Science in Information Technology Letters
Vol. 4., No. 1, May 2023, pp. 1-11
ISSN 2722-4139
http://pubs2.ascee.org/index.php/sitech

Placement model for students into appropriate academic class using machine learning

Anadi Stella Uju [a,1], Khalid Haruna [b,2,*], Ibrahim A. Lawal [b,3], Raliya Abubakar [b,4]

[a] Department of Computer Science, Bayero University, Kano, Nigeria
[b] Department of Information Technology, Bayero University, Kano, Nigeria
[1] ujustella84@gmail.com; [2] kharuna.cs@buk.edu.ng; [3] ialawan.it@buk.edu.ng; [4] raliyatsafe@gmail.com
* Corresponding Author

ARTICLE INFO

Article history
Received March, 26, 2023
Revised April, 11, 2023
Accepted April, 29, 2023

Keywords
Placement model
Appropriate academic class
Machine learning

ABSTRACT

Choosing the right academic major for junior secondary students into senior secondary school will assist both students and their teachers toward achieving the academic goal. Traditionally, students seeking admission into senior classes (Gambia, Sierra leone, Ghana, Liberia and Nigeria) must have passed stipulated examinations like Basic Education Certificate Examination (BECE) and/or West Africa Junior Certificate Examination, which are done at the end of year three (at a sitting). They must pass the exam(s) satisfactorily with no emphasis on any of Science, Art or Commercial related subjects. Some schools use “Mock exam” or “Placement exam” as the basis for their placement of students but all are done at a sitting (end of year three). Though this method is to an extent valid but associated with some challenges (bias) as it does not carry along the student’s academic history in making decision for placement into appropriate class. In this research,, we proposed a model that predicts appropriate academic class of Science, Art or Commercial for Junior students based on their progressive academic performances (history) of their predecessors on related subjects using ten supervised machine learning techniques. Two evaluation techniques were applied. The highest results of this research showed accuracy of 93% with Random Forest, 98% precision with random forest, 99% recall with Decision tree and 94% f1 score with Random forest and KNN. The correlation coefficient of the proposed model recorded 0.3 improvement to the existing methods. This research will benefit all stakeholders in education and students in particular because their academic performances over time stands a better chance for appropriate placement.

1. Introduction

The need for proper selection of academic major for students’ going into senior secondary school is of paramount importance towards achieving academic goals. It is a useful strategy to mitigate failure, promote the achievement of better results and to better manage resources in higher education. It is also established that some students discontinue their educational pursuit (dropout) due to class misplacement from their basic education level [1]. Researches have shown that students drop out of school for various reasons such as parental poverty [2], sickness [3], subjects studied at university as well as the secondary school grades [4], nature of the learning environment [5], teacher-student relationship [6], insecurity

[7], economic recession [8], accumulated failure [9] and so on. Predictive models have been proven to be an important approach toward achieving remarkable improvements in both productivity, and proficiency in almost all human endeavours [10].

We proposed placement model for students into appropriate academic class using machine learning. This model will use the progressive academic performance of students over eight terms on selected subjects to make placement. The generated datasets can be used for further research in educational field and the model can be used to predict appropriate academic class for students using their progressive academic performance.

The remaining sections of this paper are as follows: a detailed literature review is presented in section 2. Section 3 presents our proposed methodology. Results and discussion is presented in section 4. Conclusion and recommendations for future work is presented in section 5.

2. Literature Review

Most of the researches on student academic performance are geared towards the management of at-risk students at the higher education level [11] and, to a greater extent, in the distance education modality [12]. Very little information has been found on the application in basic or secondary education, where just simple analyses of the information based on statistical methods have been carried out [13]. There are some differences and/or advantages in applying data mining over just using statistical models [14]. Therefore, any predictive model designated to analyzing the students' academic performance must take cognizance of the complex and nonlinear relationships that exist among variables .

An adaptive recommendation system was proposed by Mohamed [4] for predicting a suitable education path(s) for students in college preparatory year. The adaptability was achieved by automatically applying different data mining techniques for extracting relevant features and building a tailor-made model for each education path. It recommends a suitable engineering department among seven engineering departments for each student based on his academic performance.

Vijay et al. [15] mainly focuses on the prediction of a student's university result by making use of different attributes using KNN. These attributes might be of quantitative and qualitative type. The quantitative attributes used are Internal Assessments, Attendance percentage, Number of On-Duties taken and Overall Assignments completed. The qualitative attributes include Subject feedback, Faculty feedback, and whether the student is a Day Scholar/Hosteller. KNN was used against the historical data of students for a more accurate prediction of results. This algorithm makes use of the Euclidean distance formula which is used to find the nearest record and predicts better results which help students maximize their academic output.

[16] used KNN to predict the category to which a student fails based on the skills, ability and behavior of the student after 4 weeks of 12 weeks course, similar to the work of [17] who predicted at-risk students in online course on touch-typing skill after the very first lesson.

[18] applied ensemble ML techniques for predicting dropout students in the Thai University based on their first-year academic performance. 124 students were used and the performance of the model was appreciative. [1] used the information available at the end of the first-degree year of students' academic career to segment students based on evidence of failure or high performance.

[9] applied ML in school dropout prediction based on Columbia University students' academic performance. They focus their interest in determining the multiple factors that most affect the performance of students (dropout and failure) at the different educational levels (basic, middle and higher education) through the use of the large amount of information that current computer equipment allows to store in databases. Ten classifiers were used on both demographic and academic features. Random forest emerged the highest accuracy of 85%.

[19]–[23] predicted students' academic performance by tracking the academic performance of students for a short term. This yielded a better result when compared with existing methods (one sitting exam).

From the reviewed literature, it was observed that most of the data used for analysis were obtained at a very early stage regardless of the improvement that may occur later. Predicting student performance is usually not a one-time task [20].

You cannot accurately predict a student's academic potential only through his early grades or even through his static demographic traits [24]. Later prediction yields better accuracy and is preferred when intervention is not needed [25]. This requires continuous tracking of the student's performance over a longer period of time, to reveal his/her true potentials.

3. Method

This section shows the methodology used in carrying out the research. Three different datasets were used. These datasets were generated from the same students in their various academic stages. Junior Mock Record (JMR): Records generated from one sitting placement exams at 8th term by 1500 students on thirteen subjects in junior secondary. Junior Progressive Record (JPR): Records generated from 1500 students on six subjects for eight terms progressively in junior secondary.

Senior Progressive Record (SPR): Records generated from 1500 students on six subjects for eight terms progressively in senior secondary. Fig. 1 shows the activity chart of the work flow.

3.1. Dataset

Dataset is collection of data, a dataset can be made up of various types of data, including text data, image data, and numerical data [26]. The information is typically initially marked or described to enable the algorithm to grasp the desired result [27]. Creating these datasets can pose challenges in terms of both complexity and cost due to the extensive time required for data labeling [28].

Datasets are typically divided into training, validation, and test sets, serving distinct roles in the machine learning process. Data sets are usually divided into training, validation, and test sets, which fulfill different roles in the machine learning process.

The size, format, source, and destination of the data set are important considerations, with metadata providing important contextual information. Understanding the domain and purpose of a data set is critical to ensuring responsible and informed use, while being mindful of potential issues such as privacy, bias, and data quality.

The characteristics of the dataset used in this research is presented in Table 1.

Table 1. Dataset

Dataset	Description
Subject Category	Subjects area categorized into three classes of science, art and commercial in the following subjects combinations; Science : Biology, Chemistry, Physics, Basic science and Basic technology. Art : Economics, Literature, Government, Social studies and Cultural and creative art. Commercial : Account, Commerce, Enterpreneur, Civics and Business studies.
Terms	Examination scores of eight terms (three session) for both senior and junior students
Students	1521 student recors were collected. Twenty-one (1.4%) were dropped during preprocessing. 1050 instances (70%) were for training and 450 instances (30%) were used for testing with 154 art, 149 commercial and 147 science labels

3.2. The Existing Method

The existing manual method as used by the sampled schools, subject students to a placement examination (at a sitting) at the end of year three (3). The average score of all the subjects per student were taken and students were categorized based on their overall average performance.

The classes are categorized based on the following scale :

- 65 marks and above __ science class.
- 64 – 55 marks __ commercial class.
- 54 and below __ art class

Here, science class is prioritized (higher precedence) over other classes. That is, sciences are for outstanding students and art class for weak students.

3.3. Algorithm of The Proposed Model

- Stage 1 : Data is collected from students from seven schools (Qi)
- Stage 2: - Do preprocessing by sorting student scores
  - Cleaning by dropping records less than eight terms (t < 8)
  - Apply the equation $\hat{Y} = (\sum_{i=1}^{t} Q_i)\ {}^1\!/_8$. Where $\hat{Y}$ = preprocessed record, t = number of terms, i = scores per term and Q = students. Then drop $\hat{Y}$ < 50.
- Stage3: Do classification by applying ten classifiers on Junior Mock Record (JMR), Junior Progressive Record (JPR) then Senior Progressive Record (SPR).
- Stage4: Perform evaluation using 10 fold cross validation and 70/30 splitting and measure using confusion matrix.

Activity chart of the proposed model show as Fig. 1.

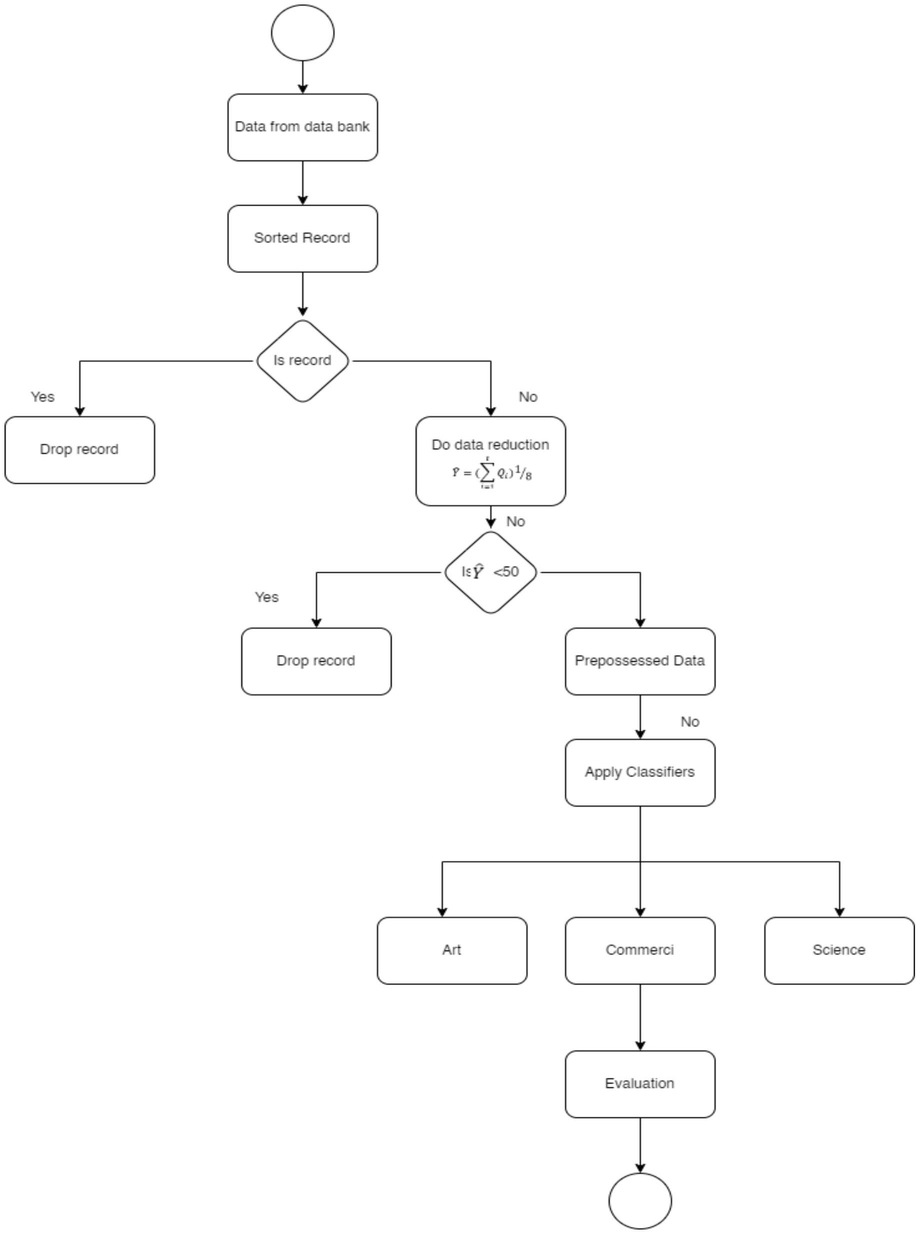

Fig. 1. Activity chart of the proposed model

3.4. Proposed Placement Model for Students Into Appropriate Academic Class Using Machine Learning

Our proposed model is divided into four stages as represented in Fig. 3. It is made up of data resource and collection, preprocessing, classification and evaluation stages.

- Stage 1: Data Resource and Collection.

  The source of the data used in this research was obtained from five private schools (Federal Airport Authority of Nigeria (FAAN); Prince of Wales Secondary School, Sierra leone; Nusrat Senior Secondary School, Gambia; Akosombo International School, Ghana and Our Lady of Grace International school, Liberia) from the five West Africa Examination Council, (WAEC) member countries . Only students who enrolled into JSS1 class in 2015 were considered with emphasis on the following subjects: Biology, Physics, Chemistry, Accounting, Commerce, Entrepreneur, Literature, Government, Economics, Basic science, Basic technology, Social studies, Civic education, Cultural and creative art and Business studies. The average performances of senior students for eight terms (SPR) their corresponding progressive average in JSS1 – JSS3 (JPR) and their mock performances (JMR) were used for the analysis. Every given student in each class has to offer three combinational subjects either for science, art or commercial as stipulated by the WAEC policy , (2013) and shown in Fig 2. Eight academic terms must be completed in JSS (Junior secondary school) and eight terms in SSS (Senior secondary school). 1521 students' records were collected for the analysis.

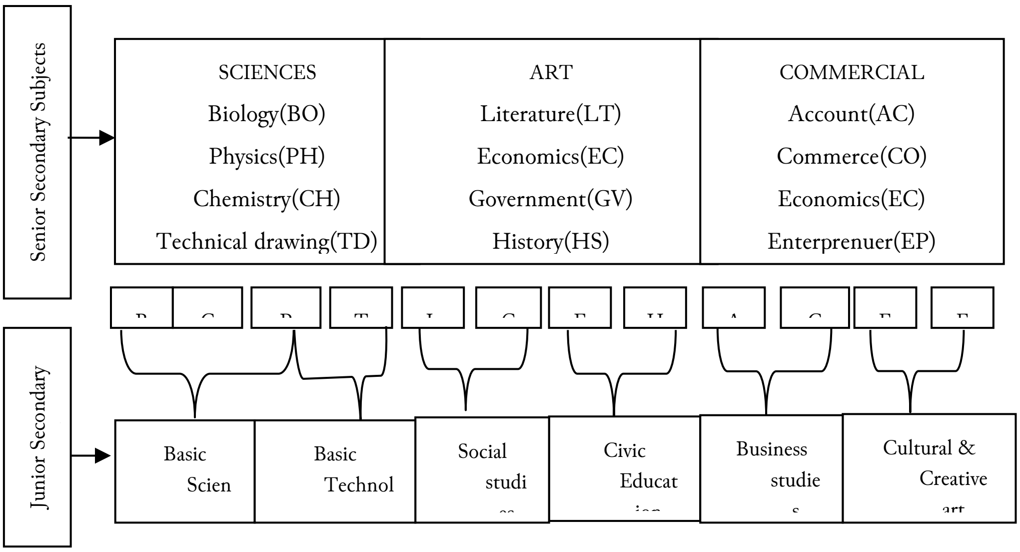

Fig. 2. Summary of relationship between subjects

- Stage 2: Data Preprocessing.

Data preprocessing is the preparation of data into a usable form. To achieve this, the collected datasets has to undergo certain stages which include :

- Data Sorting: during data collection, records of students were randomly collected irrespective of class (science, art or commercial). These records will then be sorted based on the actual class which they are presently in and labeled.
- Data Cleaning: Is the process of detecting and correcting corrupt or inaccurate record from a record set. Nine students with incomplete records (less than 8 terms) were extracted from the datasets as they represented only 0.6% of the sample size.
- Data Reduction: In ML, data reduction is done to bring records into a single representation. In this model, we are to apply the equation :

$$yi = \left(\sum_{i=1}^{t} Q_i\right) {}^{1}/_{8} \dots \dots \dots \dots . eqn(4) \quad (1)$$

Where i represent the scores obtained by student Q and t number of terms. Then, Q_i^t denote student i's academic scores at term (t). Our independent variable is the scores obtained by the students (predictors) and the dependent variable is the targeted output (science, art or commercial). The $\widehat{yi}$ for all the students will be filtered by setting a threshold of $\widehat{yi} \geq 50$. Twelve students (0.8%) whose scores fall below this range were extracted. This threshold was set inconformity to the set criteria by West Africa Examination Council law, (1995). 1500 samples were used as input to the next stage.

- Stage 3: Classification and prediction

Classification is a technique used to assign a class of unseen records as correctly as possible by using a collection of records called a training dataset, where each tuple in the training set comprises of a set of attributes, and then one of the attributes is called a class. Because is supervised learning, the class label (art:0, commercial:1 and science:2) was assigned to all the datasets for python compatibility. The preprocessed data were divided into 70:30 for training and testing respectively. Then, SMOTE (Synthetic Minority Oversampling Technique) was applied on the training datasets to balance the minority class. Ten classifiers (LR, SVM, DT, RF, NB, LoR, Adaboost, RNN, GPC and KNN) were passed on the datasets for classification.

- Stage 4: Performance Evaluation

In this phase, the result from the existing model and that of the proposed model was evaluated to ascertain the differences in their performance. Confusion matrix was used to measure accuracy, precision, recall and F1 score of all datasets. The general formula for these measures as stated by [11].

$$Accuracy = \frac{Number\ of\ Correct\ Predections}{Total\ Number\ of\ Data\ Poin} \quad (2)$$

$$Precision = \frac{True\ Positive}{(True\ Postive + False\ Positive)} \quad (3)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (4)$$

$$F1 - Score = 2^{*}\ Precission * Recall \quad (5)$$

3.5. Architecture of The Proposed Model

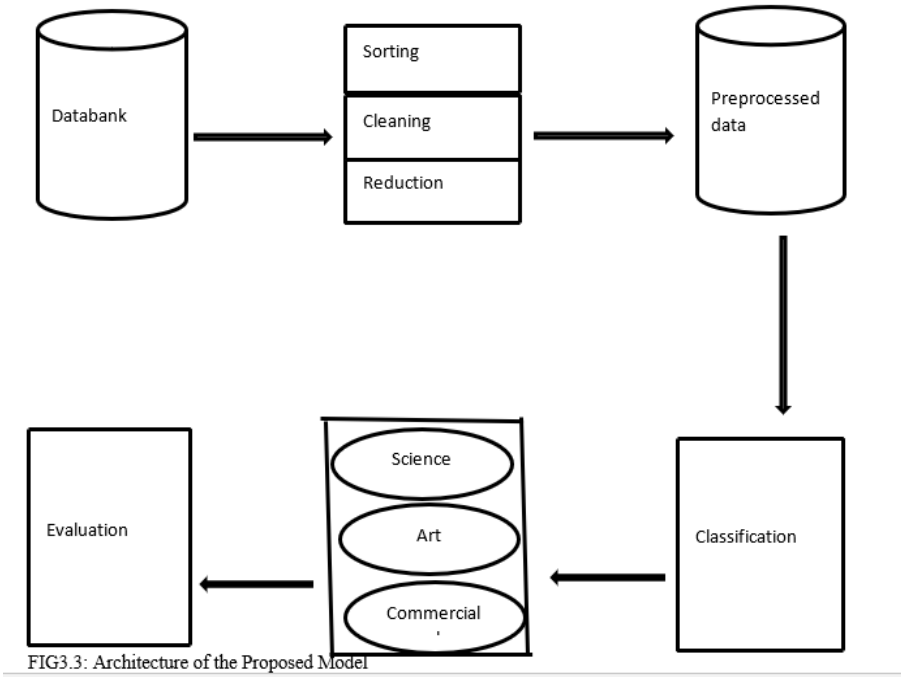

Fig. 3. Architecture of The Proposed Model

4. Results and Discussion

The results of the three experiments are explained and recorded in Table. 2 and graphically represented in Fig. 4.

Table 2. Results of Experiments

	Ensamble		Regression			Non Linear	Neural Network	Gausian		
Classifier	*RF*	*Adaboost*	*LiR*	*LoR*	*SVM*	*DT*	*RNN*	*NB*	*GPC*	*KNN*
Accuracy 1	84/95/90	68/71/70	44/74/73	44/65/60	50/65/60	80/91/84	68/72/70	48/75/72	82/93/90	81/94/90
Accuracy 2	72/93/80	71/83/80	58/76/74	46/58/54	68/70/69	73/92/88	50/57/49	64/80/49	71/90/88	71/85/92
Precission	82/98/90	74/72/71	42/83/81	66/42/39	46/66/62	80/84/82	57/50/49	43/83/80	93/82/80	95/82/80
Recall	81/92/88	71/69/67	44/67/65	66/32/43	55/69/65	81/99/93	72/68/65	49/69/65	93/82/80	94/81/80
F1 Score	81/94/89	72/70/68	42/74/72	66/36/40	50/67/65	80/90/87	63/57/55	45/75/71	93/82/80	94/81/80

In the first experiment, considering the state-of-the-art performance evaluation technique (cross validation-accuracy2), decision tree attained the highest accuracy of 73% and logistic regression lowest accuracy of 46%. Random forest attained the highest accuracy of 84%, logistic and linear regression the least performed (70/30 splitting-accuracy1).

In the second experiment (JPR), Random forest emerged with the highest accuracy of 93% and least with RNN 57% (accuracy2). Random forest scored 95% against logistic regression and Support vector machine with least score of 65% accuracy (with 70/30 splitting-accuracy1).

The third experiment (SPR) recorded 90% accuracy for RF, GPC and KNN as the highest accuracy with 70/30 splitting and 60% as lowest with logistic regression(accuracy1). KNN scored the highest accuracy of 92% with cross validation while NB and RNN as the least performed with 49% (accuracy2).

In all experiments, 70/30 splitting evaluation metric out performed cross validation performance metric in six (6) classifiers (RF, KNN, LoR, RNN, GPC & NB) but the result of cross validation is more efficient and accurate because it gives the model the opportunity to train on multiple train-test split. It was also observed that there is a general higher accuracy in the results using JPR (progressive record) against the result using JMR (one sitting exam record). Our model used subjects from junior secondary which have a direct link to subjects offered in senior secondary. Eight terms (longer time) was considered as students tend to progress with time. We also measured the correlation coefficient between the three datasets used for analysis and result shown in Table 3.

Table 3. Correlation coefficient of datasets

JMR/JPR	JPR/SPR
0.419412	0.736714

Correlation coefficient is a valuable numerical measure of association between two variables which is a value between -1 and 1 indicating the strength of the association of the observed data for the two variables [4].

From the table above, it showed that the correlation coefficient between JPR/SPR is stronger (more accurate) than that of JMR/JPR. This is because the progressive performances of these students in junior secondary and their progressive performances in senior secondary have closer aggregate, that is, those who did well in basic science and basic technology in the junior level (progressive scores) where found also doing well in the science class (considering their progressive scores in chemistry, physics and biology). The values between JMR and JPR were dispersedly distributed thereby resulting to a weaker correlation. The correlation coefficient of the proposed model is 0.317302 higher than that of the existing model.

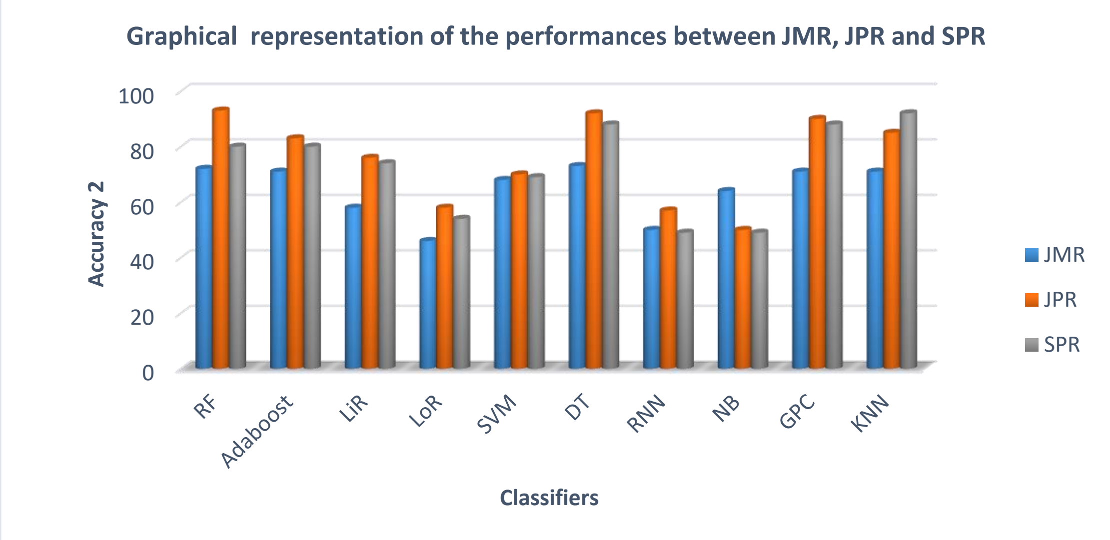

Fig. 4. Graphical representation of the performances between JMR, JPR and SPR

5. Conclusion

The results obtained after the experiments showed that our model can efficiently and accurately predict student's appropriate academic class by taking their progressive academic performance on selected subjects for longer time (JPR) compared to the result obtained using one sitting exam (MPR) on all the thirteen subjects. The generated datasets can be used for further research in educational field. The modality for handling students below the threshold score ((yi) ̂<50) is a direction for future research.

References

[1] V. L. Miguéis, A. Freitas, P. J. V. Garcia, and A. Silva, "Early segmentation of students according to their academic performance: A predictive modelling approach," *Decis. Support Syst.*, vol. 115, pp. 36–51, Nov. 2018, doi: 10.1016/j.dss.2018.09.001.

[2] O. M. O. Dada, "A Sociological Investigation of The Determinant Factors and the Effects of Child Street Hawking in Nigeria: Agege, Lagos State, Under Survey," *Int. J. Asian Soc. Sci.*, vol. 3, no. 1, pp. 114–137, 2013, Accessed: Oct. 17, 2023. [Online]. Available at: https://ideas.repec.org/a/asi/ijoass/v3y2013i1p114-137id2405.html.

[3] Å. Sagatun, S. Heyerdahl, T. Wentzel-Larsen, and L. Lien, "Mental health problems in the 10thgrade and non-completion of upper secondary school: the mediating role of grades in a population-based longitudinal study," *BMC Public Health*, vol. 14, no. 1, p. 16, Dec. 2014, doi: 10.1186/1471-2458-14-16.

[4] M. Ezz and A. Elshenawy, "Adaptive recommendation system using machine learning algorithms for predicting student's best academic program," *Educ. Inf. Technol.*, vol. 25, no. 4, pp. 2733–2746, Jul. 2020, doi: 10.1007/s10639-019-10049-7.

[5] R. Ekanem, E. Apebende, and E. Ekefre, "Learning Environment And Pupils Academic Performance: Implications For Counselling," *Couns.*, vol. 29, no. 1, pp. 31–42, 2011. [Online]. Available at: https://www.ajol.info/index.php/coun/article/view/73036.

[6] E. Maheshwari, C. Roy, M. Pandey, and S. S. Rautray, "Prediction of Factors Associated with the Dropout Rates of Primary to High School Students in India Using Data Mining Tools," in *Advances in Intelligent Systems and Computing*, vol. 1013, Springer, 2020, pp. 242–251, doi: 10.1007/978-981-32-9186-7_26.

[7] G. Ramaswami, T. Susnjak, A. Mathrani, and R. Umer, "Use of Predictive Analytics within Learning Analytics Dashboards: A Review of Case Studies," *Technol. Knowl. Learn.*, vol. 28, no. 3, pp. 959–980, Sep. 2023, doi: 10.1007/s10758-022-09613-x.

[8] U. K. Okoji, S. C. Taylor, and J. B. Lipoff, "Equity in skin typing: why it is time to replace the Fitzpatrick scale," *Br. J. Dermatol.*, vol. 185, no. 1, pp. 198–199, Jul. 2021, doi: 10.1111/bjd.19932.

[9] A. Viloria, J. García Guliany, W. Niebles Núñez, H. Hernández Palma, and L. Niebles Núñez, "Data Mining Applied in School Dropout Prediction," *J. Phys. Conf. Ser.*, vol. 1432, no. 1, p. 012092, Jan. 2020, doi: 10.1088/1742-6596/1432/1/012092.

[10] N. Patricia, U. Akanwa, and N. Norbert, "Challenges Encountered Using Cbt by 2015 Utme Candidates In Owerri Zone One, Nigeria: Test Validity Implications," *IOSR J. Res. Method Educ. Ver. I*, vol. 5, no. 5, pp. 28–35, 2015. [Online]. Available at: https://www.iosrjournals.org/iosr-jrme/papers/Vol-5%20Issue-5/Version-1/F05512835.pdf.

[11] A. O. Elfaki, K. M. Alhawiti, Y. M. AlMurtadha, O. A. Abdalla, and A. A. Elshiekh, "Supporting students' learning-pathway choices by providing rule-based recommendation system," *Int. J. Educ. Inf. Technol.*, vol. 9, no. January, pp. 81–94, 2015. [Online]. Available at: https://www.researchgate.net/publication/278032482_Supporting_students'_learning-pathway_choices_by_providing_rule-based_recommendation_system.

[12] S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah, and E. Alajrami, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology," *Int. J. Hybrid Inf. Technol.*, vol. 8, no. 2, pp. 221–228, 2015, doi: 10.14257/ijhit.2015.8.2.20.

[13] F. Okubo, T. Yamashita, A. Shimada, and H. Ogata, "A neural network approach for students' performance prediction," in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, Mar. 2017, pp. 598–599, doi: 10.1145/3027385.3029479.

[14] S. Helal *et al.*, "Predicting academic performance by considering student heterogeneity," *Knowledge-Based Syst.*, vol. 161, pp. 134–146, Dec. 2018, doi: 10.1016/j.knosys.2018.07.042.

[15] V. Muralidharan, P. M., and J. Balaji, "Result Prediction Using K-Nearest Neighbor Algorithm For Student Performance Improvement," *Int. Res. J. Electron. Comput. Eng.*, vol. 3, no. 1, p. 7, Mar. 2017, doi: 10.24178/irjece.2017.3.1.07.

[16] A. S. Nagesh, C. V. S. Satyamurty, and K. Akhila, "Predicting student performance using KNN classification in bigdata environment," *CVR J. Sci. ...*, vol. 13, no. December, 2017, [Online]. Available at: http://cvr.ac.in/ojs/index.php/cvracin/article/view/61.

[17] T. Tanner and H. Toivonen, "Predicting and preventing student failure – using the k-nearest neighbour method to predict student performance in an online course environment," *Int. J. Learn. Technol.*, vol. 5, no. 4, p. 356, 2010, doi: 10.1504/IJLT.2010.038772.

[18] N. Iam-On and T. Boongoen, "Improved student dropout prediction in Thai University using ensemble of mixed-type data clusterings," *Int. J. Mach. Learn. Cybern.*, vol. 8, no. 2, pp. 497–510, Apr. 2017, doi: 10.1007/s13042-015-0341-x.

[19] F. Ouatik, M. Erritali, F. Ouatik, and M. Jourhmane, "Predicting Student Success Using Big Data and Machine Learning Algorithms," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 12, pp. 236–251, Jun. 2022, doi: 10.3991/ijet.v17i12.30259.

[20] J. Xu, K. H. Moon, and M. van der Schaar, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 5, pp. 742–753, Aug. 2017, doi: 10.1109/JSTSP.2017.2692560.

[21] F. Benslimane and K. Askia, "Tracking and predicting student performance in university," pp. 1-14, 2018. [Online]. Available at: https://faresbs.github.io/downloads/fares2018.pdf.

[22] A. Mondal and J. Mukherjee, "An Approach to Predict a Student's Academic Performance using Recurrent Neural Network (RNN)," *Int. J. Comput. Appl.*, vol. 181, no. 6, pp. 1–5, Jul. 2018, doi: 10.5120/ijca2018917352.

[23] M. A. Amoo, O. B. Alaba, and O. L. Usman, "Predictive modelling and analysis of academic performance of secondary school students: Artificial Neural Network approach," *Int. J. Sci. Technol. Educ. Res.*, vol. 9, no. 1, pp. 1–8, May 2018, doi: 10.5897/IJSTER2017.0415.

[24] H. A. Ali, C. Mohamed, B. Abdelhamid, N. Ourdani, and T. El Alami, "A Novel Hybrid Classification Approach for Predict Performance Student in E-learning," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 147, Springer Science and Business Media Deutschland GmbH, 2023, pp. 45–52, doi: 10.1007/978-3-031-15191-0_5.

[25] A. Das and S. N. Pradhan, "Area-Power-Temperature Aware AND-XOR Network Synthesis Based on Shared Mixed Polarity Reed-Muller Expansion," *Int. J. Intell. Syst. Appl.*, vol. 10, no. 12, pp. 35–46, Dec. 2018, doi: 10.5815/ijisa.2018.12.04.

[26] A. M. U. D. Khanday, S. T. Rabani, Q. R. Khan, N. Rouf, and M. Mohi Ud Din, "Machine learning based approaches for detecting COVID-19 using clinical text data," *Int. J. Inf. Technol.*, vol. 12, no. 3, pp. 731–739, Sep. 2020, doi: 10.1007/s41870-020-00495-9.

[27] L. Drukker, J. A. Noble, and A. T. Papageorghiou, "Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology," *Ultrasound Obstet. Gynecol.*, vol. 56, no. 4, pp. 498–505, Oct. 2020, doi: 10.1002/uog.22122.

[28] S. R. Richter, V. Vineet, S. Roth, and V. Koltun, "Playing for Data: Ground Truth from Computer Games," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9906 LNCS, Springer Verlag, 2016, pp. 102–118, doi: 10.1007/978-3-319-46475-6_7.