

Analyzing Factors Affecting Academic Performance of Postgraduate Students using Data Mining Techniques

Abraham E. Ewwiekpaefe, Ph.D.*; Muhammad M. Isa, B.Sc.;
and Fiyinfoluwa Ajakaiye, M.Sc.

Department of Computer Science, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria.

E-mail: aeevwiekpaefe@nda.edu.ng*

ABSTRACT

Data mining is the process of knowledge extraction from large data repositories. Its several techniques and algorithms have been applied in various fields to extract hidden information and to find patterns and relationships among the data. The aim of this work is to predict the academic performance of postgraduate students using data extracted at the time of their admission and to further determine which classifier gives the best accuracy. The classification techniques used in this research are J48 decision tree, Naïve Bayes and PART rule-based algorithms with the aid of the WEKA tool. Cognitive, non-cognitive and demographic data of the students were used for the data analysis.

The analysis was carried out in two experiments. The first experiment was done without feature selection while the second experiment was conducted with feature selection. After feature selection, J48 classifier emerged the best with an accuracy of 69.7% when compared with Naïve Bayes and PART classifiers. The results of this research can positively influence the admission process of the Postgraduate School of the Nigerian Defence Academy (NDA), Kaduna, Nigeria.

(Keywords: academic performance classifier, data mining, feature selection, NDA, Nigeria, WEKA).

INTRODUCTION

The academic performance of university students and ascertaining the factors that influence such are vital concerns for educational institutions, academics, and students. Academic performance can be described as how well a person completes the set goals of an academic institution based on some predefined standards. Students' academic

performances are affected by several factors such as personal, socio-economic and other environmental variables. Knowledge about these factors and their effects on student performance can help manage these effects (Yassein, et al., 2017).

Data is generated on a daily basis and in large quantities from different organizations across various walks of life. Sectors where these voluminous data are constantly generated include schools, manufacturing, e-commerce, medicine, insurance, fraud detection, and bioinformatics (Badr, et al., 2016). A large array of data about students, courses offered and grades are readily available in schools. Data mining techniques can be applied to such data to acquire high level information. This information can be used to improve students' performance, retention and avert attrition of at-risk students.

Educational data mining is a feature of data mining where the main concern is to construct models for extracting hidden knowledge from students' data which may help to improve students' academic performance (Hasibur and Rabiul, 2017). Attributes from students' data will help determine which unique feature of a student will have an impact on his/her performance. Prediction and analysis of student academic performance is vital for student academic progress (Mueen, et al., 2016).

In Nigeria, higher educational institutions set criteria for entry of students and for the award of final certificate of completion into its postgraduate programs. Entry criteria are often set to ensure certain standards in the students' performance, retention of the students and to reduce students' attrition. For instance, entry requirements into most science postgraduate studies include obtaining a minimum of second-class lower division with a CGPA of 3.0/5.0 for an academic

Masters program for candidates with Bachelor's degrees from an approved university. Also, candidates must have five credit passes including English, Mathematics and three other relevant subjects at 'O' Level. These criteria determine who gets admitted into the institution. This helps to reduce the danger of admitting under qualified candidates as well as ensure that admitted students complete their studies successfully with good academic performance.

Despite these checks some PG students still abandon their course of study, sometimes after the first year, with no hope of continuing the program. Others are dazzled with a number of carry-overs thereby extending the duration of their program study. Again, some PG students struggle with their research thesis for a number of years and end up being withdrawn from such programs. A situation that often culminates in a lose-lose scenario for both parties – the students and the institution. Eventually, affected students often end up not completing the programs and achieving their dreams and the institutions, on the other hand loose on the part of revenue generation and credibility.

Therefore, to maximize students' performance and to reduce attrition in higher educational institutions, the study seeks to identify indicator variables for the prediction of postgraduate students' performance. This, the study believes, can further help to determine who gets admitted and what special intervention is required, by both the institution and students, to improve academic performance.

LITERATURE REVIEW

Asif, et al. (2017) applied data mining techniques in predicting the performance of students after graduation. In the work, the Naïve Bayes classifier had an accuracy of about 84% which made it the most accurate of all employed methods. Nonetheless, the authors lamented that despite the achieved accuracy, in future, they would have to use a white box classifier like decision tree because Naïve Bayes, a black box classifier, cannot be used by non-data mining experts to make prediction.

Ko and Fang-Yie (2018) applied machine learning techniques to discover which significant attributes a successful learner often demonstrates in an undergraduate computer course. Five machine

learning algorithms were used in the study to compare and predict students' final performance. These are: decision tree, Naive Bayes, support vector machines, multilayer perceptron and logistic regression. The study indicated that Naive Bayes was the most appropriate for predicting students' performance.

Limsathitwong, et al. (2018) built prediction models using classifier algorithms such as Random Forest and decision tree to achieve an improvement and to recall students before they drop out. Their results showed that the Decision tree classifier obtained more precision in performance prediction than Random Forest.

Xu, et al. (2017) developed a novel machine learning method for predicting student performance in degree programs that is able to address key challenges such as students' background. Selected courses and progress of students were also incorporated into the prediction system. First, a bi-layered structure comprising of multiple base predictors and a cascade of ensemble predictors was developed for making predictions based on students' evolving performance states. Secondly, a data-driven approach based on latent factor models and probabilistic matrix factorization was proposed to discover course relevance, which is important for constructing efficient base predictors. The proposed method achieved superior performance to the benchmark approaches.

Buniamin, et al. (2015) used grades of students in Mathematics, Signal and Digital System, English as well as their CGPA to construct a model using Neuro-Fuzzy classification algorithm. The algorithm was chosen after a wide study on several algorithms that perform classification. The model constructed is hoped to help teachers make predictions of their students' CGPA.

Al-Shehri, et al. (2017) estimated students' performance in final examination using two supervised learning algorithms using dataset from the University of Minho in Portugal with 395 data samples. Support Vector Machine and K-Nearest Neighbor were the algorithms used. Study results showed that Support Vector Machine achieved better results than the K-Nearest Neighbor.

Sweeney, et al. (2015) developed a system for predicting students' course grades for the next enrolment term in a traditional university setting. Historical grade data was used to predict the grades for each student in the courses they will enroll during the next term. The factorization machine (FM), and the general-purpose matrix factorization (MF) algorithms were used for this study. Their results showed that FM outperformed MF making it the choice for the next term prediction system.

Rechkoski, et al. (2018) evaluated grade prediction for future courses using the model-based collaborative filtering methods: Probabilistic Matrix Factorization (PMF) and Bayesian Probabilistic Matrix Factorization (BPMF) using Markov Chain Monte Carlo. The prediction model was evaluated in a simulated scenario of an enrolment cycle in a winter and summer semester, based on a real data-set of enrolments. The results of evaluation were presented including the distribution of deviations of the predicted grades from the actual grades, deviation across different study programs, courses and grades. Results of this study showed that PMF-based grade prediction performed better than BPMF-based grade prediction.

Abakouy (2017) focused on predicting introductory programming performance of first year bachelor students in Computer Application course by predictive data mining model using classification-based algorithms. Data used include demographics, grade in introductory programming at college and grade in introductory programming test which contains 60 questions. Five supervised learning classification algorithms used in this study included: Multilayer Perception (MLP), Naive Bayes, SMO, J48 and REPTress using WEKA. MLP performed better with 93% accuracy.

The review of literature revealed that though a number of research projects have been carried out on student academic performance not much has been done with regards to postgraduate student academic performance. Again, within the Nigerian context and particularly in northern Nigerian universities, little research has been done on postgraduate student academic performance. There is therefore the need to narrow this gap.

MATERIALS AND METHODS

The Classification Algorithms used

The study used J48 Decision Tree, PART rule based and Naïve Bayes Algorithms for predicting the academic performance of the postgraduate students. Decision Tree is a supervised learning technique that has been widely used to analyze data for classification using the divide-and-conquer rule (Yadav et al., 2012; Witten et al., 2011). Decision Tree as a classification technique has improved accuracy when non-cognitive features are added to demographic and cognitive features as attributes to predict students' performance (Sultana et al., 2017).

PART is a rule-based algorithm for solving classification problems. It is used in the field of education to solve problems bothering on the field. It is often used on dataset that are not robust. Rule-based classifiers generate classification models using a combination of "if ... then ..." rules. The algorithms are computationally inexpensive, are capable of incorporating categorical and continuous variables and the developed models are usually easy to interpret (Lehr, et al., 2011). PART is an indirect method for rule generation.

Naïve Bayes is a classification method that is based on Bayes' theorem which is used to predict class labels. This classifier is based on probability theorem and is named after Thomas Bayes who is the founder of the theorem (Goyal and Mehta, 2012; Obuandike, et al., 2015)

WEKA Data Mining Tool

The experimental procedures were conducted using the Waikato Environment for Knowledge Acquisition (WEKA), a Java-based application. WEKA contains machine learning and statistical algorithms used by experts in the data mining field. WEKA works with arff and CVS file formats. CVS format can easily be converted to arff format when imported into WEKA.

Dataset and Attributes

A dataset of sixty-six (66) Postgraduate Diploma graduated students for three academic sessions (2015/2016, 2016/2017 and 2017/2018) from the Department of Computer Science, Nigerian Defence Academy, Nigeria was used for this study. A mix of demographic, cognitive and non-cognitive features of the students was selected. The demographic features used were age and gender. Class and Degree were the cognitive data used while the non-cognitive features used were Marital Status, Senior Secondary Certificate examination (SSCE) and Employment as illustrated in Table 1. Class was used as the target variable for the students' performance with "Excellent", "Good", and "Poor" features.

Table 1: The Various Attributes Showing a Combination of Demographic, Cognitive and Non-Cognitive Features.

Category of Data	Attributes	Features
Demographic	Age	Age of students
	Gender	Male Female
Cognitive	Class	Excellent, Good, Poor
	Degree	HND, BSc, PGD
Non-cognitive	Marital Status	Single Married
	Employment	Employed Unemployed
	SSCE	1 for One sitting 2 for two sittings

Architecture of the Proposed Methodology

The architecture of the proposed methodology showing the process used in developing the classifier is represented diagrammatically in Figure 1. Demographic, Cognitive and Non cognitive data were the features used in the transformation, cleaning, and integration of the students' data. This was followed by the feature selection; the learning and testing phase of the classifier; and lastly, the evaluation of the classifier.

Preprocessing Step

This is the stage where data preprocessing takes place (Figure 1). Preprocessing is an important step in the data mining process which often removes noisy data from the data set. Preprocessing was done in order to have quality data and few but effective attributes for classification (Hamalainen and Vinni, 2011). Activities involved in preprocessing include but are not limited to data cleaning, data integration and data transformation.

The raw data collected from the Department of Computer Science were cleaned at this stage. The cleaning involved removing incomplete data i.e. the data of students on deferment and voluntary withdrawal were discarded. In the data integration stage, different features were combined to form the seven attributes used as earlier depicted in Table 1.

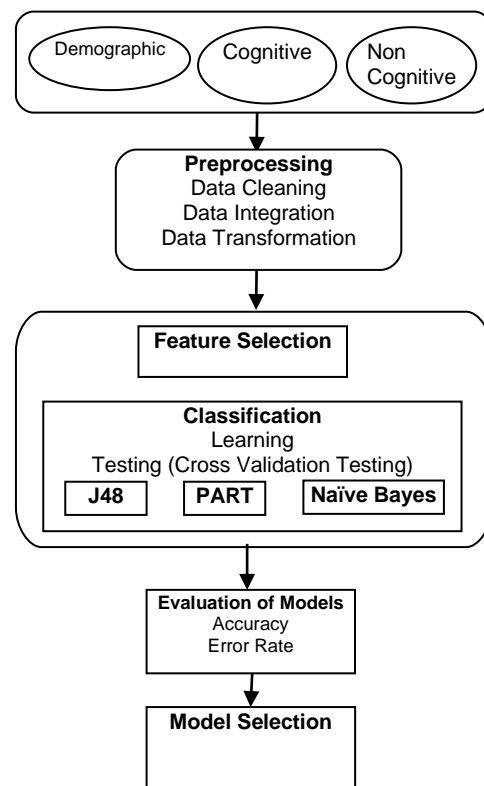


Figure 1: Architecture of Proposed Methodology.

Feature Selection Step

The students' attributes identified may not all be relevant to the prediction of the performance of the students. Therefore, it is important to remove irrelevant attributes before the classification process is carried out. Feature selection removes some attributes from the training data but leaves and uses relevant and useful attributes (Ahmad, 2015). According to Han and Kamber (2006), feature selection makes the produced pattern easier to understand and improves the speed of processing.

Classification Step

Classification of the patterns of the students features selected was done using decision tree and rule-based algorithms in WEKA. Decision Tree technique was implemented using J48 algorithm in WEKA application while rules algorithm was implemented using PART algorithm also in WEKA application. The model applied these two classification techniques on the students' demographic, cognitive and non-cognitive features which were used to build the classifiers in order to predict the students' performance. The target variable considered by the model for the students' performance is Class with "Excellent", "Good", and "Poor" features.

The classifiers were tested with the WEKA 10 fold cross validation. Maximizing the accuracy of a classifier is of utmost importance when working with classification techniques. To ensure accuracy of the classifiers, two experiments were performed. The first experiment was done with no feature selection while the second experiment was conducted with feature selection.

Evaluation Step

The evaluation was carried according to five performance metrics. These performance metrics are accuracy, error rate, precision, recall and F1 score. The classifier with the highest accuracy was selected.

Performance Metrics and Confusion Matrix

Evaluating the performance of classifiers is an important task in the overall data mining process (Oprea, 2014). The study adopted five

performance metrics namely: accuracy, error rate, precision, recall and F1 score. These measures can be derived from the confusion matrix. The confusion matrix is a useful tool for analyzing how well classifiers can recognize tuples of different classes. Confusion matrix was used to determine the performance of the classifiers. Confusion matrix is based on the following parameters: true positives, true negatives, false positives, and false negatives (Joshi, 2017).

Where,

True positives (TP): refers to the correctly identified tuples by the classifier as positive.

True negatives (TN): refers to the correctly identified tuples by the classifier as negative.

False positives (FP): refers to the incorrectly identified tuples by the classifier as positive.

False negatives (FN): refers to the incorrectly identified tuples by the classifier as negative.

Accuracy

Accuracy is simply a proportion of correctly predicted observation to the total observations or dataset.

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

Precision

Precision is used to access the percentage of tuples labelled as True Positives. Precision is the proportion of correctly predicted positive observations to the total predicted positive observations.

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

Recall

Recall is also known as Sensitivity or True Positive Rate (TPR). Recall is the measure of recognition rate (that is, the proportion of positive tuples that are correctly identified).

$$\text{Recall or Sensitivity} = \frac{TP}{TP + FN} \quad (3)$$

F1 score

F1 Score is the weighted average of Precision and Recall.

$$\text{F1 Score} = \frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (4)$$

Error

Error is simply the complement of accuracy.

$$\text{Error} = 1 - \text{accuracy} \quad (5)$$

RESULTS AND DISCUSSION

Result of First Experiment (Without Feature Selection)

The first experiment was conducted without any feature or attribute selection. The learning and testing were achieved by applying the three classification techniques on the students' performance features. Consequently, confusion matrices were produced as results of the experiments and are presented in the Table 2, Table 3, and Table 4.

Table 2: Naïve Bayes Classification Confusion Matrix without Feature Selection.

	Predicted			
	Excellent	Good	Poor	
Actual	Excellent	28	3	7
	Good	4	0	7
	Poor	5	0	12

Table 3: J48 Decision Tree Classification Confusion Matrix without Feature Selection.

	Predicted			
	Excellent	Good	Poor	
Actual	Excellent	29	1	8
	Good	4	4	3
	Poor	9	1	7

Table 4: PART Rule Based Classification Confusion Matrix without Feature Selection.

	Predicted			
	Excellent	Good	Poor	
Actual	Excellent	30	4	4
	Good	4	3	4
	Poor	7	3	7

Result of Second Experiment (With Feature Selection)

In the second experiment where feature selection was done, the model fed the attributes to WEKA where some attributes were selected and others dropped. WrapperSubsetEval attribute evaluator and Best first search method were used for the analysis being one of the effective and commonly used feature selection methods.

After applying the feature selection algorithms, three attributes (SSCE, MStatus and Degree) were selected for naïve bayes classification. For J48 Decision Tree classification, three attributes (MStatus, Employment and Degree) were also selected while five attributes (Gender, Age, MStatus, Employment and Degree) were selected for PART Rule Based classification.

The results of the experiment with feature selection are presented in the Tables 5, 6, and 7 and in Figures 2, 3, and 4.

Table 5: Naïve Bayes classification confusion matrix with feature selection

	Predicted			
	Excellent	Good	Poor	
Actual	Excellent	35	0	3
	Good	3	0	8
	Poor	9	0	8

Table 6: J48 Decision Tree Classification Confusion Matrix with Feature Selection.

		Predicted		
		Excellent	Good	Poor
Actual	Excellent	34	1	3
	Good	3	4	4
	Poor	8	1	8

Table 7: PART Rule Based Classification Confusion Matrix with Feature Selection.

		Predicted		
		Excellent	Good	Poor
Actual	Excellent	28	2	8
	Good	2	4	5
	Poor	7	1	9

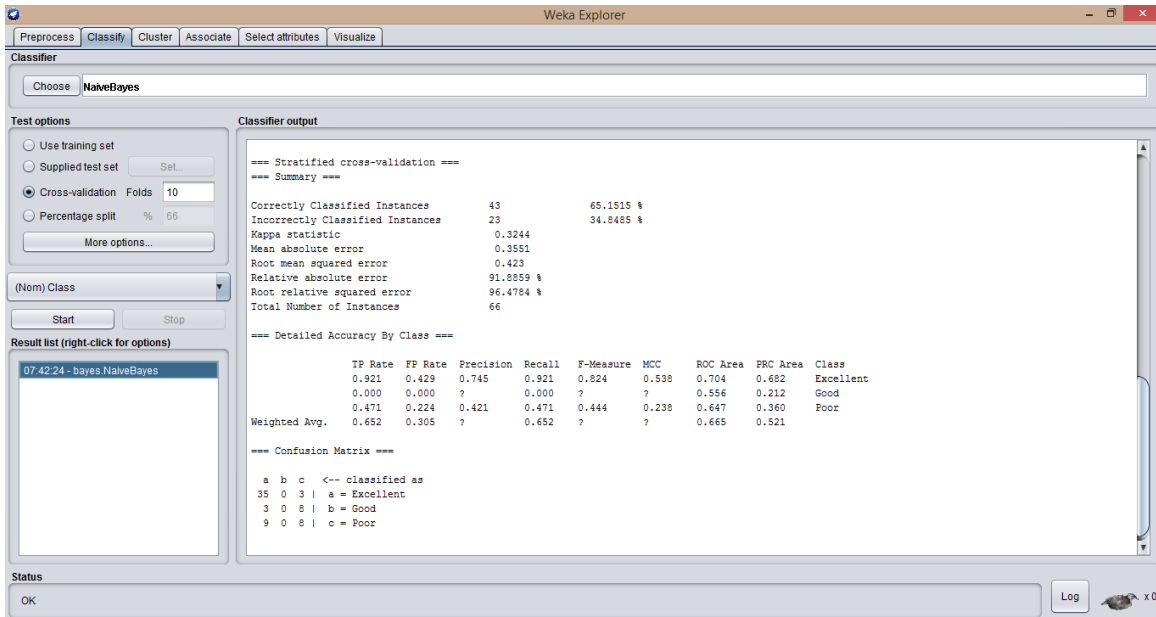


Figure 2: Screenshot of Naïve Bayes Classification with Feature Selection.

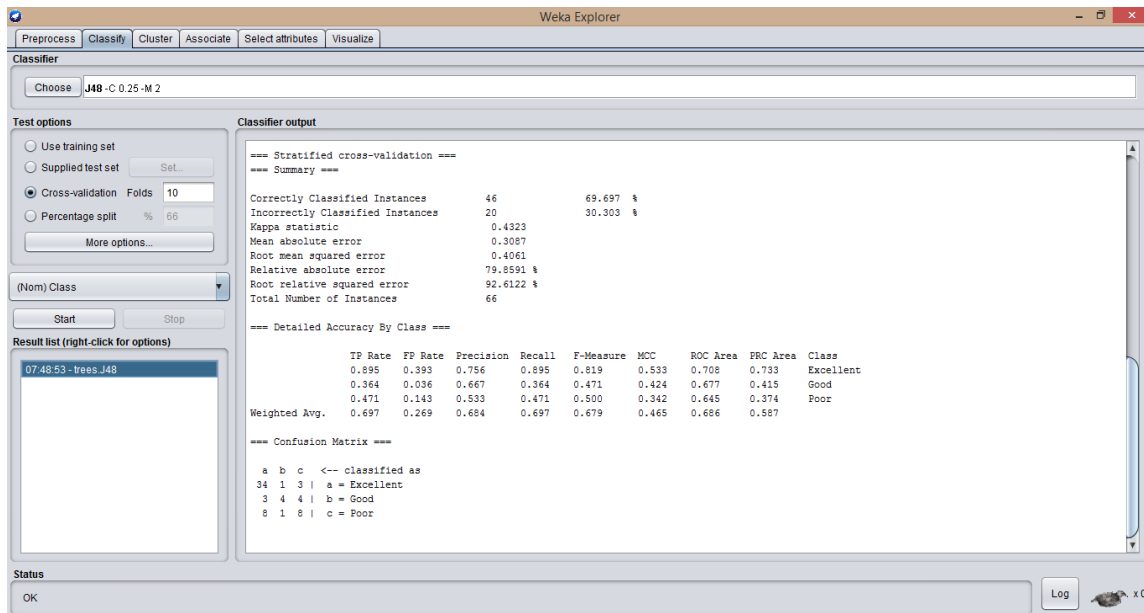


Figure 3: Screenshot of J48 Decision Tree Classification with Feature Selection.

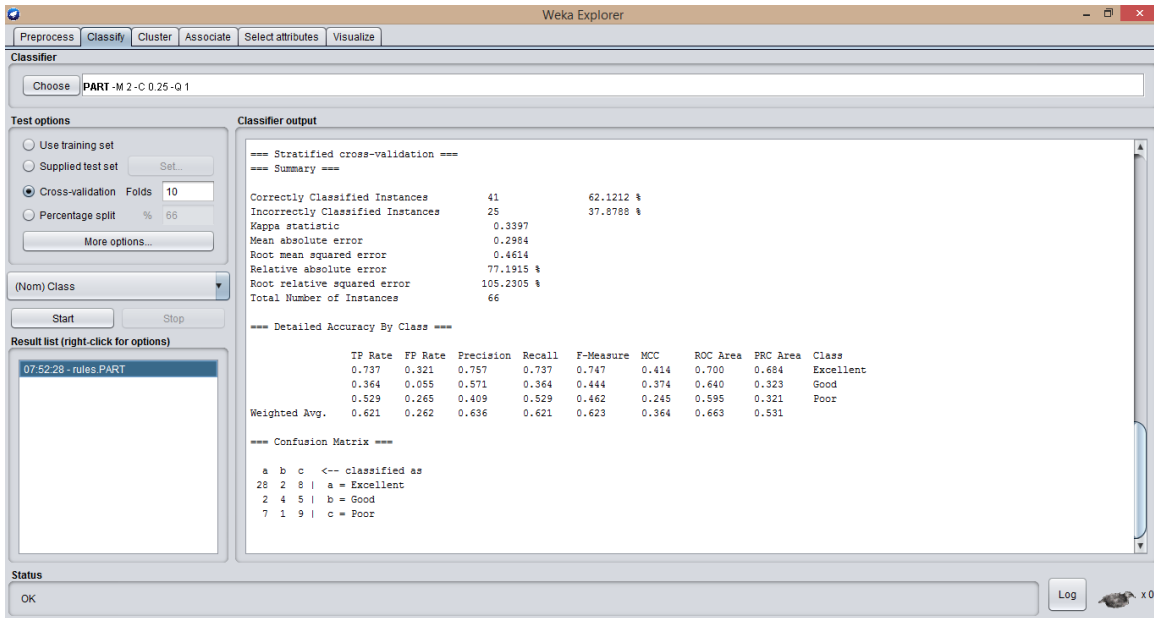


Figure 4: Screenshot of PART Rule-Based Classification with Feature Selection.

Evaluation of Classifiers

The evaluation of the classifiers from the results of the experiments are presented in Tables 8 and 9. The evaluation is guided by the performance metrics discussed in the methodology. Table 8 shows the evaluation result of the classifiers without feature selection.

Table 8: Evaluation Result of the Classifiers without Feature Selection.

Classifier	Method of training	Accuracy (%)	Precision	Recall	Error Rate (%)	F1 Score
Naïve Bayes	10-fold Cross Validation Training	60.6061	0.757	0.737	39.3939	0.747
J48	10-fold Cross Validation Training	60.6061	0.690	0.763	39.3939	0.725
PART	10-fold Cross Validation Training	60.6061	0.732	0.789	39.3939	0.759

The three classifiers, in the experiment without feature selection, have the same accuracy value of 60.6061%. Additionally, Naïve Bayes classifier

has the highest precision value of 0.757 while PART classifier has the highest recall value of 0.789. Table 9 shows the evaluation result of the classifiers with feature selection.

Table 9: Results of Classifier with Feature Selection.

Classifier	Method Of training	Accuracy (%)	Precision	Recall	Error Rate (%)	F1 Score
Naïve Bayes	10-fold Cross Validation Training	65.1515	0.745	0.921	34.8485	0.824
J48	10-fold Cross Validation Training	69.6970	0.756	0.895	30.3030	0.819
PART	10-fold Cross Validation Training	62.1212	0.757	0.737	37.8788	0.747

In the experiment after feature selection, J48 classifier gave the highest accuracy of 69.6970% out of the three classifiers. Furthermore, PART classifier gave the highest precision value of 0.757 while Naïve Bayes classifier gave the highest recall value of 0.921. Thus the best classifier with highest accuracy and best predictive power in the experiment with feature selection is J48 classifier with accuracy of 69.6970%.

DISCUSSION

The performance metrics of accuracy, error rate, precision, recall and F1 score guided the evaluation of the classifiers. The best accuracy is 1.0, whereas the worst is 0.0. A high precision indicates that algorithm returns more relevant results than irrelevant and high recall means that most of the results returned by the algorithms are relevant.

Table 9 results show that the 10-fold cross validation for J48 classification is better than that of PART and Naïve Bayes. J48 classification had accuracy, precision and recall values of 69.70%, 0.756 and 0.895, respectively as against 62.12%, 0.757 and 0.737, respectively, for PART classification and 65.15%, 0.745 and 0.921 respectively for Naïve Bayes, respectively.

Tables 5, 6, and 7 indicated the confusion matrix of the classifiers with feature selection indicating the academic performance of students that are excellent, good and poor. The J48 classifier was able to predict accurately 34 out of the 38 excellent students, 4 out of the 11 good students, and 8 out of the 17 poor students. This gives an accuracy of 89.5% for excellent, 36.4% for Good and 47.1% for the poor student prediction.

Naïve Bayes classifier predicted accurately 35 out of the 38 excellent students, none out of the 11 good students and 8 out of the 17 poor students giving an accuracy of 92.1% for excellent, 0% for good and 47.1% for the poor students prediction. The PART classifier on the other hand predict accurately 28 out of the 38 excellent students, 4 out of the 11 good students and 9 out of the 17 poor students. This gives an accuracy of 73.7 for excellent, 36.4% for Good and 52.9% for the poor student prediction. Therefore, J48 classifier performed better than Naïve Bayes and PART classifiers.

CONCLUSION

The research investigated the prediction of students' academic performance using data obtained from indigenous postgraduate students at the NDA PG School, Kaduna, Nigeria. Three most widely used data mining classification techniques, namely, J48 decision tree, Naïve Bayes and PART were used for the training and evaluation with the aid of WEKA tool. J48

classifier emerged the most accurate, with an accuracy of 69.70% and error rate of 30.30%.

REFERENCES

1. Abakouy, R., E.M. En-Naimi, and A. El Haddadi. 2017. "Classification and Prediction Based Data Mining algorithms to Predict Email Marketing Campaigns". *Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems*. ICCWCS'17. 1-5.
2. Ahmad, A.S. 2015. "Educational Data Mining to Predict Student Exam Grades in Vocational Institutes". Dissertation submitted in partial fulfilment of the requirements for the degree of MSc Information Technology Management, British University in Dubai.
3. Al-Shehri, H., A. Al-Qarni, L. Al-Saati, A. Batoaq, H. Badukhen, S. Alrashed, J. Alhiyafi and S.O Olatunji 2017. "Student Performance Prediction using Support Vector Machine and K-Nearest Neighbor". *IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, Windsor, ON. 1-4
4. Asif, R., A. Merceron, S.A. Ali, and N.G. Haider. 2017. "Analyzing Undergraduate Students' Performance using Educational Data Mining". *Computers & Education*. 113: 177-194.
5. Badr, G., A. Algobail, H. Almutairi, and M. Almutery. 2016. "Predicting Students' Performance in University Courses: A Case Study and Tool in KSU Mathematics Department". *Procedia Computer Science*. 82: 80-89.
6. Buniyamin, N., U. Mat, and P.M. Arshad. 2015. "Educational Data Mining for Prediction and Classification of Engineering Students Achievement". *IEEE 7th International Congress on Engineering Education (ICEED)*. IEEE. 49-53.
7. Goyal, A. and R. Mehta 2012. "Performance Comparison of Naive Bayes and J48 Classification Algorithms". *International Journal of Applied Engineering Research*. 7 (11).
8. Hamalainen, W. and M. Vinni, M. 2011. *Classifiers for Educational Data Mining*. In Romero, C., S. Ventura, M. Pechenizkiy, and R. Baker. 2011. *Handbook of Educational Data Mining*. CRC Press: New York, NY. 57-74.
9. Han, J. and M. Kamber. 2006. *Data Mining Concepts and Techniques*. Morgan Kaufmann: San Francisco, CA.
10. Hasibur, R., and I. Rabiul. 2017. "Predict Students' Academic Performance and Evaluate the Impact

- of Different Attributes on the Performance Using Data Mining Techniques". 2nd International Conference on Electrical & Electronic Engineering (ICEEE), Bangladesh.1–4.
11. Joshi, R. 2017. "Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures". Available online @blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/. Retrieved February 5, 2018.
 12. Ko, C. and L. Fang-Yie. 2018. "Analyzing Attributes of Successful Learners by using Machine Learning in an Undergraduate Computer Course". IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). 801–806.
 13. Lehr, T., J. Yuan, D. Zeumer, S. Jayadev and M.D. Ritchie (2011). "Rule based classifier for the analysis of gene-gene and gene-environment interactions in genetic association studies". *BioData Mining*. 4(4).
 14. Limsathitwong, K., T. Kanda and Y. Tanasin. 2018. "Dropout Prediction System to Reduce Discontinue Study Rate of Information Technology Students". 5th International Conference on Business and Industrial Research (ICBIR). 110–114.
 15. Obuandike G.N., A. Isah and J. Alhasan. 2015. "Analytical Study of Some Selected Classification Algorithms in WEKA Using Real Crime Data". *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*. 4(12): 44-48.
 16. Oprea, C. 2014. "Performance Evaluation of the Data Mining Classification Methods". Annals of the Constantin Brancusi University of Targu Jiu, Economy Series, Special Issue-Information Society and Sustainable development. *Academica Brancusi Publisher*. 249-253.
 17. Rechkoski, L., V.V. Ajanovski, and M. Mihova. 2018. "Evaluation of Grade Prediction using Model-based Collaborative Filtering Methods". IEEE Global Engineering Education Conference, EDUCON. 1096–1103.
 18. Sultana, S., S. Khan, S., and M.A. Abbas. 2017. "Predicting Performance of Electrical Engineering Students using Cognitive and Non Cognitive Features for Identification of Potential Dropouts". *International Journal of Electrical Engineering Education*. 1-14.
 19. Sweeney, M., J. Lester and H. Rangwala 2015. "Next-term student grade prediction". IEEE International Conference on Big Data (Big Data). 970–975.
 20. Witten, I.H., E. Frank and M.A. Hall. 2011. *Data Mining Practical Machine Learning Tools and Techniques*. Morgan Kaufmann: Burlington, VT.
 21. Xu, J., K.H. Moon and M. van der Schaar. 2017. "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs". *IEEE Journal of Selected Topics in Signal Processing*. 11(5): 742–753.
 22. Yadav, S.K. and S. Pal. 2012. "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification". *World of Computer Science and Information Technology Journal*. 51-56.
 23. Yassein, N. A., R.G.M. Helali, and S.B. Mohomad, 2017. "Predicting Student Academic Performance in KSA using Data Mining Techniques". *Journal of Information Technology & Software Engineering*. 7(5): 1-5.
 24. Mueen, A., B. Zafar and U. Manzoor. 2016. "Modeling and Predicting Students' Academic Performance Using Data Mining Techniques". *International Journal of Modern Education and Computer Science*. 11: 36-42.

ABOUT THE AUTHORS

Dr. Abraham E. Evwiekpaefe, is a Senior Lecturer at the Department of Computer Science, Nigerian Defence Academy, Kaduna State, Nigeria. He holds both a Master of Science (M.Sc.) degree and a Doctor of Philosophy (Ph.D.) degree in Computer Science from the University of Benin, Benin City, Nigeria. His research interests are in IT adoption, Software Engineering, Internet of things, and Data Science.

Muhammad M. Isa, is a Lecturer at the Department of Computer Science, Nigerian Defence Academy, Kaduna State, Nigeria. He holds both a Bachelor of Science (B.Sc.) degree in Computer Science and a Professional Masters degree in ICT from the Bayero University, Kano, Nigeria. His research interests are in the areas of machine learning, artificial intelligence and Computer networks.

Fiyinfoluwa Ajakaiye, is a Lecturer at the Department of Computer Science, Nigerian Defence Academy, Kaduna State, Nigeria. She holds a Master of Science (M.Sc.) degree in Computer Animation from the University of Westminster, UK. Her Ph.D. is in view. Her

research interests are in Machine Learning, Artificial Intelligence, Edge Computing, Internet of things, Computer Graphics and Software Engineering.

SUGGESTED CITATION

Evwiekpaefe, A.E., M.M. Isa, and F. Ajakaiye. 2020. "Analyzing Factors Affecting Academic Performance of Postgraduate Students using Data Mining Techniques". *Pacific Journal of Science and Technology*. 21(2):203-213.

 [Pacific Journal of Science and Technology](http://www.akamaiuniversity.us/PJST.htm)