International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

An Enhanced Multiple Classfier System for Predicting Students' Academic Outcome

Ahmed Zainab Jibril[1], Umar Adamu Bello[2], A.F. D Kana[3]

[1,2]Yusuf Maitama Sule University, Kano, Nigeria,
[3]Ahmadu bello University Zaria, Nigeria

Abstract
Data-driven decision support approaches have been increasingly employed in recent years to unveil purposeful task-oriented patterns from accumulated students' academic records. Educational data mining and Knowledge Discovery in Database (KDD) provide a viable solution to decipher implicit knowledge through predictive modelling. Most approaches used the induction of the individual classification models also known as the single classifiers and few efforts focuses on the ensemble methods; which are unstable, overfits and susceptible to skewed data or a poorly performing classifier. To overcome this work examines the existing multiple classification algorithms, Pandey and Taruna approach and proposes an enhanced MCS that uses bagging strategy to reduce variance and perhaps improves accuracy of the resulting model to improve the accuracy. The approach is based on the definition of a model that integrates several classification techniques to predict academic performance of undergraduate students. Data were collected, cleaned, preprocessed and integrated using MS (excel and access) and Weka software. Experiments were carried out using WEKA and EMCS improved using python library. The result shows that the proposed (EMCS) has an increase of 1.03% and 0.63% of accuracy over the Bagging MCS and Pandy& Taruna respectively

KEYWORDS: Educational Data mining EDM, Multiple classifier/ensemble, Bagging

1. INTRODUCTION

Data mining involves revealing intricate patterns within extensive datasets using machine learning, statistics, and databases. Recently, a surge in employing Data Mining and Knowledge Discovery in Databases (KDD) for educational purposes has been witnessed, particularly in extracting task-oriented patterns from student academic records [Al-Barrak & Al-Razgan, 2016; Dutt et al., 2017; Dwivedi & Singh, 2016; Livieris et al., 2016; Santana et al., 2017; Abdulsalami A. O 2016].

Educational Data Mining (EDM) has emerged as a novel avenue, aiming to predict academic performance and unveil insights within historical academic data [Dutt et al., 2017; Dwivedi & Singh, 2016; Livieris et al., 2016; Ognjanovic et al., 2016; Pandey & Taruna, 2016; Santana et al., 2017; Satyanarayana & Nuckowski, 2016; Sivakumar et al., 2016]. Data mining's influence spans domains such as healthcare (Kavakiotis et al., 2017), business (Massaro, 2018), and education (Adekitan, 2018; Alyahyan et al., 2020).

EDM's advanced methods play a crucial role in enhancing learning environments, evaluating both educational settings and machine learning techniques [Albreiki et al., 2021]. Previous research confirms that factors like emotions, family dynamics, and study schedules significantly impact students' academic performance in cognitive activities [S. Ahmad et al., 2022].

This study focuses on enhancing a multiplier classifier algorithm to predict student performance, specifically within the departments of Computer Science, Biology, Crop Science, Statistics, and Civil Engineering at Kano University of Science and Technology Wudil (KUST). By improving performance prediction, the study aims to contribute to the understanding of student outcomes and potentially provide insights for tailored educational interventions.

2. RELATED WORK

Educational Data Mining (EDM) is an emerging field that builds models to uncover unique data from educational settings, enhancing understanding of students and their influencing factors (Santana et al., 2017). EDM develops methods and algorithms to detect patterns, utilizing computational and psychological strategies. New interactive learning tools and systems offer insights into student trends and behaviors, fostering discoveries and hypotheses about learning (Santana et al., 2017).

Knowledge Discovery in Database (KDD) is a multidisplinary approach for changing data into knowledge (Fayyad *et al.*, 1996; Shapiro, *et al.*, 1996; Tan *et al.*, 2008). The major steps of KDD procedure for knowledge acquisition are data pre-processing, data mining and evaluation of the discovered knowledge.

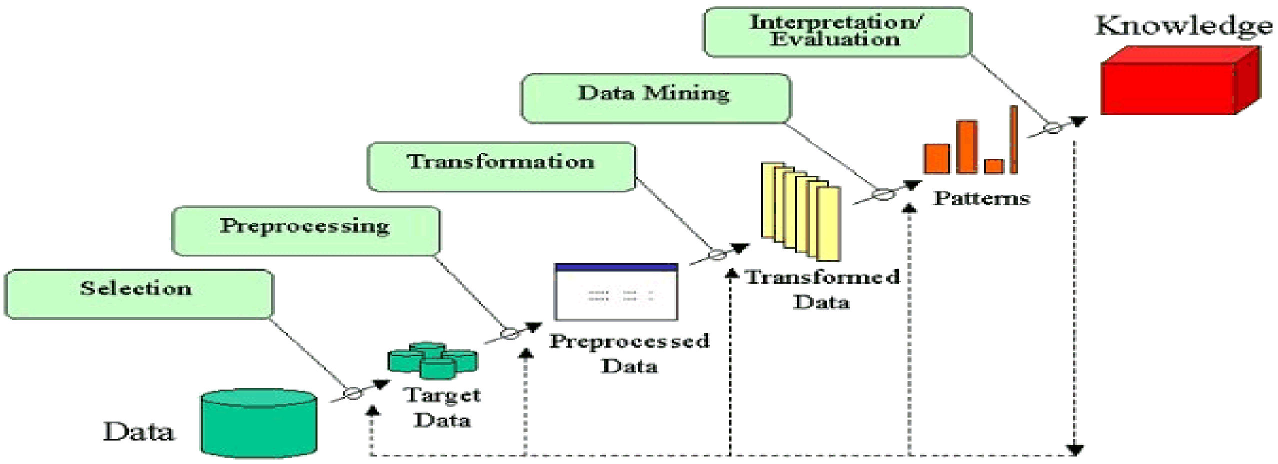

Figure 1: The Knowledge Discovery Process (Fayyad et al., 1996)

Various related works based on KDD classification models have been utilized to predict students' academic outcomes. Kovacic (2010) explores socio-demographic variables (age, gender, ethnicity, education, work status, and disability) and the study environment (course program and course block) that may influence students' persistence or dropout in higher education. Yadav et al. (2012) induce two models classifying successful and unsuccessful students, using data like attendance and assessment scores to build a decision tree model. They describe several methods and approaches within the EDM context. Kabakchieva (2013) centers their research on constructing data mining models to predict student performance, incorporating personal, pre-university, and university performance attributes. They employ OneR rule learner, decision trees, neural networks, and k-nearest neighbor (k-NN) classifiers. The results indicate that university entry scores and year 1 failures play a role in determining success. Kaur et al. (2015) focus on the application of predictive models using classification algorithms to identify and discover slow learners. They collect their dataset from a high school and apply preprocessing techniques using the WEKA workbench. Anuradha & Velmurugan (2015) aim to predict students' performance in the final semester of university. They employ techniques such as decision tree C4.5 (J48), Bayesian classifiers, k Nearest method, and rule learners. Xing et al. (2015) propose a student performance prediction model that is both practical and comprehensible, utilizing genetic algorithms for interpretability. They suggest the possibility of extending the model using ensemble classifiers and measuring its accuracy performance. Abdulsalami A.O (2016) develops a model for early detection of students at risk of attrition. They employ various classification algorithms, including Multi-Layer Perceptron, Naïve Bayes, J48 Decision Tree, Sequential minimal optimization, K-Nearest Neighbor, and a modified K-Nearest Neighbor. Results indicate the strong performance of J48 decision tree with an average accuracy of 97.9%, and modified nearest neighbor algorithm with an average accuracy of 97.3%. The use of ensemble classifiers to further improve performance is suggested for future work. Pandy and Taruna (2016) present a diverse multiple classifier (heterogeneous) based framework that fuses classifiers such as AODE, IBK, and J48 using voting methodology. They propose a single compound model for use and suggest its potential for building decision support systems. Al-Barrak and Al-Razgan (2016) utilize classification techniques, particularly decision trees, to predict students' final GPA based on grades in previous courses. They suggest potential extensions using other techniques such as MCS, neural networks, and clustering. Dutt et al. (2017) provide an explicit schema of learning methods for students, considering attributes like time spent on tasks, group discussions, student behavior, classroom decoration, and learners' inspiration. Clustering offers insights into pertinent attributes differentiating clusters.

However, educational data is typically multi-level hierarchical and non-independent, necessitating researchers' careful consideration when selecting clustering algorithms that align with research questions to yield valid and reliable results. Not all clustering algorithms are applicable in the EDM context.

2.1 MULTIPLE CLASSIFIER SYSTEM (MCS)

MCS consists of a set of classifiers whose decision outputs are combined using an aggregation function to achieve the final decision outcome. The concept of MCS has been shown to improve performance, often not attainable using single classifiers (Dietterich, 2000; Polikar, 2006). MCS can have an adverse effect by either skewed data or an inadequately performing classifier affecting the overall performance. MCS is sensitive to classifiers.

The classifiers that made up MCS are referred to as the base-classifiers. Theoretical report has shown that to achieve improvement in performance a MCS must consist of diverse base-classifiers having unrelated classification errors (Kuncheva & Whitaker, 2003). However recent studies have shown the adoption of MCS in EDM

MCS also known as ensemble methods, committee, classifier fusion, combination, aggregation, Integration is said to improve predictive accuracy. The most powerful MCS techniques in data mining include: Boosting, Bagging, and Random Trees

2.2 DIVERSIFICATION OF CLASSIFIERS

There two basic methods of integrating classifiers the heterogeneous and homogeneous methods; Heterogeneous ensemble of classifier refers to combine the predictions of multiple base models. While heterogeneous refers to inclusion of different decision techniques such as classification or regression to decide process. However, this study focuses on the Homogeneous ensemble methods use the same base learner on different distributions of the training set, e.g. bagging and boosting

Two classifiers are diverse, if they make different errors on a new object, Assume a set of three classifiers {h1, h2, h3} and a new object x.

- If all are identical, then when h1(x) is wrong, h2(x) and h3(x)will be also wrong (making the same decision)
- If the classifier errors are uncorrelated, then when h1(x) is wrong, h2(x) and h3(x) may be correct → a majority vote will correctly classify x!

Ensemble methods are multiple classifier combination method used to create powerful / explicit classification tree models. This is accomplished by integrating weak classification models to develop stronger versions. However, Bagging, Boosting, stacking and Random Forest methods are powerful alternatives that could provide relevant data classifications.

Boosting

Boosting is a powerful ensemble algorithm which is inclined to reduce both bias & variance, and also promotes the conversion of weak learners (i.e., classifiers having correlation classification) to strong learners (i.e classifiers with uncorrelated error). Boosting produces powerful classification models by training models to focus on misclassified records from earlier models; in addition, classifiers here are also combined using weighted majority vote. The algorithm computes the weighted sum of votes for all the class and allots the best classification to the record. Boosting is not capable of parallelization; moreover, if dataset is very large and has a significant number of weak learning algorithms, then boosting may not be *the most suitable method.*

Random Trees

Random forests or random decision forests are ensemble learning method for classification, regression and other tasks, deals with generating a multitude of decision trees at training time and outputting the mode/mean classes prediction of the individual decision trees. Voting and averaging is used for prediction in the case of classification and regression respectively.

Random Trees abstracts the bagging concept of random feature selection to build models of decision tree with controlled variance. Random Tree process employ training datasets to generate several decision trees and using the mode from each class to build a powerful classifier. The advantages of using this method include; it is parallelizable, it prevents overfitting, and building models is usually faster compared to Bagging. Its drawback is that the improved speed is caused by the number of selected features in each iteration; thus, result not *comprehensive*

Bagging

Bagging (aka bootstrap aggregating) is an easy but strong ensemble algorithm that promotes the increased strength & accuracy of classification models. Bagging procedure is demonstrated by creating multiple training datasets using random samples with replacement, inducing the algorithm to each dataset, then selects the majority vote among the built models to determine data classifications. Bagging is peculiarly known to reduces variance, prevent overfitting, and it can easily be applied on large datasets. In addition, both bagging and boosting uses the majority vote approach. Therefore, bagging homogeneous method of MCS is used in this work.

The literature shows this has mainly been achieved using single classification modelling techniques. In most cases comprehensible classifier learning techniques, which return a set of propositional rules (IF–THEN rules), or tree-based classifiers are most preferred (Al-Barrak & Al-Razgan, 2016; Dejaeger et al., 2012; Rathee A, 2013; Xing et al., 2015). This is due to the explanatory nature of such model, which allows professional to check if the output of the model is in line with the accepted factors that define educational quality.

To enhance the accuracy and reliability of classification models in EDM, researchers have advocated for Multiple Classifier Systems (MCS) (Pandey & Taruna, 2014, 2016; Satyanarayana & Nuckowski, 2016), which integrates several classifiers to form a powerful complex model. Although, the emergence of the concept of integrating classifiers to form MCS has been regarded as a viable solution to overcome the limitation of single classifier (Kuncheva, 2004; Polikar, 2006)

2.3 LITERATURE GAPS AND CONTRIBUTION OF THIS WORK

Most of the works found in EDM literature focused on single classifiers. Quite few efforts in EDM have been expanded on powerful combination methods of classifiers i.e. the MCS. Most of research efforts are tailored towards the prediction accuracy using either the heterogeneous or homogeneous existing combination methods of classifiers as such there is need for performance prediction model that is highly accurate, reliable and a decision support system for educational experts to evaluate. This work tends to develop a model that its accuracy supersedes the existing approaches and can support decision making policies The research is to extend research efforts titled: Multiple classifier and combination of models to predict academic performance (Pandey & Taruna, 2016).

3. METHODOLOGY

EDM and KDD process steps have been used to provide a viable solution through predictive modelling. It begins with the initial data collection from the university ARKEPS, applying the preprocessing techniques; first is to import the libraries that will be needed,which involves data cleaning to remove noise, handling missing values, Encoding categorical data also known as determining data quality, spliting the dta into training and test sets, feature scaling which simply transform all variables into the same scale known as normalization/standardization. The Figure below shows the pictorial presentation of the architecture of the intended system. The entire architecture comprises the data mining process and the MCS.

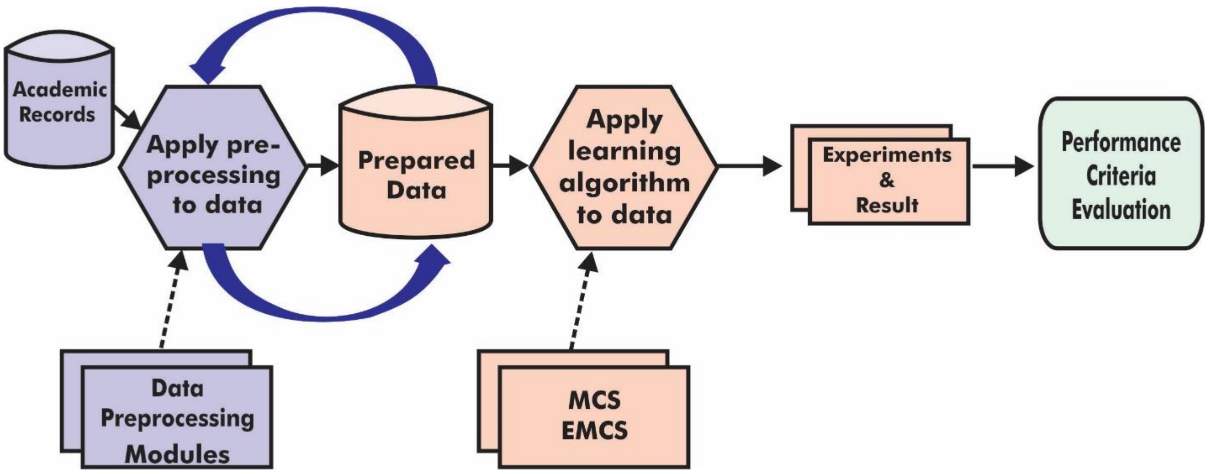

Figure 2: system Architecture

DATA COLLECTION

Dataset: The dataset used was from KUST Wudil, Kano. The dataset consists of academic information of undergraduate students. The dataset consists of approximately complete (2000) instances across the selected programs chosen: Biology, civil engineering, computer science crop science, electrical engineering and statistics. There attributes selected which includes (GPA's, CGPA's, Gender and a notable class attribute of GRADUATE or NOT_GRADUATE) dataset of student performance with a total instance of (1525) was used by Pandy and Taruna (2016)

Data Preprocessing Techniques Used:

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. The techniques employed in this work includes the following:

Data Cleaning:

The reason for data cleaning is to remove unwanted noise i.e. inaccurate or incomplete record from data, certain tasks like harmonization and standardization of data were carried out. Which involves: missing data and noisy data: All of this procedure was carried out using Weka software.

Data Transformation

Data transformation and feature selection approach; where employed which makes the patterns easier to recognize.

Strategies involved includes the following:

1. **Normalization**, here attributes are scaled to fall within the smallest range, such as −1.0 to 1.0, or 0.0 to 1.0. In this work the attribute class was defined to be either of **GRADUATE**/ **NOT_GRADUATE** and this was determined using *mean and standard deviation* v' =(v-Mean)/StDev. The final cumulative grade point (cgpa) of the student was used to determine if the student will be graduate or not graduate.
2. **Attribute construction** (or *feature construction*), here new attributes are created and added from the set of given attributes to facilitate the mining process. In this work, newer attributes were constructed from existing ones; an example is deriving the value cumulative grade point from attributes like credit unit and grade point of a particular course the student offered.
3. **Aggregation**, here summary operations are applied to the dataset. E.g, total credits units registered (TCUR) or earned (TCUE) in a particular session/ semester. This step is commonly used in building a dataset for data analysis at several abstraction levels.
4. **Discretization**, here numeric valued attributes (for example, *age*) were replaced using interval labels (e.g., 0–10, 11–20, etc. in this dissertation we mapped values for student entry level like 100, 200 and 300 level to UTME, Direct entry and admissions on transfer respectively.

MODEL VALIDATION

The dataset used in this work is partitioned into two (train and test set) using the k-fold cross validation (where k = 10) this technique was applied to all machine learning algorithms used for classification purpose; Ninety percent (90%) of the data was chosen to be the training set, and 10% to be the testing set in each iteration.

PROPOSED ALGORITHM FOR ENHANCED MULTIPLE CLASSIFIER SYSTEM (EMCS)

The primary contribution of this research is to enhance the MCS techniques with the view to improve accuracy that provides the best decision-making policies in an academic environment. EMCS approach is used to enhance the performance of classifiers. The modified algorithm for multiple classifier performance criteria for evaluating the classifiers uses: classification accuracy, specificity, sensitivity/recall, precision, absolute relative root square error, Time.

Proposed Model for EMCS:

Inputs: Training set S; Ensemble Selection classifier E; Integer N (number of bootstrap samples)

Basic procedure:

1. *for i = 1 to N {*
2. *S_b ← bootstrap sample from S (sample with replacement)*
3. *S_{oob} ← out of bag sample*
4. *Generate classifiers in E—›S_b*
5. *E_i —›do ensemble selection based on base classifiers' performance on S_{oob}*
6. *}*
7. *Predicting class label for new instance:*

$$E\ (e_i, s_j) = \sqrt{\textstyle\sum_{i=1}^{n} \propto\ E_i\,(e_i, s_j)^2}$$

Example E_i is classified to the class S_j in accordance to the number of votes obtained from particular classifiers E_i.

In this case precision can be obtain by computing:

Precision $\beth^{\mu} = \frac{TP}{TP+FP} = \frac{\sum_{j=1}^{|C|} TP_j}{\sum_{j=1}^{|C|}(TP_j + FP_j}$ ………………..... Equation 1

Recall $\rho^{\mu} = \frac{TP}{P+FP} = \frac{\sum_{j=1}^{|C|} TP_j}{\sum_{j=1}^{|C|}(TP_j + FN_j}$ ……………………… Equation 2

PERFORMANCE EVALUATION CRITERIA

The method for analyzing classifier performance is the confusion matrix, the easiest and most common form of the confusion matrix is a two-class matrix as shown in the fig below given two classes (Positive and Negative classes). True Positives are positive instances that were correctly classified, True Negatives instances that are correctly classified for the negative class while incorrectly classified positive instances are called false positive and incorrectly classified negative instances are known as false negatives. TP (True Positives) TN (True Negatives) FP (False Positives) FN (False Negatives).

		Predicted class	
		yes	no
Actual class	yes	true positive	false negative
	no	false positive	true negative

Figure 3: confusion matrix

4. RESULTS AND DISCUSSION

Table below shows a summary of accuracy (%) achieved by the MCS in building the models for all the programs. The average of each classifier's accuracy metric recorded during the experiments is given. EMCS performs best based on this set of data and has an average of 94.0%.

Classifiers	**Bio**	**Civil**	**CompSci**	**Crop**	**Electrical**	**Stat**	**Average**
RANDOM TREES	**88.8%**	**92.8%**	**95.8%**	**78.2%**	**95.3%**	**79.7%**	**88.43%**
BOOSTING	**90.0%**	**94.0%**	**95.0%**	**79.5%**	**96.4%**	**80.1%**	**89.17%**
BAGGING	**91.4%**	**96.0%**	**96.1%**	**87.5%**	**96.9%**	**89.9%**	**92.97%**
PANDY&TARUNA	**93.0%**	**95.6%**	**96.6%**	**89.0%**	**96.0%**	**90.0%**	**93.37%**
EMCS	**94.2%**	**96.5%**	**96.7%**	**88.5%**	**98.0%**	**95.1%**	**94.0%**

Table 1: ***Performance accuracy comparison for all six datasets***

From the result of all the comparism it appears that the proposed algorithm performs better and reflect to be more reliable and accurate perhaps behaves almost same as that of Pandey &Taruna, (2016) and proves to support decision making policies by assisting educational personnel to provide quick intervention to students at risk which turns to improve graduation rate.

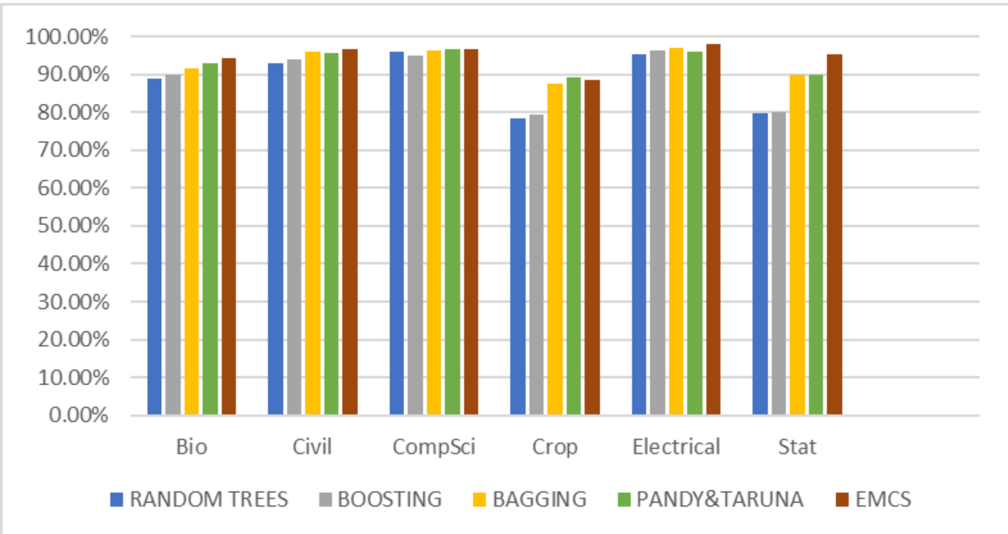

Figure 3: performance accuracy comparison of models

The figure above shows that the EMCS and all MCS performs similarly and that EMCS has a relatively higher accuracy when compared on the data set employed in this work

5. RECOMMENDATION AND FUTURE WORK

This endeavor marks the initial phase of Educational Data Mining research at Kano State University of Technology. There are several avenues for further enhancement, including:

- **Incorporating Extracurricular Data:** Expanding the scope by incorporating data from students' extracurricular activities. Research has shown that very few contributions have mathematically modeled emotional attributes, making this an area ripe for exploration.
- **Integration with University Systems:** Enhancing the utility of the developed model by integrating it with university databases, portals, and e-learning platforms. This integration could provide a comprehensive overview of student performance and facilitate informed decision-making.
- **Early Predictive Application:** Extending the application of this model from the outset, such as during the student admission process, by utilizing data like O'level/JAMB scores and Putme. This predictive approach can alleviate admission-related pressures, leading to an improved graduation rate.

This work serves as a foundational step in the realm of Educational Data Mining, with promising avenues for future advancements and broader impact.

REFERENCES

1. Abdulsalami A. O (2016) Detection of Students at Risk of Attrition Using Data Mining Approach, MSc thesis Department of Mathematics ABU Zaria
2. Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting student’s final GPA using decision trees: a case study. *International Journal of Information and Education Technology, 6*(7), 528.
3. Albreiki, B.; Zaki, N.;Alashwal, H. A Systematic Literature Review of Student’ Performance Prediction Using Machine Learning Techniques. Educ. Sci. **2021**, 11, 552. https://doi.org/10.3390/educsci11090552
4. Adekitan, A. I. (2018). *“Data mining approach to predicting the performance of first year student in a university using the admission requirements,” no. Aina 2002*.
5. Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education, 17*(1). doi:10.1186/s41239-020-0177-7
6. Ahmad, S., El-Affendi, M. A., Anwar, M. S., & Iqbal, R. (2022). Potential future directions in optimization of students’ performance prediction system. *Computational Intelligence and Neuroscience*, *2022*, 1-26. doi:10.1155/2022/6864955
7. Anuradha, C., & Velmurugan, T. (2015). A comparative analysis on the evaluation of classification algorithms in the prediction of students performance. *Indian Journal of Science and Technology, 8*(15).
8. Dejaeger, K., Goethals, F., Giangreco, A., Mola, L., & Baesens, B. (2012). Gaining insight into student satisfaction using comprehensible data mining techniques. *European Journal of Operational Research*, *218*(2), 548-562.
9. Dietterich, T. (2000). Ensemble Methods in Machine Learning *Multiple Classifier Systems* (Vol. 1857, pp. 1-15): Springer Berlin / Heidelberg.

10. Dutt, A., Ismail, M. A., & Herawan, T. (2017). A Systematic Review on Educational Data Mining. *IEEE Access*.
11. Dwivedi, T., & Singh, D. (2016). Analyzing Educational Data through EDM Process: A Survey. *International Journal of Computer Applications, 136*(5), 13-15.
12. Fayyad, U., Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine, 17*, 37-54.
13. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and Structural Biotechnology Journal*, *15*, 104–116.
14. Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. *Cybernetics and Information Technologies*, *13*(1), 61-72.
15. Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector. *Procedia Computer Science, 57*, 500-508.
16. Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades. *Artificial Intelligence Review*, *37*(4), 331-344.
17. Kovacic, Z. (2010). Early prediction of student success: Mining students' enrolment data.
18. Kuncheva, L. I. (2004). *Combining Pattern Classifiers Methods and Algorithms*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc.
19. Livieris, I. E., Mikropoulos, T. A., & Pintelas, P. (2016). A decision support system for predicting students' performance. *Themes in Science and Technology Education*, *9*(1), 43-57.
20. Massaro, A., Maritati, V., & Galiano, A. (2018). Data mining model performance of sales predictive algorithms based on Rapidminer workflows. *International Journal of Computer Science & Information Technology*, *10*(3), 39–56.
21. Ognjanovic, I., Gasevic, D., & Dawson, S. (2016). Using institutional data to predict student course selections in higher education. *The Internet and Higher Education*, *29*, 49-62. doi: http://dx.doi.org/10.1016/j.iheduc.2015.12.002
22. Pal, A. K., & Pal, S. (2013). Analysis and mining of educational data for predicting the performance of students. *International Journal of Electronics Communication and Computer Engineering*, *4*(5), 1560-1565.
23. Pandey, M., & Taruna, S. (2014). A Comparative Study of Ensemble Methods for Students' Performance Modeling. *International Journal of Computer Applications*, *103*(8).
24. Pandey, M., & Taruna, S. (2016). Towards the integration of multiple classifier pertaining to the Student's performance prediction. *Perspectives in Science*, *8*, 364-366
25. Polikar, R. (2006). Ensemble based systems in decision making. *Circuits and Systems Magazine, IEEE*, *6*(3), 21-45.
26. Rathee A, M. R. (2013). Survey on Decision Tree Classification algorithm for the Evaluation of Student Performance. *International Journal of computers & Technology*, *4*(2), 244–247.
27. Santana, M. A., Costa, E. B., Fonseca, B., Rego, J., & de Araújo, F. F. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*.
28. Satyanarayana, A., & Nuckowski, M. (2016). Data Mining using Ensemble Classifiers for Improved Prediction of Student Academic Performance.

29. Sivakumar, S., Venkataraman, S., & Selvaraj, R. (2016). Predictive modeling of student dropout indicators in educational data mining using improved decision tree. *Indian Journal of Science and Technology, 9*(4).
30. Yadav, S. K., Bharadwaj, B., & Pal, S. (2012). Data mining applications: A comparative study for predicting student's performance. *arXiv preprint arXiv:1202.4815*.
31. Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47*, 168-181.