E-ISSN: 1115 - 8425 P-ISSN: 3027-012X
MPRI-JAEE | INTERNATIONAL JOURNAL OF ASSESSMENT & EVALUATION IN EDU.
VOL. 13 NO. 8 – JULY. 2026

BEHAVIORAL-FACTOR DECISION SUPPORT MODEL FOR PREDICTING ACADEMIC PERFORMANCE

MOHAMMED NASIR MUSA; SUFIYANU B. TUGGA; VICTORIA GOJE GREGORY; & KAMAL SALIHU

Department of Computer Science, School of Science and Technology, Adamawa State Polytechnic, Yola, Nigeria
Corresponding Author: nas9ja@gmail.com
DOI: https://doi.org/10.70382/mejaee.v13i8.0139

Abstract

Early identification of students at academic risk remains a major concern in higher education because timely intervention can improve retention, learning outcomes, and institutional effectiveness. Educational data mining research has shown that not only grades and test scores but also demographic, family, social, and behavioral variables that can be analyzed for decision support shape student performance. This study developed a behavioral-factor-based decision support model for predicting the academic performance of 178 undergraduate students in the Department of Computer Science, Adamawa State University, Mubi, Nigeria. This study was adapted from the first objective of a 2023 master's thesis dataset and model framework. The model integrated family support, student motivation, engaged time, extra-curricular participation, study behavior, and related contextual variables. Academic performance was operationalized as the cumulative grade point average, classified into two groups: ≥ 2.0 and ≤ 2.0. After preprocessing and a 70:30 training-testing split, the behavioral-factor framework was evaluated in Python Jupyter Notebook and Weka. The findings showed that Naïve Bayes provided the strongest predictive performance, recording 92.6% accuracy, 89.3% precision, 96.2% recall, and 92.6% F1-score in Python, whereas K-means performed weakly with 38.9% accuracy. This study demonstrates that a behavioral-factor model can serve as a practical

Keywords: academic performance, behavioral factors, decision support system, educational data mining, higher education, Naïve Bayes.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

decision-support mechanism for the early identification of academically vulnerable students. This study recommends the institutional adoption of behavioral monitoring and SPA for student support and academic planning.

Introduction

Predicting student academic performance has become a major research priority in higher education because institutions are increasingly seeking evidence-based strategies to reduce failure, dropout, and delayed completion. Predictive systems can help universities identify students who require early academic support, thereby improving institutional planning and student outcomes (Golding & Donaldson, 2006). Educational data mining has further expanded this possibility by enabling the extraction of meaningful patterns from large educational datasets for classification, clustering, and performance forecasting (Anuradha et al., 2015; Sayana, 2015).

Despite these advances, many academic performance models still rely heavily on conventional variables, such as examinations, quizzes, assignments, and attendance. Although such variables are important, they may not fully explain why students succeed or struggle in higher education. Previous research has shown that parental support, motivation, socioeconomic conditions, institutional environment, and study behavior can substantially influence performance outcomes (Chohan & Khan, 2010; MolokoMphale & Mavis, 2014; Amjed, 2018). This indicates that a behavior-oriented model may provide a more useful basis for intervention than one based solely on outcome measures.

The underlying problem addressed in this paper is that many institutional systems do not incorporate behavioral, family, and social factors into routine academic prediction and therefore provide limited support for proactive intervention. The source study specifically identified the need for a model that uses students' behavioral factors to predict academic performance and functions as a decision support system for stakeholders (Musa, 2023). This article refines that objective into an international journal format and focuses on the design, structure, and practical utility of the behavioral-factor model.

Accordingly, this study presents a behavioral-factor decision support model for predicting undergraduate academic performance and demonstrates its practical usefulness using empirical results from the underlying study. This study contributes to the literature by organizing student-related predictors into a coherent conceptual framework and showing that such a framework can support accurate classification when coupled with a suitable supervised learning method.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

Literature Review

Research on academic performance prediction has consistently shown that the task is multidimensional. Golding and Donaldson (2006) emphasized the importance of predictive modeling in HE environments for identifying patterns that may explain academic success or underperformance. Clustering and classification methods have been applied to identify hidden structures in student data and to forecast educational outcomes (Anuradha et al., 2015; Sayana, 2015).

A key insight from the literature is that classroom tests and examination scores do not exclusively determine academic performance. Chohan and Khan (2010) found that parental support significantly and positively influences academic achievement and self-concept among students. Similarly, MolokoMphale and Mavis (2014) reported that declining academic performance is associated with parental involvement, resource availability, staff morale, and student preparedness. These studies support the inclusion of family and contextual variables in predictive models.

Related evidence also shows that broader student information benefits predictive systems. Amjed (2018) demonstrated that student information system records can be effectively mined to predict academic performance while highlighting the relevance of demographic history, prior academic records, and general student information. In a review of clustering in educational data mining, Anuradha et al. (2015) argued that grouping students based on shared characteristics can assist institutions in identifying capability patterns and planning academic support. However, the same literature indicates that predictive success depends on the selection methods appropriate to the data structure and decision problem.

Oladele, Aro, Adegun, and Ogundokun (2019) demonstrated that test, quiz, and assignment scores are major predictors of academic performance, reflecting the traditional emphasis on direct academic indicators. While useful, such approaches may underrepresent the behavioral processes that shape achievement before the final grades are produced. Therefore, this study extends this line of inquiry by focusing on the behavioral, social, family, and motivational dimensions in a decision support model for higher education.

Study design and setting

This study was based on the first objective of the master's thesis titled Predicting Student's Academic Performance Using K-Means and Naïve Bayes Algorithms (Musa, 2023). The empirical setting was the Department of Computer Science, Adamawa State

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

University Mubi, Nigeria, while the thesis was submitted to Modibbo Adama University, Yola, Nigeria. The design was quantitative and predictive, with the intention of developing a model that could serve as a decision support system for student academic performance. Data were collected using a structured questionnaire from 178 undergraduate respondents in the 2022/2023 academic session at the 400 level. The source thesis reported the use of stratified random sampling and organized the questionnaire around student performance, interest, study behavior, engaged time, beliefs, family support, and grades (Musa, 2023).

Behavioral Factors in the Decision Support Model

The model was conceived as a behavioral-factor framework in which multiple student-related domains as independent variables influence academic performance. The conceptual structure reported in the underlying thesis grouped predictors into family support and student motivation, engaged time and extra-curricular activities, interest and study behaviors, and related personal, social, and school attributes (Musa, 2023).

Family support encompassed parental education, family communication, family size, and financial support to students. Engaged time reflected classroom attention, regular attendance, assignment completion, and the relationship between the lecturer and the student. Interest and study behavior included course interest, interest in teaching and learning methods, revision habits, study time, study-group participation, and use of extra learning references. These variables were combined to produce a decision-oriented representation of factors associated with academic outcomes.

Variables and outcome definitions

The dataset included demographic, family, behavioral, social, and academic variables. These included gender, age, address, state, occupation, parental status, mother's education and occupation, father's education and occupation, sponsor, family size, family relationship, school-choice reason, travel time, study time, past failures, school support, family support, participation in activities, internet access, romantic involvement, free time, class absence, going out, course understanding, course interest, lecturer relationship, and semester grades (Musa, 2023). Academic performance was operationalized as the cumulative CGPA grouped into two classes: GT2 for CGPA ≥ 2.0 and LE2 for CGPA ≤ 2.0.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

Preprocessing and data analysis

The source thesis reported standard preprocessing procedures, including data cleaning, categorical encoding, and feature handling, before the construction of the model (Musa, 2023). The dataset was partitioned into training and testing subsets using a ratio of 70:30. The model was implemented and validated in Python Jupyter Notebook and Weka to assess the analytical consistency across platforms.

Although this study focuses on the behavioral-factor model design, the practical validity of that model was assessed through predictive application. Naïve Bayes was used as the principal supervised classifier, and K-means was used for comparison as an unsupervised clustering approach. Accuracy, precision, recall, F1-score, and number of correct and incorrect classifications were used to evaluate performance.

Structure and utility of the behavioral-factor model

The behavioral-factor model provided a coherent decision support structure by integrating variables that are often considered separately in academic evaluation. Instead of treating performance as the product of test and examination scores, the model organized family support, motivation, engaged time, interest, and study habits into a unified predictor framework linked to academic outcomes (Musa, 2023). This structure is important because it supports the early identification of risk conditions before academic decline becomes fully visible in the final results.

The empirical usefulness of the model is reflected in the predictive results obtained when the framework was analyzed computationally. In particular, the model performed well when implemented with Naïve Bayes, indicating that the selected factor structure captured meaningful relationships between behavioral indicators and academic performance.

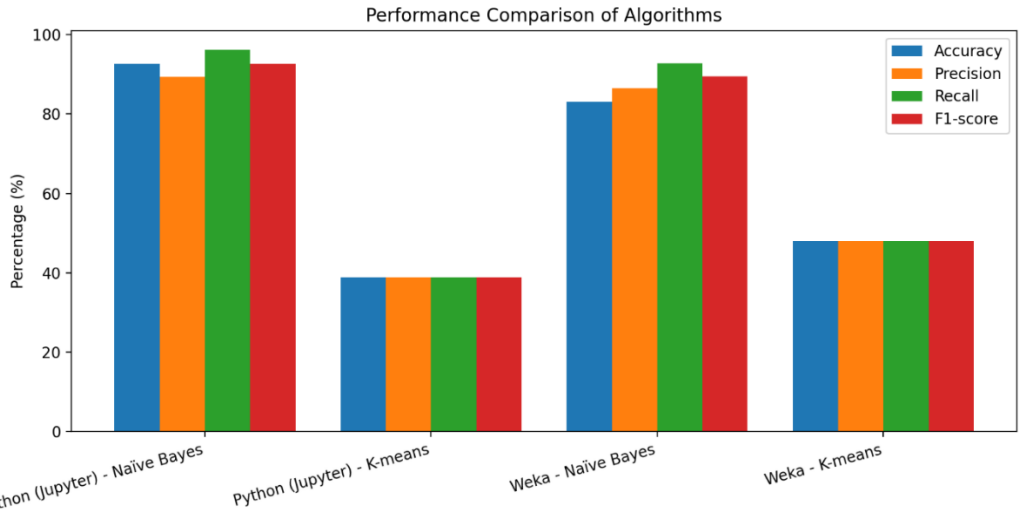
Predictive performance of the model

Table 1 presents the performance metrics reported for the BF model under Naïve Bayes and K-means implementations in Python and Weka (Musa, 2023).

Implementation	Algorithm	Accuracy	Precision	Recall	F1-score
Python (Jupyter)	Naïve Bayes	92.6%	89.3%	96.2%	92.6%
Python (Jupyter)	K-means	38.9%	38.9%	38.9%	38.9%
Weka	Naïve Bayes	83.1%	86.5%	92.8%	89.5%
Weka	K-means	48.0%	48.0%	48.0%	48.0%

Table 1. Predictive performance of the behavioral-factor model.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]



The source study further reported that, in Python, Naïve Bayes produced 165 correct predictions and 13 incorrect predictions, whereas K-means produced 69 correct predictions and 109 incorrect predictions. In Weka, Naïve Bayes achieved 148 correct and 30 incorrect predictions, whereas K-means produced 85 correct and 93 incorrect predictions (Musa, 2023). These findings indicate that although the behavioral-factor model is operationally useful, its success strongly depends on the analytical method employed.

Discussion

The findings show that a behavioral-factor model can provide a practical basis for academic decision support in higher education. This finding is consistent with the broader literature indicating that student achievement is shaped by more than just academic scores. For example, parental support and contextual conditions have been linked to improved academic outcomes (Chohan & Khan, 2010; MolokoMphale & Mavis, 2014), while educational data mining research has demonstrated the value of combining student-related attributes for prediction and institutional planning (Anuradha et al., 2015; Amjed, 2018).

A particularly important implication of this study is that the use of behavioral factors moves academic monitoring from a reactive to a proactive orientation. Instead of waiting until poor grades are fully established, institutions can monitor variables such as study time, attendance, lecturer relationship, course interest, and family support to identify students who may need earlier intervention. This aligns with the original rationale of Golding and Donaldson (2006) for predictive academic systems.

The comparison between Naïve Bayes and K-means also offers a methodological lesson. Although clustering approaches can be useful for exploring student groupings in

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

educational datasets (Anuradha et al., 2015; Sayana, 2015), the results of this study show that supervised classification was much more effective for the specific decision problem addressed in this study. Because the output variable was clearly defined in terms of the academic performance class, Naïve Bayes was better suited to learning the relationship between predictors and outcomes than unsupervised clustering. This result also differs from claims that modified K-means structures may offer strong predictive value in some settings (Sayana, 2015), indicating that model suitability remains context-dependent.

Nevertheless, the study has some limitations. The sample was limited to 178 students from a single department in one university, which limits the generalizability of the findings. The outcome measure was also reduced to two CGPA categories, which improves operational simplicity but limits granularity of interpretation. In addition, the source study did not present a detailed ranking of the importance of the variable, making it difficult to determine which behavioral factors contributed the most to prediction. Therefore, future studies should expand the dataset, incorporate multi-institutional samples, and apply feature-importance or explainable AI techniques to improve interpretability and policy value.

Conclusion

This paper presented a behavioral-factor decision support model for predicting the academic performance of higher education undergraduates. Drawing on data from 178 students and a model structure centered on family support, motivation, engaged time, study behavior, and related contextual variables, the study demonstrates that behavioral factors can be systematically organized for predictive academic decision-making. The strongest empirical results were achieved with Naïve Bayes, which substantially outperformed K-means. Therefore, this study concludes that HEIs can improve the early identification of academically vulnerable students by combining behavioral monitoring with SPA.

Recommendations

Higher education institutions should integrate behavioral-factor monitoring into student support and academic advisory systems.

In routine data collection, variables closely related to classroom engagement and study behavior, such as attendance, study time, course interest, and lecturer relationship, should be prioritized.

Indicators of family support should be considered in intervention planning because the literature and the present findings that the home and social contexts and school-based factors shape academic outcomes.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]

Future research should validate the model on larger and more diverse samples and compare additional SL techniques with EM approaches.

References

Amjed, T. M. A. (2018). Mining student information system records to predict students' academic performance. British University in Dubai. https://www.britishuniversity.com

Anuradha, C., Velmurugan, T., & Anandavally, R. (2015). Clustering algorithms in educational data mining: A review. International Journal of Power Control and Computation, 7(1), 47–52. https://doi.org/10.1016/j.ijpcc.2015.09.010

Chohan, B. I., & Khan, R. M. (2010). Impact of parental support on the academic performance and self-concept of the student. Journal of Research and Reflections in Education, 4(1), 14–26. https://doi.org/10.1016/j.jref.2010.09.010

Golding, P., & Donaldson, O. (2006). Predicting academic performance in School of Computing and Information Technology (SCIT). In Frontiers in Education: 36th Annual Conference, San Diego, CA, USA. https://doi.org/10.1109/FIE.2006.322661

MolokoMphale, L., & Mavis, B. M. (2014). An investigation on students' academic performance for junior secondary schools in Botswana. European Journal of Educational Research, 3(3), 111–127. https://doi.org/10.1016/j.ejeres.2014.09.010https://doi.org/10.1016/j.ejeres.2014.09.010

Musa, M. N. (2023). Predicting student's academic performance using K-means and naïve Bayes algorithms. Master's thesis, Yola-Nigeria: Modibbo Adama University.

Oladele, T. O., Aro, T. O., Adegun, A. A., & Ogundokun, R. O. (2019). Prediction of student's academic performance using k-means clustering and multiple linear regressions. Journal of Engineering and Applied Sciences, 14(22), 8254–8260. https://doi.org/10.1016/j.jengs.2019.08.016https://doi.org/10.1016/j.jengs.2019.08.016

Sayana, T. S. (2015). Prediction of students academic performance using data mining: Analysis. International Journal of Engineering Research & Technology (IJERT), conference proceedings. https://doi.org/10.1016/j.ijrt.2015.09.010.

EDITOR-IN-CHIEF
Prof. A.I. Mteiye [MALAWI]