University of Ibadan
Journal of Science and Logics in ICT Research

Evaluation of Three Prediction Models for Identification of At-Risk First-Year Students

[1]OLADAYO, O., [2]AYENI, J. A. AND [3]OSUNADE O.

[1,2]Dept of Computer Science Aajayi Crowther University, Oyo, Nigeria.
[3]Dept of Computer Science, University of Ibadan, Nigeria.
[1]osatuoreoladayo@gmail.com, [2]ja.ayeni@acu.edu.ng, [3]seyiosunade@gmail.com

Abstract
Education is a tool and means of achieving life goals that requires knowledge acquisition through diligence, practice and demonstration. Tertiary institutions adopt many techniques to educate students thus the success rates differ. Some students perform better when visual, textual or auditory based methods are used for teaching. Private tertiary institutions with limited student population are interested in retaining existing students' and not expelling them. Additional services such as guidance, counselling and mentoring have been introduced to reduce failure rates. These additional services are only useful if the weak students are identified early. In this work, data about first-year students in Computer Science was used to predict their academic performance and identify the students at-risk of failure in the first year. Naïve Bayes Algorithm, Decision Tree Algorithm and Support Vector Machine Algorithm were used to develop the predictive models. The results of the models demonstrated a prediction accuracy of 100%, a 0% classification error and a runtime of 505 milliseconds for the best model. Early identification of weak students will enable appropriate help to be activated for such students early in their academic life.

***Keywords**: Classification algorithm, Educational data mining, Decision tree, Predictive model, Machine Learning Algorithm*

1. INTRODUCTION

The development of any nation is largely dependent on the strength and quality of her educational sector. This is because the educational sector, most especially the tertiary educational sector has the responsibility of producing graduates who have been equipped with the physical tenacity, emotional intelligence and the intellectual capacity to handle life's challenges and contribute to the development of the nation. The Federal Republic of Nigeria through her national policy on education defines a tertiary institution education as any form of education obtained after a secondary education [1].

According to Kaufman and Bradbury [2], an "at-risk" student is generally defined as a student who is likely to fail at school and in this context, school failure is typically seen as dropping out of school before high school graduation or from a tertiary institution. In their report [2], it was further stated that as a result, the characteristics of at-risk students have traditionally been identified through retrospective examinations of high school dropouts' family and school histories and those characteristics associated with dropping out of school then become the defining characteristics of at-risk students.

Parente and Sherer [3] defined an algorithm in data mining (or machine learning) as a set of heuristics and calculations that creates a model from data stated further that to create a model, the algorithm first analyzes the data you provide, looking for specific types of patterns or trends.

The results of this analysis is then applied over many iterations by the algorithm to find the optimal parameters for creating the mining model and thereafter applied these parameters across the entire data set to extract actionable

Oladayo, O., Ayeni, J. A. and Osunade, O. (2021). Evaluation of Three Prediction Models for Identification of At-Risk First-Year Students. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 7 No. 1, pp. 29 -38

patterns and detailed statistics [3]. Choosing an algorithm in data mining is often very challenging as Data Mining allows building of multiple models on a single mining structure and within a single data mining solution one could use a clustering algorithm, a decision trees model, and a Naïve Bayes model to get different views of the data [3]. A multiple algorithms could also be used within a single solution to perform separate tasks [3]. Again, different algorithms can be used to perform the same business task, each algorithm producing a different result, and some algorithms can produce more than one type of result [3].

In the development of the predictive models, three known algorithms were used; Naïve Bayes Algorithm, Decision Tree Algorithm and Support Vector Machine Algorithm. Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem [4] with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. It is not a single algorithm but a family of algorithms where they all share a common feature and every pair of features being classified is independent of each other [4].

Sunil [5] defined "Support Vector Machine" (SVM) as a supervised machine learning algorithm that can be used for both classification or regression challenges but used in classification problems. In the SVM algorithm [5], each data item is plotted as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate and then, perform classification by finding the hyper-plane that differentiates the two classes [5].

Akshay [6] defined Decision Tree Algorithm as a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. The decision tree Algorithm belongs to the family of supervised machine learning algorithms that can be used for both a classification problem as well as for regression problem [6].

Akshay [6] further stated that "the goal of this algorithm is to create a model that predicts the value of a target variable, for which the decision tree uses the tree representation to solve the problem in which the leaf node corresponds to a class label and attributes are represented on the internal node of the tree" [6].

The rate of graduation of students in private and public universities is very low due to poor academic performance and finances. The poor academic performance can be improved if such students are identified early and remedial action taken. The aim of this study is to develop data mining models to predict the academic performances of undergraduate students in their first-year using standard academic performance data.

2. Literature Review

The performance of a person is defined as the measurable or observable behaviour of that person in a given situation [7]. This could be in a test or an examination setting. The academic performance of students is of great interest to every academic institution as students are the basic reason for their existence. Tienken and Wilson [8] belong to this school of thought as they believe that the preparation of students academically should be the primary focus of every tertiary institution.

There are others who believe that other social institutions such as the community and family should be included into the educational processes of students [9]. The encouraging words of a father or mother can go a long way in motivating a student towards academic excellence. The academic performance of a student lies on the shoulders of educational institutions, the parents and the lecturers [9].
In the past, several researchers have carried out studies on the various methods employed in measuring the academic performances of students and include grade point average, standardized test scores and drop-out rates [10, 11].

Findings from research studies carried out identified students' prior knowledge [12], prior achievement [13] and standard test [14] as the major factors that influence the academic performance of students. Prior knowledge in a domain according to [15] and Dochy *et al.* [12], is said to be the knowledge an individual has of a particular area or field such as Mathematics or Physics.

Dochy *et al.,* [12] concurring with the above definition given by Alexander [15] and Dochy *et al.* [12], believe that the prior knowledge an individual has in a domain is that knowledge that was already obtained and resident in the mind of that individual and available to be used before that individual carried out a learning process and that the available knowledge contains conceptual and meta-cognitive knowledge components. Such knowledge can be measured by examining all areas needed as a foundation in order to be able to carry on in whatever field chosen to study.

The prior knowledge in a domain is considered as all compulsory subjects needed to be passed with at least a credit score in either the Senior School Leaving Certificate Examination (SSCE) or in the National Examination Council (NECO), before a candidate can be qualified to study in a particular programme at the University. This forms the foundational knowledge the individual needs to excel in that department.

Hailikari *et al.,* [16], conducted a research to find out the degree to which a prior knowledge in a particular domain can affect the performance of students. The result obtained showed that students' prior knowledge in a domain was the strongest variable when compared with 'self-belief' that affects the academic performance of students. Also worthy of note here is the fact that the quality of the prior knowledge in a domain will determine the success of a student.

Prior Achievement is defined here as the cumulative Grade Point Average (GPA) and not the grades students obtain in each course. Hicks and Richardson [17] carried out a descriptive analysis of some accounting student data to find out the effect of prior knowledge and prior achievement on student performance. They found out that there was a correlation between overall GPA and their diagnostic score which is the prior knowledge in a class. They also found a very strong relationship existing between students' overall GPA and the final grade.

Another similar research carried out by Harachiewicz *et al.,* 18], revealed that the performances of students from the start have an impact on their long-term success academically. From the results of these researchers, it could be deduced fact that the Cumulative Grade Point Average CGPA of first-year students has an impact on the long-term success and even determines whether or not students will drop out of a program.

In order to gain admission into tertiary institutions in Nigeria, a student must pass the unified tertiary matriculation examination (UTME) which is conducted by the Joint Admission Matriculation Board and also a Post UTME examination. This implies that the UTME and Post UTME examinations are the two main examinations aside WAEC and NECO a student must write and pass to be qualified for admission into any of the tertiary institutions in Nigeria. The UTME exam is a set of multiple-choice objective questions [19].

A student seeking admission, into the higher institution is expected to write the UTME examination in three subject areas relevant to his or her field of proposed study including the English language as compulsory. As highlighted by Burd [20], several advantages of the UTME examination include the potential of using it as a field diagnostic:

i. The potential of using it as a field diagnostic test is high.
ii. It gives allowance for more objective scoring.
iii. Both the high and low language ability groups of students are favoured
iv. They are relatively more reliable than the essay test.

Such tests can be easily scored by computers and other scoring machines and even by unskilled personnel. Gbore [21] conducted a research to find out the correlation between the SSCE, UTME, National Diploma and National Certificate of Education (NCE) using data from six Universities in the south west of Nigeria and concluded that there was a low correlation between UTME and students' CGPA.

Ogbebor [14] opined that the UTME was a more effective predictor of students' academic performance than the post UTME and even went further to conclude that the conduct of individual examinations for students seeking admission by some universities was a mere charade. Osakuade [22] believes differently stating that the post UTME was a better way to predict students' academic performance than the UTME, also encouraging JAMB to continue in the practice of conducting the UTME as a pre-qualifying examination into the university.

The academic performance of a student is the measurement of his achievement in various subjects over a period of time. The measurement of students' academic performance is crucial for the following reasons:

i. It is a way to actually find out if students are understand what lecturers teach; it helps to determine the extent to which the learning outcomes are being accomplished;
ii. It also helps in determining the effectiveness of a curriculum/program and how to improve on them;
iii. It helps educational institutions determine the best teaching approach to take in order to aid learning process; and
iv. Students are able to know their performance level and are able to make decisions that will improve their academic performance.

Hargis [23] agrees with the idea of measuring the performances of students as he believes it will help students stay motivated and goal oriented. As such, students are motivated towards their study and create personal educational goals that will propel them towards academic excellence and that motivation is a key factor in determining student scores in an examination [24].

Poor academic performance could be defined as one that falls below a set minimum performance standard in a test, examination or continuous assessment. The problem of poor academic performance is a major concern to several stakeholders including parents who spend a fortune in sending their wards to private university with the promise of a brighter future. There are several factors that contribute to the poor academic performance of students. Some of these factors include:

a. The lack of interest in the course and as a result,
b. They have little or no motivation to study,
c. Unstable emotional and behavioral attitude.

In a situation where students begin to perform poorly in school, the resulting consequences are usually even more alarming and damaging as such students can be withdrawn from the university. Some of them refuse to return home but rather hang around the school premises, getting involved in illegal and immoral activities which can cost them their lives. This is a problem that can be avoided by detecting "at-risk" students early enough to render the needed assistance towards the improvement of their academic performance.

3. Educational Data Mining and Machine Learning Technologies

Several experiments have been conducted to buttress the importance and potentials of data mining and machine learning in the educational sector. This is because of the fact that universities are more and more now being held responsible for the success of students. Varghese *et al.*, [25] clustered 8000 students in respect of five variables: input average in the university, average scores of the test/exams, average scores of papers, seminar notes, and a note of the work by frequency. The K-means algorithm was used and the results showed that there was a strong relationship between the attendance of students and their performances. Another interesting work was carried out by

Grafsgaard *et al.* [26] to recognize the facial expressions of students in the classroom based on frustrations or the understanding of students with the use of algorithms to discover unspoken behaviours and associate them to the information they acquired. The use of the Naïve Bayes algorithm, to predict the performances of students using thirteen variables especially "at risk" students thus a guidance and counselling program was activated to mitigate the effect in Bhardwaj and Pal [27].

Luan [28] also made use of data mining techniques to categorize students in order to determine students who can easily pile up course work and those who can take courses for a longer period of time. In a research carried out by Minaei-Bidgoli Sharma [31] are of the opinion that knowledge obtained via data mining has the potential to enhance the educational system in orientation, student performance and organization management.

4. Methodology

In this study, data mining and machine learning predictive models were developed. The three models are Naves Bayes, Decision Tree and Support Vector Machine. The models were used to predict the performances of students academically in the Computer Science Department using the combination of three variables.

The variables used are the student's first year result i.e. Grade Point Average (GPA) or Cumulative Grade Point Average (CGPA), Unified Tertiary Matriculation Examination (UTME) score and their O'level subject results. The data analysis was done using RapidMiner version 9.2.Data was collected from Ajayi Crowther University for four consecutive sessions for a total of four hundred and sixty (460) students from the Department of Computer Science. The data consist of all the elements needed for the research. Distance-base and local-base outliers' factors were also checked with none found. The data provided has three variables:

a. Grade Point Average (GPA) or Cumulative Grade Point Average (CGPA) indicates the full measure of a student cognitive level after the first and second semester of the first year.
b. University Tertiary Matriculation Examination (UTME) score. The score measures a student's cognitive level in four subjects that are required to obtain admission into a course at the University. The subjects are Mathematics, English, Chemistry and Physics for Computer Science.
c. O' Level Subject results for five science subjects at the ordinary level examination conducted by a national or regional body. This variable represents the totality of the knowledge a student possesses as a pre-requisite to studying Computer Science at Ajayi Crowther University.

The collected data were in different measurement scales with an established standard unit for all variables. All numerical values of the variables were normalized. The normalized data for GPA and CGPA was obtained by dividing the initial value of data by its range i.e. between 0 and 5.00. Hence for a student that obtained a CGPA of 3.55, the normalized CGPA of such a student will be 3.55/5.00 = 0.71. The University Matriculation Tertiary Examination (UTME) was obtained by dividing the score by the total mark possible. Thus, a student with UTME score of 250 will have a normalized value of 250/400 = 0.625.

The O' (Ordinary) Level subject results from regional and national examination bodies such as West African Examination Council (WAEC) and National Examination Council (NECO)use the same letter grades for results. The results of the five required subjects tendered to the university which indicates the domain prior knowledge were added together and divided by 500 to get the normalized score.

The normalized data was used in the development of the predictive models using RapidMiner. For developing and training the models in RapidMiner, 70% of the student data collected was used while the remaining 30% of the data was used in the testing and validation of the models. To identify at-risk students within the 30% data, the first semester results in the session i.e. GPA was used to predict the second semester results i.e. CGPA. Available CGPA at the end of the first session was used to validate the models.

The evaluation metrics for the prediction models was classification error, prediction accuracy and runtime. Classification error is the margin of error in grouping the data correctly and in percentages. To determine the prediction accuracy, the percentage of accurate predictions was calculated as the number of accurate predictions divided by the total number of predictions. Runtime is the time it takes the model to process the data given to it.

5. Modelling and Results

The modelling was carried out using Rapid Miner and the results are provided as tables and screenshots. In Table 1, the range of values that are generated for student CGPAs during normalization is provided. This table shows how the student performance in the university is classified by the three algorithms. Students with a low interpretation are inferred to have academic problems in the first semester or session.

Table 1: Range of Values for GPA/CGPA Normalization and Interpretation

GPA/CGPA	Normalized GPA/CGPA	Interpretation
0.00 – 1.xx	0.00 - 0.30	Low
1.xx – 2.xx	0.31 – 0.49	Average
2.xx – 5.00	0.50 – 1.00	Good

Student admission into the university is based on obtaining average or good interpretation as shown in Table 2. This implies that most students will have normalized JAMB scores in the range 0.39 to 1.00. The regulatory body sets the JAMB score used for admission and this is usually 150.

Table 2: Range of Values for JAMB Score Normalization and Interpretation

JAMB Score	Normalized JAMB Score	Interpretation
0 – 150	0.00 - 0.38	Low
151 – 200	0.39 – 0.50	Average
200 – 400	0.51 – 1.00	Good

The interpretation of results shown in Table 3 is based on five subjects required to study Computer Science at the tertiary level. The subjects are Mathematics, English Language, Physics, Chemistry and any other subject. Table 3 is generated using the scoring table of West African Examination Council (WAEC). Fail indicates that all results were F9 (i.e., 0 – 39%). The Pass category is made up of C6 to D8 i.e., 40 – 54%, while the good category is for marks from 55 – 100% i.e., C5 to A1. The five subjects must have been passed prior to admission hence no student is expected to have normalized values below 0.39.

In Table 4 sample normalized data for the three variables are shown. The GPA/CGPA varies from probation to excellent i.e. 0.282 to 0.712.

The UTME score is above low indicating that all students met the minimum score for entrance into the university. The O‘ level results for the five science subjects are in the pass range for the students.

Table 3: Range of Values for O’ Level Subject Results Normalization and Interpretation

O’ Level results for 5 subjects	Normalized O’ Level subject results	Interpretation
F9	0.00 - 0.39	Fail
E8 – C6	0.40 – 0.54	Pass
C5 – A1	0.55 – 1.00	Good

Table 4: Sample of Normalized Student Data

Student ID	GPA/ CGPA	UTME Score	O’ Level
1.	0.408	0.485	0.610
2.	0.282	0.418	0.580
3.	0.712	0.418	0.580
4.	0.248	0.405	0.550
5.	0.534	0.458	0.620
6.	0.462	0.465	0.620
7.	0.580	0.485	0.610
8.	0.480	0.498	0.570

The correlation matrix in Table 5 was obtained using K-means algorithm. The correlation matrix determines the significance of each variable and their dependence on each other. Table 5 indicates that all variables are statistically significant in the prediction of student academic success. The result is consistent with previous research findings that the three variables; the CGPA at the end of the second semester, the student UTME Score and O’ Level result determine the academic capacity of a student.

From Table 5, the UTME score a student obtains has a79% significance on the CGPA a student obtains at the end of the first year. The O’ Level subject result has a 75% significance on a student CGPA at the end of the first year.

This implies that the better a student performance is in the UTME and O' Level subject results, the higher the possibility of that student not being at-risk due to poor academic performance.

Table 5: Correlation Matrix of Normalized Variables

	GPA/CGPA	UTME	O-level
CGPA	1	0.799	0.754
UTME	0.799	1	0.825
O-level	0.754	0.825	1

Table 6: Prediction Accuracy of the Different Algorithms

MODEL	ACCURACY	STANDARD DEVIATION	RUNTIME	CLASSIFICATION ERROR
Naïve Bayes	100%	± 0.0%	504ms	0%
Decision Tree	96.7%	±7.5%	522ms	3.3%
Support Vector Machine	96.7%	±7.5%	376ms	3.3%

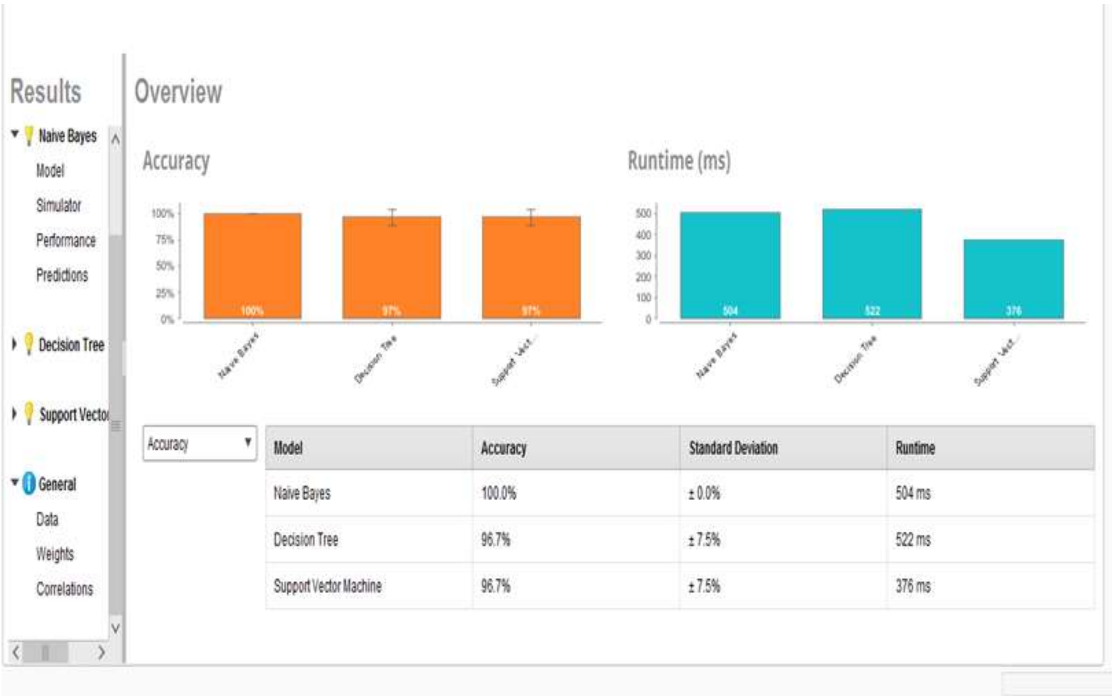

Figure 1: Prediction Accuracy of the three algorithms

Table 7 shows a sample output from the predictive model.

Student ID	CGPA	JAMB Score	O'Level	Prediction
1.	0.408	0.485	0.610	Average
2.	0.282	0.418	0.580	At-risk
3.	0.712	0.418	0.580	Good
4.	0.248	0.405	0.550	At-risk
5.	0.534	0.458	0.620	Good
6.	0.462	0.465	0.620	Average
7.	0.580	0.485	0.610	Good
8.	0.480	0.498	0.570	Average

Table 6 shows results for prediction models built with Support Vector Machine, Naïve Bayes and Decision Tree algorithm. The results were compared and the algorithm attached to the model with the highest prediction accuracy was selected as the best model. The Naïve Bayes algorithm classified new students having the highest prediction accuracy of 100% and a classification error of 0% but with a runtime of 505 milliseconds. Support Vector Machine had a prediction accuracy of 96.7%, a standard deviation of ±7.5%, and a classification error of 3.3% and a run time of 376 milliseconds. The decision tree had an accuracy of 96.7%, a standard deviation of ±7.5%, and a classification error of 3.3% and a runtime of 522 milliseconds. Based on the results in Figure 1, the Naïve Bayes algorithm model was selected as the best predictive model.

One of the objectives of this study is to identify academically at-risk students. The predictive model developed by the Naïve Bayes algorithm was able to predict the performances of all 48 students.

6. Results Ad Discussions

From the analysis and results presented earlier on the developed models, Naïve Bayes algorithm classified new students having the highest prediction accuracy of 100% and a classification error of 0% but with a runtime of 505 milliseconds. Support Vector Machine had a prediction accuracy of 96.7%, a standard deviation of ±7.5%, and a classification error of 3.3% and a run time of 376 milliseconds. The decision tree had an accuracy of 96.7%, a standard deviation of ±7.5%, and a classification error of 3.3% and a runtime of 522 milliseconds.

The predictive model developed by the Naïve Bayes algorithm was able to predict the performances of all 48 students given to it to predict, which included academically at-risk students with an accuracy of 100%. Amongst the test data fed to the Naïve Bayes model, a total of 8 students were academically at risk having a probation status and the Naïve Bayes model predicted them all correctly.

Adopting this predictive model will give Ajayi Crowther University time to make decisions to take proactive actions to improve the performance of academically at-risk students. Proactive actions can include extra classes, one-on-one sessions and group discussions or peer tutoring. The developed model is not transferable to another institution but can serve as an example.

7. Conclusion

In this research, the aim and objects have successfully been realized with the development of a mathematical model with the combination of predictor variables that can predict the academic performance of students in the University (using Ajayi Crowther University as a test-bed). Three data mining and machine learning algorithms were used to predict performances of students and hence identify academically at-risk students. The algorithms (SVM, Naïve-Bayes and Decision Tree algorithms) were trained with data taken over a period of academic session and the results compared against each other with the best algorithm identified as the Naïve Bayes algorithm. The model would be useful to Ajayi Crowther University in predicting and identifying academically at-risk students.

Acknowledgement

Mr. E. O. Oyebode, the Coordinator of the Department of Computer Science, Ajayi Crowther University, Oyo, Nigeria is appreciated for providing access to the data so also the contributions of other PG students of the Department of Computer Sciences.

References

[1] Gambari, A.I., Obielodan, O. O. And Kawu, H. (2017). Effects of Virtual Laboratory On Achievement Levels and Gender of Secondary School Chemistry Students In Individualized And Collaborative Settings In Minna, Nigeria. The *Online Journal of New Horizons In Education* - January 2017 Volume 7, Issue URL:https://www.tojned. Net/Journals/tojned/Articles/V07i01/V07i01 - 11.pdf

[2] Phillip Kaufman and Denise Bradbury (1995). "Characteristics of At-Risk Students". National Education Longitudinal Study. Contractor Report. In U. S. Department of Education Office of Educational Research and Improvement NCES 92-042

[3] Parente, J. and Sherer, T. (2020). Data Mining Algorithms (Analysis Services – Data Mining. https://docs.microsoft.com/en-us/analysis-services/data-mining/data-mining-algorithms-analysis-services-data-mining?view=asallproducts-allversions.

[4] Sunil, R. (2020). 6 Easy Steps to Learn Naive Bayes Algorithm with codes in Python and R. https://www.analyticsvidha. com/blog/2017/09/naive-bayes-explained Accessed 20/11/2021

[5] Sunil, R. (2020). Understanding Support Vector Machine (SVM) algorithm from examples (along with code). https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/Accessed 20/11/2021

[6] Akshay Sharma (2021). Machine Learning 101: Decision Tree Algorithm for Classification https://www.analyticsvidhya. com/blog/2021/02/machine-learning-101-decision-tree-algorithm-for-classification/

[7] Simpson, J. A. & Weiner E. S. C. (1989). The oxford English dictionary (2nd ed.). Vol. 1. Oxford: Clarendon Press

[8] Tienken, C., & Wilson, M. (2001). Using state standards and tests to improve instruction. Practical Assessment, Research & Evaluation, 7 (13), 74-99.

[9] Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. Educational Policy Analysis Archives, 8 (1), 67-93.

[10] Burns, A., & Darling, N. (2002). Peer pressure is not peer influence. The Education Digest, 68, 4-6.

[11] Hijazi, S. T., & Naqvi, S. M. M. R. (2006). Factors affecting students' performance: A case of private colleges. *Bangladesh E-Journal of Sociology*, 3 (1), 65-99.

[12] Dochy, F., De Rijdt, C., & Dyck, W. (2002). "Cognitive prerequisites and learning". The Institute for Learning and Teaching in Higher Education and SAGE Publication,3(3) 265-284.

[13] Harachiewicz, J.M., Barron, K.E., Tauer, J.M., & Elliot, A.J. (2002). "Predicting success in College:A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshmen year through graduation. Journal of Educational Psychology, 94(3) 562-575.

[14] Ogbebor G.G. (2012). "Mode of Admission and Undergraduate Academic Performance; A Comparative Study in Delta State University". Journal of Educational and Social Research 2(3) ISSN 2240-0524195-204.

[15] Alexander, P.A. (1992). "Domain knowledge as a determinant for future learning: the use of knowledge state tests and knowledge profiles. Utrecht, London: Lemma.

[16] Hailikari, T., Ninakatajavuori, N. and Lindblom-Ylanne, S. (2008): The Relevance of Prior Knowledge and Instructional Design. American Journal of Pharmaceutical Education. November 2008. Source PubMed.

[17] Daraseyseh, M. (2018). Student Education Specific account. Researchgate.Net

[18] Harachiewicz, J.M., Barron, K.E., Tauer, J.M., & Elliot, A.J. (2002). "Predicting success in College: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshmen year through graduation. Journal of Educational Psychology, 94(3) 562-575.

[19] Kaplan, R.M., Saccuzzo, D.P. (2005). 'Psychological testing Principles, applications, and issues. United States; Thomson Wadsworth'.

20] Burd, R. (2012). "Test standardization". Retrieved from http://asr.ipmahr.org/2012/06/Tests Standardization

[21] Gbore, L.O. (2013). "Relationship between cognitive entry characteristics and the Academic performance of university undergraduates in South West, Nigeria.

Journal of Education and Social Research, 3(1) 19-24

[22] Osakuade, J.O. (2011). “Effectiveness of University Matriculation Examination and Post-University Matriculation Examination on the Academic Performance of Nigerian Undergraduate Students”. *Journal of Education and Practice,* 2(4) 59-65

[23] Hargis, C.H. (2003). Grades and Grading Practices. Obstacles to Improving Education to Helping At-Risk Students (2nd ed.) Springfield, IL: Thomas.

[24] Covington, M. V. (2000). Goal theory, motivation, and school achievement: An Integrative review. Annual Review of Psychology, 51, 171−200.

[25] Varghese, J., Bindiya, M., Tommy, J., & Unnikrishnan, A. (2010). “Clustering Student data to Characterize Performance Pattern”. *International. Journal. Adv. Computer Science Applications, 138-140.*

[26] Grafsgaard, J.F., Wiggins, J.B., Boyer, K.E., Wiebe, E.N. & Lester, J.C. (2014). “Predicting Learning and Affect from Multimodal Delta Streams in Task-Oriented Tutorial.

[27] Bhardwaj Brijesh Kumar and Pal Saurabh. (2011) “Data Mining: A prediction for Performance Improvement Using Classification”. *International Journal of Computer Science and Information Security*. IJCSIS 9(4)

[28] Luan, J. (2002). “Data Mining and Knowledge Management in higher education-Potential applications. In proceedings of AIR Forum, Toronto, Canada.

[29] Minaei-Bidgoli, B., Kortemeyer, G. Punch, W. (2004). “Enhancing Online Learning Performance: An Application of Data Mining Methods”. In Proceeding of Computers and Advanced Technology in Education.

[30] Cortez, P., Silva, A. (2008). “Using Data Mining to Predict Secondary School Student Performance”. 5th Annual. Future. Business. Technology. Conference. 2003(2000) 5-12.

[31] Gulati, P. & Sharma, A. (2012). “Educational data mining for improving Educational quality”. *International Journal of Computer Science and Information Security* 2(3) 648-650.