Evaluating the Performance of Engineering's Students in Mathematic Subject based on Academic Decision-Making Techniques

Abbas Atwan Mhawes
Ministry of Education, Baghdad, Iraq.
E-mail: abbas_atwan@gmail.com

Ahmed Yousif Falih Saedi
Department of Computer Engineering, Al-Mustansiriyah University, Baghdad, Iraq.
E-mail: ahmed.yousif@uomustansiriyah.edu.iq

Ali Talib Qasim Al-Aqbi
Department of Computer Engineering, Al-Mustansiriyah University, Baghdad, Iraq.
E-mail: lali.al_aqbi@uomustansiriyah.edu.iq

Lamees Abdalhasan Salman
Department of Computer Engineering, Al-Mustansiriyah University, Baghdad, Iraq.
E-mail: lameesiteng@uomustansiriyah.edu.iq

Received March 14, 2021; Accepted July 08, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18I2/WEB18313

Abstract

Data mining is characterized as a quest for useful knowledge via large quantities of data. Some basic and most common techniques for data extraction are association rules, grouping, clustering, estimation, sequence modeling. For a wide range of applications, data mining techniques are used. Techniques of data analysis are essential to the preparation and implementation of the administration of the learning system, including behavioral guidance and personal behavior appraisal. The article applies data analytical methods to the role of student classification. Several tests are used for the interpretation of the findings. In keeping with the methodology proposed in the paper, the classification using cognitive skills provides more detailed results than the findings of other study published. Five algorithms were used (J48, Naïve Bayes, Multilayer Perception, K Star and SMO). This essay discusses and measures the application of the various algorithms so that factors affecting the success and failure of students can be identified, student performance can be estimated, and the significant consequences of the mathematics system for the second university year can be identified. However the number of exams can be minimized using data mining techniques. In terms of time and consequences, this shortened analysis plays a key role.

Keywords

EDM, J48, Naïve Bayes, Multilayer Perception, K Star, SMO, Data Mining.

Introduction

A huge amount of data can be collected and analyzed using algorithms that are technologically assisted learning environments. Learning analytics roles are important to the scheduling and adjustment of learning system management, including adaptable guidance and personal learning behavior assessment. In the discovery of information and data mining, there are many different methods for identification used. Every technique or method has its own advantages. The article applies methods of data analysis to the student classification task. Several assessments are used to analyze the results achieved. Classification using the features of cognitive abilities according to the framework suggested in the article gives more detailed outcomes than the findings reported in other studies published by available sources.

Actual changes in academic digitalization contribute to the processing of large amounts of data that show correlations that can be used to make decisions to improve education performance. Intelligent Analysis Data (IAD) is the exploration of such useful data patterns. Intelligent Analysis of Educational Data mining (EDM) was developed as part of the experience in the implementation of IAD methods in education in 2008 in quite a new direction. For example, studies in the analysis of educational data are: the determination of relationships between utilization ratings and additional student performance, assessment of the quality of courses based on a student survey, and quality assessment of classroom papers.

Previous Studies

Educational Data Mining (EDM) develops strategies for academic decision-making studies. In order to find different patterns in the data DM and EDM are involved. However, the presence of mining is what makes EDM different. EDM tries to improve the educational process by all means, guide students in the right direction, and recommend teaching staff, in addition, to examine the essence of educational phenomena to understand how we nevertheless absorb information and gain skills and skills. EDM approaches are listed as statistical techniques, analysis of interaction methods, structure seeking and discovery methods for models and human governance data distillation.

In 2017, Rahman, Sultan, and others gathered extensive data on the quality of a cohort of 20 students from Jubail University College in the Kingdom of Saudi Arabia to discover students at risk of failure at the final exam. In the Student Assessment Results Report, use the decision-making tree of the data mining classification system to identify students who need help to perform well in the final exam. There have been two methods of selecting that feature: obtaining the information and the percentage of gain and using it to select the gain ratio (Atta-Ur-Rahman *et al.*, 2018). In 2015, Jacob, Sultan and others explored several strategies for the processing of academic information and how they can be used to support all educational system participants. Regression analysis was used to build a model containing a dependent variable and multiple separate variables. Techniques in data mining such as regression and decision-making tree have been studied and applied to forecast academicals success and to meet successful student performance forecasts and to predict academic failure. Assembling students was successfully used in classes based on academic strengths and weaknesses (Jacob *et al.*, 2016). In 2017 Rahila Umer and Suriadi Suriadi suggested a method to make early assessments to improve student learning experience in open-line courses. The research employed four methods for classifying machine learning (Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF) and K neighbor). The study includes four approaches of identification. By integrating system mining functions and standard features, full accuracy (F1 score, region below the curve) of computer training techniques can be increased (Umer *et al.*, 2017). In 2015 Al-Saleem and Al-Kathiry and others suggested the template that estimates student performance on the basis of the number of former students in core and elective courses in future courses. The algorithms for Model ID3 and J48 decision tree classification are used; the advantages of the system and the performance of certain algorithms are presented. They found that J48 was more accurate than the ID3 algorithm. The J48 algorithm was therefore used to forecast digitally, so that one or more loops of forecasts could be made for students (Al-Saleem *et al.*, 2015). In 2019, Hashemineiad and Sarvmili proposed to extract the hidden rules which can be used for the prediction of students ' final results by the classification method based on the S3PSO algorithm. Moodle results from the University of Tarbiat, Tehran, Iran, have been included in S3PSO performance evaluation in computer science courses. Compared to other classification methods including SVM, KNN, Naif Bayes, Neural Network, and APSO, S3PSO increased 9 per cent of the Moodle dataset's accuracy quality and received promising results for the estimation of final score for its learners. Experimental results suggest that S3PSO's detecting bases quality in assistance, reliability and inclusivity measures is higher than that of another rule-based identification system like C4.5, ID3 and CART (Hasheminejad and Sarvmili, 2018). In 2018 Kiu proposed a framework for extracting supervised academic information. The decision tree model J48 was used in order to combine students into the final grade forecast, based on the learning process and events taken from the event logs of the online learning program. He used a Weka data mining tool to explore student learning with simple and effective visuals. Experimental results demonstrate that the models of the workbook and the identification rules set forth in the J48 decision-making tree are reliable for student performance. It provides a way of making useful advice on the enhancement of the education of students and the quality of academics, both for students with and without training (Kiu, 2018).

In our study, we predicted the results of second-stage mathematics students using a number of algorithms to achieve the best expectations among them. This article presents an experimental study of J48, Naive Bayes, and Multilayer perception, K Star, and SMO classification algorithms and compared among these classifications to prediction Student performance in mathematics according to the attributes taken. The experimental results show a satisfactory performance of K star algorithm in comparison with the other approaches in pred.

The exploration of the reasons of academic failure in a specific studying subject is a major problem to identify and describe pupil learning difficulties. In the early stages of the learning process, the research seeks to find students with disabilities during school as soon as possible. The goal is to identify student groups according to their probability of academic failure and success.

Collecting Data Method

In this study, we have analyzed data in the subject mathematic obtained from department of computer engineering, Mustansiriya University, from the students of the Second stage the data was analyzed. The data set is made up of attributes such as living with parents, having a job, using social media, time for learning math, etc. In total we have collected data from 84 students, 60 of whom have completed the tests, and the others have failed.

When the data set is pre-processed, irrelevant values are discarded, the missing value is refilled and the outer value on the outer samples are deleted / refilled. The properties of the datasets we used in our study are shown in Table 1.

Table 1 Dataset Attributes

Attribute	Data Type	Attribute Rule
Math absence 1C1	Integer	Regular
Math absence 1C2	Integer	Regular
Math absence 2C1	Integer	Regular
Math absence 2C2	Integer	Regular
Math 1SC1	Integer	Regular
Math 1FC1	Integer	Regular
Math 1SC2	Integer	Regular
Math 1FC2	Integer	Regular
Math 2SC1	Integer	Regular
Math 2FC1	Integer	Regular
Math 2SC2	Integer	Regular
Math 2FC2	Integer	Regular
Sex	Binominal	Regular
Living with parents	Binominal	Regular
Dou have a job	Binominal	Regular
Family members	Integer	Regular
Seq.inthefa1ily	Integer	Regular
Using social1edia	Integer	Regular
No. hours Mathstudy1	Integer	Regular
Extra Materia l Math 1	Binominal	Regular
Electronic education Math 1	Binominal	Regular
No. hours Math study 2	Integer	Regular
No. hours Math study 2	Integer	Regular
Extra Material Math2	Binominal	Regular
Electronic education Math 2	Binominal	Regular
No. hours Control study C1	Integer	Regular
Extra Material Control C1	Binominal	Regular
Electronic education Control C1	Binominal	Regular
Class	Binominal	Label

Classification Algorithms

Classification algorithm used mainly in historical data-based predictions. Classification is a prediction technique that is monitored in nature. This methodology is capable of predicting the tag for classes if sufficient numbers of examples of learning can be identified.

A variety of classification algorithms can be used to classify a number of models for the prediction of unknown category tags. The dataset is grouped into two groups: a training set (dependent set) and a test set (dependent set). The machine learning algorithm is initially based on training, which is followed later by the test set by the predicting model.

Labeling usually refers to data elements being allocated to predefined classes and groups (Dunham, 2006). Also known as supervised learning, it includes classification and learning. The training data is analyzed in the learning phase by the classification algorithm; test information is used in the classification phase to estimate the exactness of the classification rules (Taruna and Pandey, 2014). The common algorithms discussed in this paper are:

J48 Algorithm

J48 is used for both classification and prediction operations. For classification, J48 (based on the C4.5 algorithm of machine learning) was selected given that this algorithm is one of the most used tools in Weka which provides stability between accuracy, speed and interpretability of results. In addition, this algorithm categorizes data into a decision tree in which we can easily identify vulnerable students. Classification learning as part of EDM was also implemented to predict student performance (Monika and Vohra, 2012).

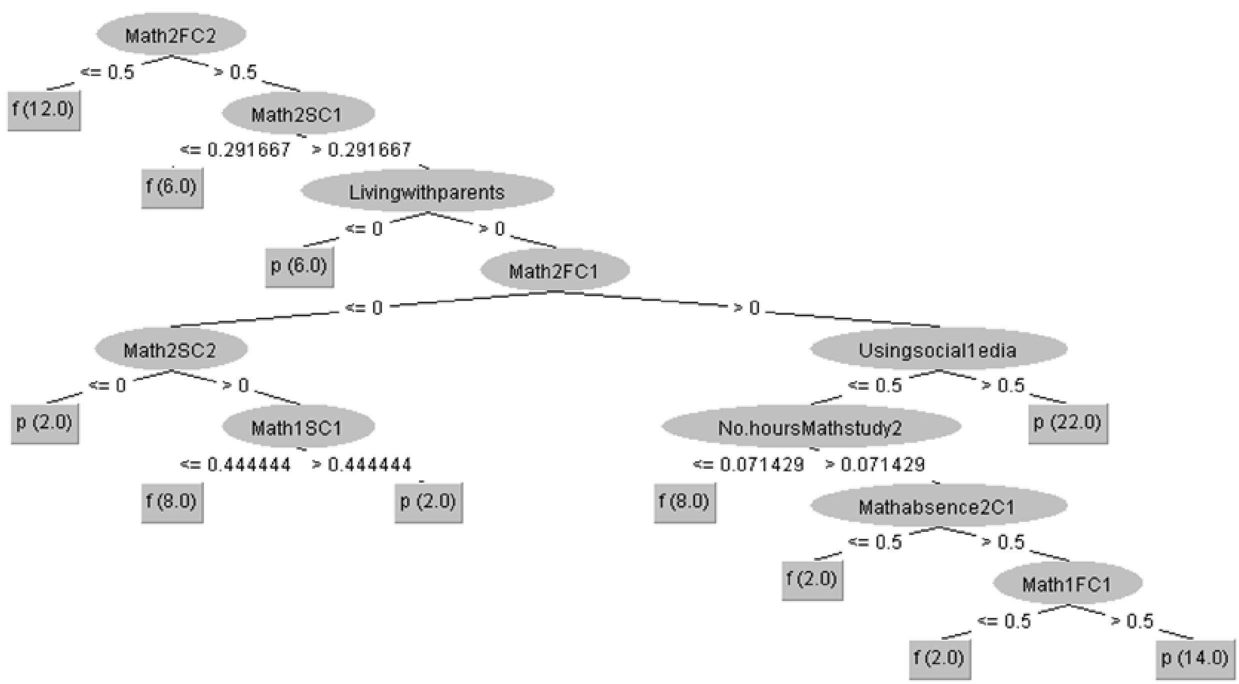

Fig. 1 J48

Naive Bayes

Naive Bayes is a most effective and efficient inductive algorithm for learning machines and data mining. It is surprisingly competitive in category, as the contingent presumption that it is built on independence is rarely true for real world implementations. The dependence between attributes can cancel each other with appropriate and required conditions for the optimality of naive Bayes (Zhang, 2004). The classifiers of Naive Bayes are highly scalable and allow a range of linear parameters for a learning problem in terms of the amount of variables (features and predictors).

$$P(C_k|X) = \frac{P(C_k)*P(X|C_k)}{P(X)} \quad (1)$$

Where:

P(c|x) is the posterior probability of class (c, target) given predictor (x, attributes).

P(c) is the prior probability of class.

P(x|c) is the likelihood which is the probability of predictor given class.

P(x) is the prior probability of predictor.

Multilayer Perception

Multilayer perceptron's (MLPs) are perception networks, linear classification networks. A unique group of sensor nodes connecting inside network make up decision (Saleh, Al-Aqbi and Saedi, 2018; Falih *et al.*, 2020; Saleh *et al.*, 2021). We can, in addition, use "hidden layers" to enforce arbitrary decision limits. Multilayer perceptron's (MLPs), a feed-forward neuronal processing network that are clustered in layers and linked by weaved links, are the most common neural network type. Weka is fitted with a graphical interface to construct a network system with as many perceptions and links.

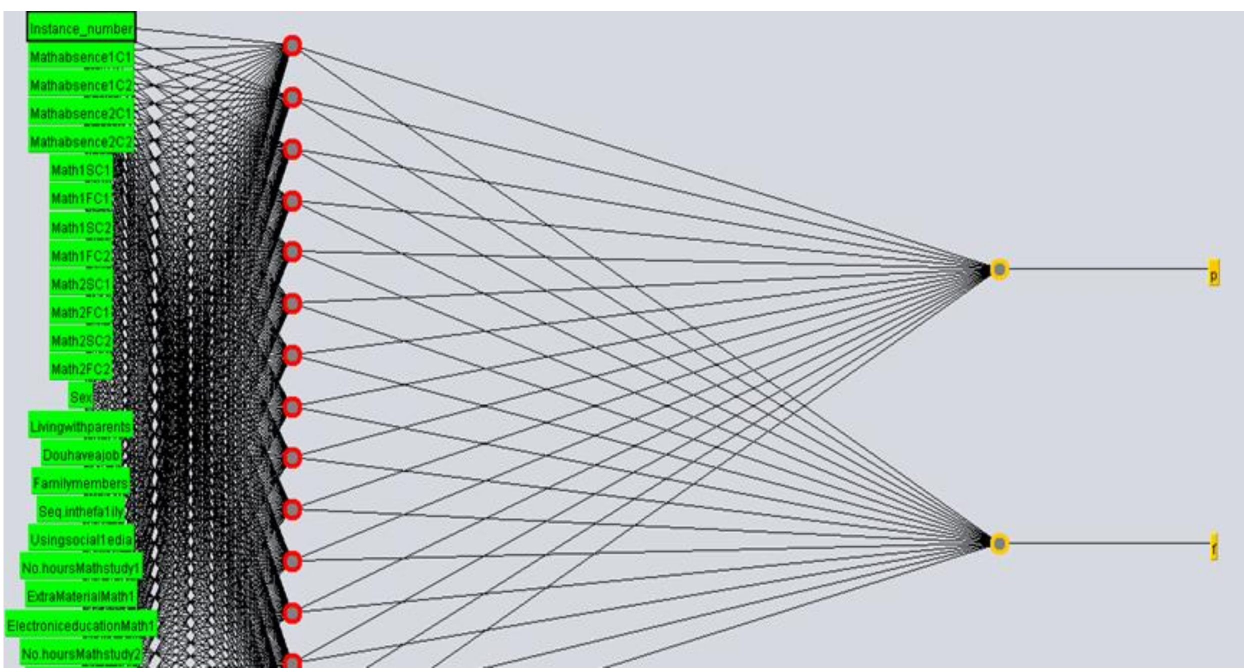

Fig. 2 Multilayer Perceptron

K Star

K-star is a classifying case. The evaluation instance's class is based on the similar training instances as defined by a function of similarity. The distinction is that it uses a computational complexity-based distance function from other instance-based classes. Instance-based students define an instance by matching it with a pre-classified sample set. The basic premise is that similar cases are treated in similar ways. The problem is how "similar case" and "similar classification" should be described. A distance function determining how similar both instances are and the classified function which indicates how example similarities give the final categorization of the instance is the correlating components of a case-based learner (Cleary and Trigg, 1995).

SMO

SMO is a simple algorithm which can solve SVM QP problem quickly without additional storage matrixes and without the use of numerical steps to optimize QP. At every step, SMO tries to solve the least possible problem of optimization. SMO chooses two Lagrange multipliers at every point in the process to function together to maximize them, determine optimum values for these multipliers and upgrade the SVM to represent optimized new values.

Methodology

In our article, three algorithms has been used. The WEKA method is used to apply perception algorithms. WEKA is an open source java software that researchers at Walkato University for New Zealand have developed. This offers several different algorithms for machine learning, including Decision Tree, Naïve Bayes, MLP and more. We used a data set with known output values and use this data set to build a model to generate the model with the aid of a training data model. Nevertheless, this type of system uses a complete trainings package and splits it into two data sets, that is to say that data set are taken and put into our training set for the project. The second data set is then placed in a data set, and we use this data immediately after the model was generated to test the consistency of model. In WEKA the following steps are also to be taken to introduce the two algorithms:

Step-1: First, import and preprocess the dataset into the device.
Step-2: Classify data with the classification algorithm of Naïve Bayers.
Step-3: Remove previous buffer value and predict with the J48 algorithm decision tree.
Step-4: Use J48 decision tree algorithm to classify the data. Split using the same percentage split. And so on with third algorithm.

In comparison, the output of K Star algorithm is compared to Naïve Bayes, J48, Multilayer Perception, and SMO. After evaluating the performance of all five algorithms, the K Star algorithm has been hypothesized to be more powerful than other algorithms. K Star algorithms scored mean absolute error 0.07%, with Multilayer perceptron, J48, Naïve Bayes, and SMO only rating 11%, 18%, 35%, and 27% respectively.

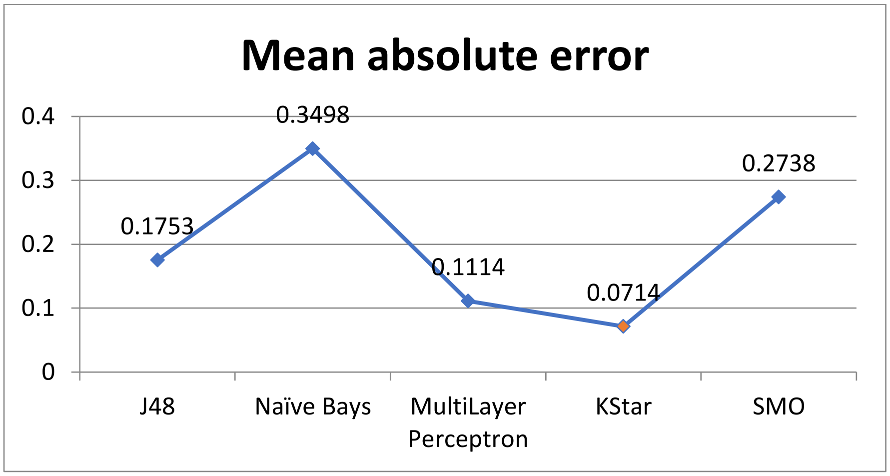

Fig. 3 Mean absolute error

Results

The Weka Tool provides unified algorithms, which help to apply different classifiers in a simple and flexible way. Five algorithms are used before and after deletion of attributes in this stage (J48, Naïve Bayes and Multilayer perception, K Star and SMO). The process of elimination of the attributes is very effective to explore the effectiveness and reliability of these attributes. Output information for the Five attributes of the algorithms (J48, Naïve Bays and MultiLayer Perceptron, K Star and SMO) namely True Positive (TP), False Positive (FP), Precision and Recall, are given. Three output metrics are used to test the model obtained: accuracy, precision and recall. Accuracy refers to the percentage in the test sample of correctly classified data. Precision tests the proportion of data that are actually good listed in the formula. Recall the calculation of the true positive rate of acceptance, and they measured as follows:

$$\text{Accuracy} = (TP + TN) / (P + N) \qquad (2)$$

$$\text{Precision} = TP / P \qquad (3)$$

$$\text{Recall} = TP / (TP + FN) \qquad (4)$$

Where, P means the number of Pass records. N means the number of Fail records. TP, TN are record numbers that correctly classified respectively as Pass and Fail. FN is the misclassified records as Fail. The assessment criteria for each model are summarized in Table 2 & 3. The results show the most accurate model based on the Naïve Bays algorithm.

Table 2 Simulation Result of Each Algorithm

Algorithm (Total instances, 48)	TP Rate	FP Rate	Precision	Recall	Class	Accuracy
J48 Algorithm	0.804	0.184	0.841	0.804	P	81 %
	0.816	0.196	0.775	0.816	F	
Naïve Bays	0.717	0.368	0.702	0.717	P	68 %
	0.632	0.283	0.649	0.632	F	
MultiLayer Perceptron	0.913	0.132	0.894	0.913	P	89 %
	0.868	0.087	0.892	0.868	F	
KStar	0.913	0.053	0.955	0.913	P	93 %
	0.947	0.087	0.900	0.947	F	
SMO	0.761	0.316	0.745	0.761	P	72 %
	0.684	0.239	0.703	0.684	F	

Table 3 Comparison of classification algorithms

Algorithm	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Total Number of Instances
J48 Algorithm	80.9524%	19.0476%	0.6173	84
Naïve Bays	67.8571%	32.1429%	0.3498	84
MultiLayer Perceptron	89.2857%	10.7143 %	0.7833	84
KStar	92.8571%	7.1429%	0.8565	84
SMO	72.619 %	27.381%	0.4461	84

Conclusion

This experiment seeks to investigate and test the multiple algorithm implementation process in order to determine the factors that affect the success and failure of students. The best solution to forecast the success of students can be data mining algorithms as they provide high precision and a route map for both academic and pupils. On the basis of the tests, multiple variables (attributes) can influence the quality of the perception outcome and the academic overall performance of students.

In this paper, we applied the five classification algorithms to the records of 84 students from the computer engineering department at the Bachelors program. The goal was to predict student performance and to recognize important impacts of the second-year undergraduate mathematics curriculum. Our observations have shown that the value classification model in the K Star algorithm is the best one as shown in figure 4 below. Attributes such as Math absence, Work, Family members, Sequence in the family, Using social media, No. hours Math study had less effect on student success, whereas GPA, Living with parents, Extra Material Math, and Electronic education Math had the most significant effect on the final class. The results showed the methods used were capable of predicting the students ' success at an early.

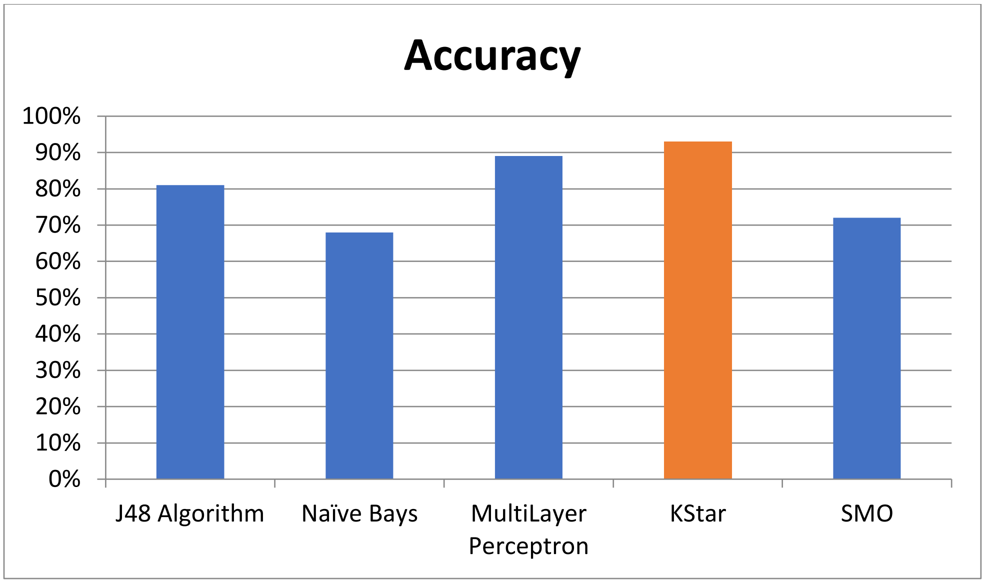

Figure 4 Accuracy Result of Each Algorithm

Our work contributed to investigate and test the implementation of multiple algorithms and identify the factors that affected students' success and failure. We proved through experiments that it is possible to predict student success through data-mining algorithms which provided for both academics and students with the benefit of focusing on some features through their impact on the quality of students' general academic performance.

Our work was distinguished from the previously mentioned studies in explaining that features such as absence from mathematics lectures, work during study, family members, use of social media, and number of hours had less impact on students' success. We also explained that features such as GPA, and living with parents, special lessons in mathematics, and adding e-learning in mathematics were more significant impact on the final grade.

Acknowledgments

The authors are extremely grateful to the University of Mustansiriyah / Engineering Faculty / Computer Engineering Department for their collaborative and providing us with dataset.

References

Al-Saleem, M., Al-Kathiry, N., Al-Osimi, S., & Badr, G. (2015). *Mining educational data to predict students' academic performance. In International Workshop on Machine Learning and Data Mining in Pattern Recognition*, Springer, Cham, 403-414. https://doi.org/10.1007/978-3-319-21024-7_28

Atta-Ur-Rahman, K.S., Aldhafferi, N., & Alqahtani, A. (2018). Educational data mining for enhanced teaching and learning. *Journal of Theoretical and Applied Information Technology, 96*(14), 4417-4427.

Cleary, J.G., & Trigg, L.E. (1995). ‘K*: An Instance-based Learner Using an Entropic Distance Measure’. *In Machine Learning Proceedings 1995,* 108–114. http://doi.org/10.1016/b978-1-55860-377-6.50022-0

Dunham, M.H. (2006). Data Mining-Introductory and Advanced Concepts. *2nd IEEE International Conference on Information Management and Engineering, 1,* 190–192.

Falih, A.Y., Qasim, A.T., Saleh, B.J., & Tahlok, W.S. (2020). Design and Implementing System to Study Environmental Noise Pollution Using Microcontroller. *Indian Journal of Forensic Medicine & Toxicology, 14*(2), 721–725. http://doi.org/10.37506/ijfmt.v14i2.2946

Hasheminejad, S.M., & Sarvmili, M. (2019). S3PSO: Students’ performance prediction based on particle swarm optimization. *Journal of AI and Data Mining, 7*(1), 77-96. http://doi.org/10.22044/jadm.2018.5506.1662

Jacob, J., Jha, K., Kotak, P., & Puthran, S. (2015). Educational data mining techniques and their applications. *In International Conference on Green Computing and Internet of Things (ICGCIoT),* 1344-1348. IEEE. http://doi.org/10.1109/ICGCIoT.2015.7380675

Kiu, C.C. (2018). Supervised Educational Data Mining to Discover Students’ Learning Process to Improve Students’ Performance. *In Redesigning Learning for Greater Social Impact,* 249–258. http://doi.org/10.1007/978-981-10-4223-2_23

Monika, G., & Vohra, R. (2012). Applications of Data Mining in Higher Education. *International Journal of Computer Science Issues, 9*(2), 113–120.

Saleh, B.J., Saedi, A.Y.F., Al-Aqbi, A.T.Q., & Abdalhasan Salman, L. (2021). Optimum Median Filter Based on Crow Optimization Algorithm. *Baghdad Science Journal, 18*(3), 614-627. http://doi.org/10.21123/BSJ.2021.18.3.0614

Saleh, B.J., Al-Aqbi, A.T.Q., & Saedi, A.Y.F. (2018). A novel biogeography inspired trajectory-following controller for national instrument robot. *In Communications in Computer and Information Science,* 171–189. http://doi.org/10.1007/978-3-030-01653-1_11

Taruna, S., & Pandey, M. (2014). An empirical analysis of classification techniques for predicting academic performance. *In IEEE International Advance Computing Conference (IACC),* 523-528. http://doi.org/10.1109/IAdCC.2014.6779379

Umer, R., Susnjak, T., Mathrani, A., & Suriadi, S. (2017). On predicting academic performance with process mining in learning analytics. *Journal of Research in Innovative Teaching & Learning, 10*(2), 160–176. http://doi.org/10.1108/jrit-09-2017-0022

Zhang, H. (2004). The optimality of Naive Bayes. *In Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2004,* 562–567.

Janavi, E., Nadi-Ravandi, S., & Batooli, Z. (2020). Impact of researchgate on increasing citations and usage counts of hot papers in clinical medicine indexed in web of science. *Webology, 17*(1), 130-139.