate decision.

The study by Tsiakmaki et al. [16] aimed to predict students' marks in the final exams of the courses of the second semester based on the first semester grades. They used a dataset from the Business Administration department of the TEI of Western Greece from 2013 to 2017 which contains 592 students. They only applied methods of regression using the WEKA tool, namely Linear Regression, Bagging, M5 algorithm, Gaussian processes (GPs), M5-Rules, Sequential Minimal Optimization (SMO), Random Forest and 5NN. The evaluation measure used in their study was MAE. After all the experiments they had done on the data set, they concluded that all the algorithms had achieved fair accuracy.

Pérez et al. [17] presented their initial results that prediction of attrition of students from a large dataset of Systems Engineering (SE) undergraduate students after six years of registration at a Colombian university, the dataset includes 762 students. In their study, they applied four algorithms which are Decision Tree, Random Forest, Naive

Bayes and Logistic Regression. Then, they found that SE courses performance is linked to mathematics and physics courses performance where they obtained the best AUC from Random Forest by 97% in the 3rd semester. These results showed them plainly that the courses which related to Systems Engineering have a dominant effect in predict dropout.

The objective of the study by Adekitan and Salau [18] was to perform predictive analysis to determine the final CGPA of graduation using their GPA of the previous three years as well as to define the class to which the student belongs at graduation. They applied six algorithms namely Decision Tree, Tree Ensemble, Random Forest, Naive Bayes, Logistic Regression and the Probabilistic Neural Network on the dataset which was gathered from Covenant University at Nigeria for the engineering students. The tools used in their study were KNIME and MATLAB. Their results demonstrated that the logistic regression obtained the best accuracy by 89.15%. Hence, they pointed out that students' results can be predicted the last year of their study using their performance in the previous three years. On the other hand, a few of studies included large datasets with records ranging from 14,333 records to 21,314 records.

Yulianto et al. [19] applied classification algorithms to student data to identify features that affect student achievement. Besides, they expected that the results of the analysis would be able to find the reasons that led to the delay of some students in the study period. They used two models of algorithms k-Nearest Neighbor and Decision Tree C4.5. They concluded that the k-Nearest Neighbor got better accuracy than the other. Quinn and Gray [20] used data from the Moodle to predict students' grades whether they will succeed or fail in the course. They applied the classification algorithms Random Forest, Gradient Boosting, k-Nearest Neighbours and Linear Discriminant Analysis using R. They summarized that the use of data from Moodle gives the ability to early detection of students at risk. The aim of the study by Walia et al. [21] was to build classification models to predict academic performance for students through the use of classification algorithms Naive-Bayes, Decision Tree, Random-Forest, JRip, and ZeroR. The results indicated that the school and study time were influential factors in the students' final grades.

A comparison of classification algorithms has been applied in several fields like emotion classification, precipitation, Spatial modelling of storm dust provenance etc. Fauziastuti et al. [22], classified students' graduation on time or overtime, by used two classification algorithms to compare their performance: Naive Bayes Classifier and K-Nearest Neighbor, but using a small dataset.

A similar research paper was in emotion classification to find the best classifiers amongst a set of classifiers [23], whereas in this paper we are concentrating on the extraction of the hidden knowledge embedded in the academic data of undergraduate students by a set of classifiers to find the best classifier for getting the hidden knowledge from this kind of data. The study achieved by Lazri et al. [24], focused on estimating precipitation from Meteosat Second Generation images, by combining six models of classification. They also used a linear regression model. Likewise, a study conducted by

Gholam et al. [25], was applied eight classification algorithms, for spatial maps to predict the source of dust in Khuzestan.

III. METHODOLOGY

The methodology used focuses on the use of classification algorithms in analyzing student performance to discover hidden patterns that help officials make the necessary decisions in the educational process.

The knowledge discovery process consists of four phases: data collection, pre-processing, data mining technique (classification), and interpretation of results, as in Fig. 1. The tool used in this study is the Orange data mining platform.

A. Data Collection

Data was collected from the College of Business and Economics from 2014-2018, which contains 72259 records for male and female students from several majors. The dataset contains the following attributes: Semester, Course code, Course name, CRD hours, Gender name, Entry date, Confirmed mark, Grade, Cumulative GPA (CGPA), Semester GPA (SGPA), Student status, Major name and Student level.

B. Data Preprocessing

Real data is usually incomplete and inconsistent due to individual errors or computer errors. Therefore, before starting to use data mining techniques, pre-processing of data is required. The process of data pre-processing includes first, clean the data from missing values, as the data was cleaned by the Orange program through the use of Impute widget. After cleaning the data, the number of records became 52,430. Second, data transformation, where the students' GPA was classifying into five categories as follows:

- 1) Excellent (GPA >=4.5)
- 2) Very Good (GPA >=3.75)
- 3) Good (GPA >=2.75)
- 4) Average (GPA >=2.00)
- 5) Fail (<2.00)

This classification was done using the Feature Constructor Widget on the Orange platform based on the CUM_GPA attribute and the new attribute was named as Class_GPA. The first class is Excellent, second class is Very Good, third class is Good, the fourth class is Average and the fifth class is Fail.

C. Data Mining Techniques (Classification)

The classification method is known as supervised classification, where, the data are organized into given known

classes. The dataset in classification is divided into a training dataset and test dataset. The classification algorithm is trained through a training dataset to build a model and test the model by test dataset since this model is used later to classify new data [26]. For example, predict students' performance by using the classification of GPA to good or bad. Algorithms used in classification such as Decision Trees, Random Forests or Bayes models. The classification techniques that were used in this study include Decision Tree, Naïve Bayes and Random Forest.

The classification algorithms are connected to Predictions widget to shows models predictions on the data. Hence, to evaluate the performance of models, we have focused on four different metrics which are CA, F1-score, Precision and Recall, given in Equations (1-4), where true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Moreover, the dataset was divided into a training set, and a test set by using a fixed proportion of data in Data Sampling widget, 75% of the data were used for the training set and 25% for the test set. The target variable is Class_GPA. Fig. 2 shows the model workflow of the classification task.

$$CA = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

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

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

Where,

TP: the model predicts correctly the positive class.

TN: the model predicts correctly the negative class.

FP: the model predicts incorrectly the positive class.

FN: the model predicts incorrectly the negative class.

Predictions widget was used to recognize model predictions on the data. Data Sampler widget is used to sample the data by using a fixed proportion of data. The dataset was divided into training data by 75% and test data by 25%. The data sample was sent to three algorithms widgets by Data Sampler widget so that they can produce the corresponding model; after that, the models were sent into Predictions widget while the remaining data was directly sent to Predictions widget.

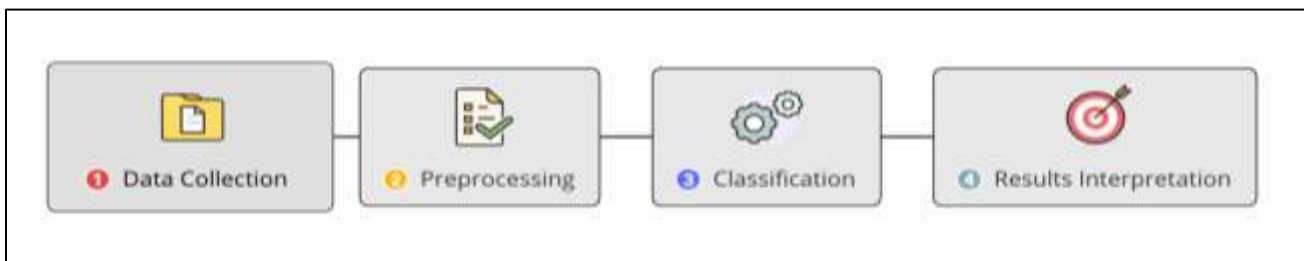


Fig. 1. Knowledge Discovery Process.

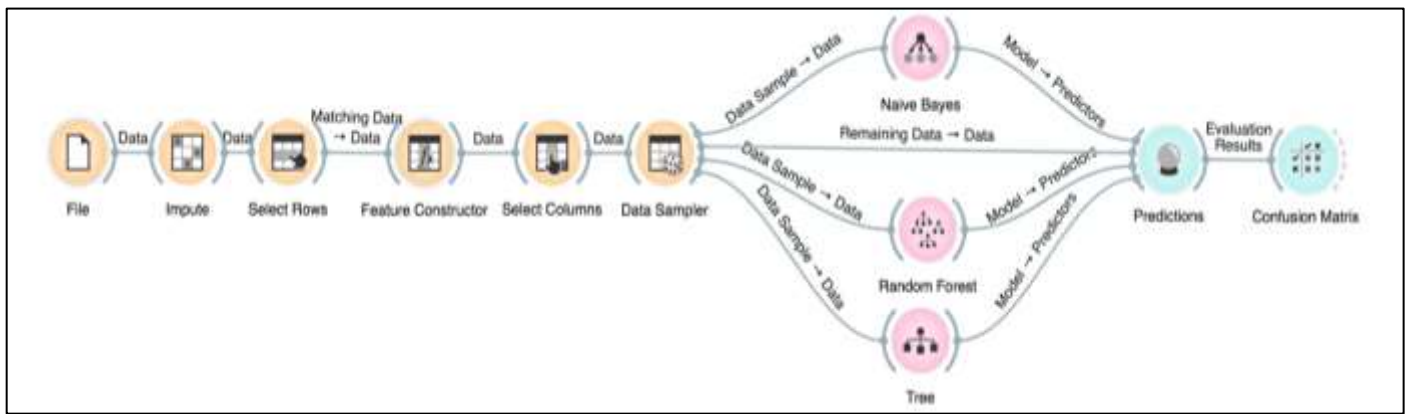


Fig. 2. The Model Workflow of Classification.

IV. RESULTS AND DISCUSSION

In this section, we presented the evaluation results of Decision Tree, Random Forest and Naïve Bayes. Fig. 3 shows the Predictions widget, which presented data with added predictions and the results of testing classification algorithms.

The widget received the dataset and then constructed a predictive model with Decision Tree, Random Forest and Naïve Bayes widgets, and it found the probabilities in predictions.

Table I presents the evaluation results of the classification. As we can see from the table, the Random Forest was the best classifier with 71.5% of Precision, 71.3% Recall and CA, and also it got 71.2% of F1-score. While the worst algorithm was the Naïve Bayes with 60.5% of Precision, 59.4% of CA and Recall, 59.5% of F1-score.

Confusion Matrix will be displayed that aims to assess the predictive performance of the models for each class to recognizing prediction of TP, FP, TN, and FN. The class labels are Excellent, Good, Acceptable and Fail. Table II, Table III, and Table IV illustrate the confusion matrices for Naïve Bayes, Random Forest and Decision Tree, respectively.

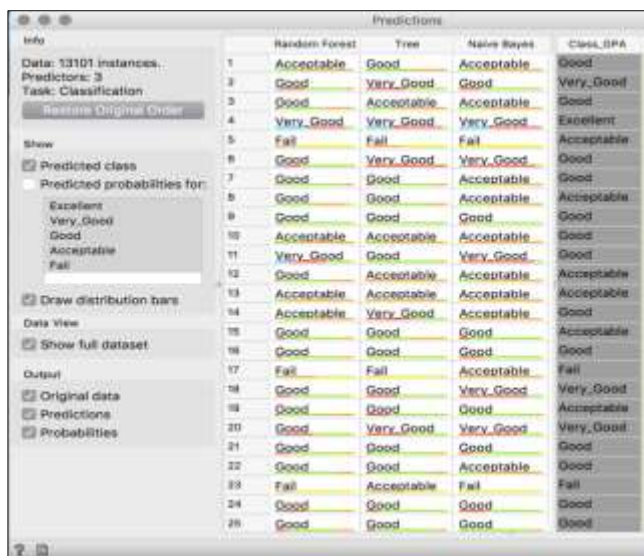


Fig. 3. Predictions Widget.

TABLE I. THE EVALUATION RESULTS OF PREDICTION

Model	CA	F1-score	Precision	Recall
Random Forest	0.713	0.712	0.715	0.713
Decision Tree	0.698	0.697	0.699	0.698
Naïve Bayes	0.594	0.595	0.605	0.594

TABLE II. CONFUSION MATRIX OF THE NAÏVE BAYES

		Predicted					Σ
		Model	Excellent	Very Good	Good	Acceptable	
Actual	Excellent	700	276	13	0	0	989
	Very Good	601	1254	615	28	3	2501
	Good	113	875	3200	916	136	5240
	Acceptable	4	62	1021	2021	491	3599
	Fail	0	0	3	159	610	772
Σ		1418	2467	4852	3124	1240	13101

TABLE III. CONFUSION MATRIX OF THE RANDOM FOREST

		Predicted					Σ
		Model	Excellent	Very Good	Good	Acceptable	
Actual	Excellent	617	344	27	1	0	989
	Very Good	162	1573	756	10	0	2501
	Good	10	412	4140	671	7	5240
	Acceptable	0	18	986	2468	127	3599
	Fail	0	0	7	208	557	772
Σ		789	2347	5916	3358	691	13101

TABLE IV. CONFUSION MATRIX OF THE DECISION TREE

		Predicted					Σ
		Model	Excellent	Very Good	Good	Acceptable	
Actual	Excellent	711	265	13	0	0	989
	Very Good	254	1593	641	13	0	2501
	Good	31	601	3989	608	11	5240
	Acceptable	2	41	1045	2316	195	3599
	Fail	0	0	18	218	536	772
Σ		998	2500	5706	3155	742	13101

Fig. 4 shows the evaluation of the three models Naïve Bayes, Random Forest and Decision Tree. The models were evaluated using four measures CA, F1-score, Precision and Recall Through the figure, we notice that the Random Forest outperform in all measures than other algorithms, followed by decision tree algorithm. As for the worst model was Naïve Bayes.



Fig. 4. Evaluation of the Models.

V. CONCLUSIONS

Educational institutions often require an analysis of student data to obtain useful knowledge that contributes to enhancing the learning process in addition to achieving quality in education. For this reason, data mining techniques were used to extract hidden knowledge from student data, and a comparison was made between three classifiers, Naïve Bayes, Random Forest and Decision Tree. Experimental results showed that Random Forest exceeded other classifiers with the accuracy of 71.3%, followed by the Decision Tree by 69.8%, then the last classifier was the Naïve Bayes by 59.4%. This study helps to know students' performance in advance by relying on previous results to improve their achievement in the future. Also, educational institutions must provide an academic adviser to failed students to enhance their academic performance.

ACKNOWLEDGMENTS

The authors would like to thank the College of Business and Economics at Qassim University that provided the data required for this research.

REFERENCES

- [1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Elsevier, 2012.
- [2] E. C. Abana, "A decision tree approach for predicting student grades in Research Project using Weka," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 7, pp. 285–289, 2019, doi: 10.14569/ijacsa.2019.0100739.
- [3] N. Rehman, "Data Mining Techniques Methods Algorithms and Tools," *Int. J. Comput. Sci. Mob. Comput.*, vol. 6, no. 7, pp. 227–231, 2017, [Online]. Available: <http://www.ijcsmc.com>.
- [4] A. H. Awlla, "Performance Analysis and Prediction Student Performance to build effective student Using Data Mining Techniques," *UHD J. Sci.*

- Technol.*, vol. 3, no. 2, p. 10, Jun. 2019, doi: 10.21928/uhdjt.v3n2y2019.pp10-15.
- [5] K. I. M. Ramaphosa, T. Zuva, and R. Kwuimi, "Educational Data Mining to Improve Learner Performance in Gauteng Primary Schools," in 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems, *icABCD 2018*, Aug. 2018, pp. 1–6, doi: 10.1109/ICABCD.2018.8465478.
- [6] E. A. Amrieh, T. Hamtini, and I. Aljarah, "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods," *Int. J. Database Theory Appl.*, vol. 9, no. 8, pp. 119–136, 2016, doi: 10.14257/ijta.2016.9.8.13.
- [7] A. Abdulrahman Al-Noshan, M. Abdullah Al-Hagery, H. Abdulaziz Al-Hodathi, and M. Sulaiman Al-Quraishi, "Performance Evaluation and Comparison of Classification Algorithms for Students at Qassim University," *Int. J. Sci. Res.*, vol. 8, no. 11, pp. 1277–1282, 2018, doi: 10.21275/ART20202907.
- [8] N. Al-Mufadi and M. A. Al-Hagery, "Using prediction methods in data mining for diabetes diagnosis," *Using Predict. Methods Data Min. Diabetes Diagnosis*, vol. 1, no. 4, p. 2014, 2014.
- [9] M. Abdullah Al-Hagery, A. Saleh Alfaiz, F. Suliman Alorini, and M. Saleh Althunayan, "Knowledge Discovery in the Data Sets of Hepatitis Disease for Diagnosis and Prediction to Support and Serve Community," *Int. J. Comput. Electron. Res.*, vol. 4, no. 6, pp. 118–125, 2015, [Online]. Available: <http://ijcer.org>.
- [10] S. Al-qarzaie, S. Al-odhaibi, B. Al-saeed, and M. Al-hagery, "Using the Data Mining Techniques for Breast Cancer Early Prediction," *Symp. Data Min. Appl.*, vol. 1, no. May, p. 2014, 2014.
- [11] M. Abdullah and H. Al-Hagery, "Classifiers' Accuracy Based on Breast Cancer Medical Data and Data Mining Techniques," *Int. J. Adv. Biotechnol. Res.*, vol. 7, no. 2, pp. 976–2612, 2016, [Online]. Available: <http://www.bipublication.com>.
- [12] M. Hasibur Rahman and M. Rabiul Islam, "Predict Student's Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques," in 2nd International Conference on Electrical and Electronic Engineering, *ICEEE 2017*, 2018, no. December, pp. 1–4, doi: 10.1109/ICEEE.2017.8412892.
- [13] S. Roy and A. Garg, "Predicting academic performance of student using classification techniques," in 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics, *UPCON 2017*, 2017, vol. 2018-Janua, pp. 568–572, doi: 10.1109/UPCON.2017.8251112.
- [14] M. S. Guerra, H. A. Neto, and S. A. Oliveira, "A Case Study of Applying the Classification Task for Students Performance Prediction," *IEEE Lat. Am. Trans.*, vol. 16, no. 1, pp. 172–177, Jan. 2018, doi: 10.1109/TLA.2018.8291470.
- [15] A. Badr, E. Din, and I. S. Elaraby, "Data Mining : A prediction for Student' s Performance Using Classification Method," *World J. Comput. Appl. Technol.*, vol. 2, no. 2, pp. 43–47, 2014, doi: 10.13189/wjcat.2014.020203.
- [16] M. Tsiakmaki, C. Pierrakeas, G. Kostopoulos, S. Kotsiantis, G. Koutsonikos, and O. Ragos, "Predicting university students' grades based on previous academic achievements," in 2018 9th International Conference on Information, Intelligence, Systems and Applications, *IISA 2018*, Jul. 2019, pp. 1–6, doi: 10.1109/IISA.2018.8633618.
- [17] B. Pérez, C. Castellanos, and D. Correal, "Predicting student drop-out rates using data mining techniques: A case study," in *Communications in Computer and Information Science*, 2018, vol. 833, pp. 111–125, doi: 10.1007/978-3-030-03023-0_10.
- [18] A. I. Adekitan and O. Salau, "The impact of engineering students' performance in the first three years on their graduation result using educational data mining," *Heliyon*, vol. 5, no. 2, p. e01250, Feb. 2019, doi: 10.1016/j.heliyon.2019.e01250.
- [19] L. D. Yulianto, A. Triayudi, and I. D. Sholihati, "Implementation Educational Data Mining For Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5," *J. Mantik*, vol. 4, no. 1, pp. 441–451, May 2020.
- [20] R. J. Quinn and G. Gray, "Prediction of student academic performance using Moodle data from a Further Education setting," *Irish J. Technol.*

Enhanc. Learn., vol. 5, no. 1, pp. 1–19, Oct. 2019, doi: 10.22554/ijtel.v5i1.57.

- [21] N. Walia, M. Kumar, N. Nayar, and G. Mehta, “Student’s Academic Performance Prediction in Academic using Data Mining Techniques,” *SSRN Electron. J.*, pp. 1–5, Apr. 2020, doi: 10.2139/ssrn.3565874.
- [22] V. T. Fauziastuti and L. A. Rakhman, “A Review of Students’ Graduation Classification: A Comparison of Naive Bayes Classifier and K-Nearest Neighbour,” in *1st International Multidisciplinary Conference on Education, Technology, and Engineering (IMCETE 2019)*, Mar. 2020, pp. 219–221, doi: 10.2991/assehr.k.200303.052.
- [23] M. AbdullahAl-Hagery, M. AbdullahAl-Assaf, and F. MohammadAl-Kharboush, “Exploration of the best performance method of emotions classification for arabic tweets,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 19, no. 2, pp. 1010–1020, 2020, doi: 10.11591/ijeecs.v19.i2.pp1010-1020.
- [24] M. Lazri, K. Labadi, J. M. Brucker, and S. Ameer, “Improving satellite rainfall estimation from MSG data in Northern Algeria by using a multi-classifier model based on machine learning,” *J. Hydrol.*, vol. 584, p. 124705, May 2020, doi: 10.1016/j.jhydrol.2020.124705.
- [25] H. Gholami, A. Mohamadifar, and A. L. Collins, “Spatial mapping of the provenance of storm dust: Application of data mining and ensemble modelling,” *Atmos. Res.*, vol. 233, p. 104716, Mar. 2020, doi: 10.1016/j.atmosres.2019.104716.
- [26] R. Lawrance and V. Shanmugarajeshwari, “An assay of teachers’ attainment using decision tree based classification techniques,” in *Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2017*, Apr. 2017, pp. 1–6, doi: 10.1109/ICCPCT.2017.8074382.

AUTHORS’ PROFILE

Ebtehal Ibrahim Al-Fairouz: received her BSc in Computer Science from the Qassim University, Buraydah, KSA. She is a teaching assistant in the Department of Management Information System (MIS) at the College of Business and Economics (CBE) and a Master’s student in Computer Science Department, Qassim University, KSA. Her research interests include data mining, data analytics, data visualization and machine learning.



Mohammed Abdullah Al-Hagery: received his BSc in Computer Science from the University of Technology in Baghdad Iraq-1994. He got his MSc. in Computer Science from the University of Science and Technology Yemen-1998. Al-Hagery finished his Ph.D. in Computer Science and Information Technology, (Software Engineering) from the Faculty of Computer Science and IT, University of Putra Malaysia (UPM), November 2004. He was the head of the Computer Science Department at the College of Science and Engineering, USTY, Sana’a from 2004 to 2007. From 2007 to this date, he is a staff member at the College of Computer, Department of Computer Science, Qassim University, Buraydah, KSA. He published more than 31 papers in various international journals. Dr. Al-Hagery was appointed the head of the Research Centre at the Computer College, and a council member of the Scientific Research Deanship Qassim University, KSA from September 2012 to October 2018. Currently, he is teaching the master degree students and a supervisor of four master thesis. He is a jury member of several PhD and master thesis, as an internal and external examiner in his field of his specialist.