Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 205 (2025) 15-28 EISSN 2392-2192

Bridging Educational Insights with Public Policy: Leveraging Predictive Analytics and Explainable AI for Student Academic Performance Outcomes Using PolySHAP

Gladys Chinyere Olumba[1,2,*], Happy Nkanta Monday[2,3], Grace Ugochi Nneji[2,3], Godwin Mark David[2,4], Wisdom Chima Olumba[2,5], Richard Iherorochi Nneji[2,6], Edwin Sunday Umana[2,7], Daniel Agbonifo[2,8]

[1]Department of Agricultural and Natural Resources, Abia State Local Government, Abia, Nigeria

[2]Intelligent Computing Lab, Hace SoftTech, Lagos, Nigeria

[3]OBU Computing, Chengdu University of Technology, China

[4]Business Administration and Management, University of Cross River State, Nigeria

[5]Department of Electrical Engineering, Reliable Steel and Plastics Industry, Anambra, Nigeria

[6]Department of Engineering, Ikelinsco Global Company, Nigeria

[7]Department of Mathematical Science, Federal University of Technology Akure, Ondo, Nigeria

[8]Department of Exploration, Neconde Energy Limited

*Author for Correspondence: adawissy@gmail.com

(Received 9 May 2025; Accepted 14 June 2025; Date of Publication 4 July 2025)

ABSTRACT

The advancement of predicting the performance of student is fast growing area of interest in educational domain, with significant implications for educational strategy and public policy. Machine learning (ML) has proven effective in investigating student data to recognize those at risk of poor outcomes, permitting for early support and intervention. Nevertheless, the complexity of these classifiers often poses interpretability issues, making it difficult for educators and policymakers to comprehend the underlying factors driving predictions. This paper proposes PolySHAP to enhance the transparency and interpretability of ML classifiers in educational settings. Numerous models are employed to identify and explain at-risk group of students based on certain features including academic, demographic, and behavioral. Proposing the PolyFeature technique is employed to improve model accuracy. The results demonstrate that the ensemble models perform better than single models and also the SHAP values effectively and efficiently decompose predictions into feature contributions, making the models' decisions interpretable for the development of policy. By connecting model insights to actionable interventions, this study provides architecture for data-driven educational policies aimed at enhancing student performance, giving room to equitable resource distribution and decreasing attrition. The investigations emphasize the requirement for a balance between explainability and predictive performance in constructing policies that support success of the students and equal resource allocation in educational domain.

Keywords: Student academic performance, machine learning**,** public policy, PolySHAP, public development, intervention strategies, predictive analytics, Explainable AI.

1. INTRODUCTION

Academic performance of student is a vital metric for parents, educators, schools, institutions and the policy makers as it solely not only showcase the success of the educational domain but also on the contribution to the society and economy at large [1]. Recognizing students at risk of poor academic performance or dropout is necessary for creating timely and effective interventions. With the introduction of big data and machine learning (ML) approaches, educational data mining has stand with the test of time as a powerful tool for predicting student academic performance based on numerous features like academic history, age, environment and socio-economic factors [2].

While conventional statistical models have been widely applied to predict academic outcomes, their predictive outcome is most times limited by assumptions of linearity and the incapability to handle complicated patterns in data [3]. In contrast, machine learning algorithms have revealed superior predictive capabilities by building complicated non-linear relationships in the data. Nevertheless, the "black-box" nature of these ML models poses an issue for interpretability, especially when these models are employed to inform public policy and assist interventions [4]. In the aspect of policymakers, not only knowing if the student is at risk but also understanding the reason behind certain conditions that influence their predicted outcomes.

Explainable AI models, such as SHapley Additive exPlanations (SHAP) values, resolve the issue by offering way to understand, justify and interpret the ML model predictions [5]. By disintegrating model predictions into the contributions of each attributes, XAI models maintain transparent decision-making and support specific policy development. This study leverages many supervised learning models, enhanced through feature selection approaches, to predict student academic performance. More so, the utilization of SHAP values covers the gap between interpretability and model accuracy, providing actionable insights for educational policymakers to implement data-driven strategies [6].

The main contributions of our work are as follow:

- The paper proposes PolySHAP values to enhance the interpretability of machine learning models, providing a transparent framework that explains how specific features contribute to a student's predicted risk of failure.
- The study introduces PolyFeature to enhance the predictive power of models while reducing dimensionality and focusing on the most impactful features.
- A diverse set of machine learning models are evaluated to predict student academic performance based on a wide range of demographic, academic, and behavioral attributes.
- Models are evaluated on multiple metrics offering a holistic view of model performance and applicability to real-world educational settings.
- Our work connects predictive analytics with actionable educational strategies, showcasing how explainable AI can aid public policy by offering insights into which students are most likely to be at risk and the reason behind it. This ensures more efficient and effective allocation of resource, early resolution, and specific interventions to enhance student outcomes.
- The study achieves a good balance between interpretability and model accuracy, maintaining that the ML predictions are both understandable and dependable to educators, parents and policymakers for well-informed decision-making.

The paper is structured as follows: Section 2 analysis the related works by some researchers and the implementation of public policy AI-based on student academic performance using ML. Section 3 discusses our methodology, including data collection and preprocessing steps, model implementation and SHAP explainability. Section 4 reviews the experimental results and analysis of the investigations. Lastly, Section 5 concludes the paper and presents the limitations and potential future trends for research in bridging educational insights with public policy.

2. LITERATURE REVIEW

The utilization of ML in predicting student academic performance has become increasingly prominent as a result of its potential to recognize students at risk early and ensure timely mitigations. Prior researches on data mining based on education primarily focused on regression and descriptive statistics analyses to assess factors affecting student success. Nevertheless, these approaches sometimes failed to capture complicated patterns in data, restricting their predictive power [1].

Current research on predicting student academic performance through the application of ML models has provided substantial contributions to the educational domain. Fan et al. [7] suggested the utilization of interpretable models to predict student achievement within different learning formats, including offline, online, and blended modes. This work highlights the importance of explainability in ML classifiers to facilitate understanding among concerned stakeholders. Also, it pays attention on different factors, including demographics, past academic performance of the student, and engagement, offering insights on how various models and features impact accuracy prediction. Furthermore, the research suggested how AI can significantly enhance educational outcomes and provides practical recommendations for model selection and application within higher educational domains.

Another vital contribution by Serrano et al. [8], examined college students' academic successful academic accomplishment prediction by utilizing ML classifiers. They utilized the CRISP-DM methodology to recognize essential features in student data and checkmated different classifiers for their predictive abilities. XGBoost model was seen to be highly effective and efficient in achieving an AUC of 87.75%, while Decision Tree classifier was highlighted for its interpretability potential. The study argues for the practical implementation of such classifiers to provide support in student retention approaches, university management and lastly, the reinforcement of policy interventions in education. To complement this, Albreiki et al. [9] checkmates the application of data mining and learning analytics over 10 years. The study reveals a holistic method to predicting student academic performance, pinpointing predictors like emotions, assessments, and so on and also emphasize on the increasing importance of ML in educational domain.

Numerous researches have particularly addressed the issue of forecasting student academic performance and dropout rate. Albreiki et al. [10] suggested a review of educational data mining approaches while paying rapid attention on utilizing ML to recognize student at-risk and enhancing the performance of the accuracy prediction. Their study focused on the prediction of the student academic performance using University dataset, thereby offering substantial opportunities for early support and mitigation. Likewise, Coussement et al. [11] investigated the utilization of logit leaf algorithms to predict student dropout rates, pinpointing on the essential of decision support systems to track and respond to real-time of the student academic performance. These strategies have been re-implemented by Gray & Perkins [12], who used early data engagement data and ML algorithms to forecast student academic risks, thus maintaining timely and effective educational strategies.

More so, the study of Alelyani et al. [13] investigated the numerous ML methods utilized to predict student academic performance, recognizing broader attributes that plays vital role in predicting outcomes. Despite the challenges faced in ascertaining the dataset benchmark, ML's application in educational domain still provides promising prospects for improving student academic performance and well-informed decision making. Generally, there still remain growing needs for explainable models with well robust ML models in order to maximize the prediction accuracy capabilities while maintaining interpretability for policy decisions.

3. METHODOLOGY

This section will elucidate the data collection and data preprocessing approaches, the optimized ML models, data splitting, evaluation metrics and environment setup utilized to assessed the proposed model performance.

3.1. Data Collection and Preprocessing

The dataset utilized in this study is obtained from the publicly available "Student Performance" dataset in the Kaggle dataset, which consists of student academic and demographic information, including their final grade, performance indicators, and socio-economic status. To reframe the analysis into a binary classification problem, students' final grades are recoded into binary outcomes, representing "Pass" and "Fail" categories. Specifically, a grade greater than 9 is labeled as "Pass" and a grade equal to or below 9 as "Fail." This allowed for a focused prediction of academic success or risk.

The dataset included both numerical and categorical features. Categorical features, such as “school,” “sex,” and “address,” are transformed using One-Hot Encoding, and numerical features are standardized using MinMaxScaler to ensure all values are non-negative and within a comparable range. Missing values are imputed using the mean for numerical features and the most frequent value for categorical features to maintain data consistency.

3.2. Feature Selection

To improve the predictive power of the model, our novel polyfeature is applied to create interaction terms between features. PolyFeature set at degree of 2 is used to generate higher-order features, capturing potential non-linear relationships in the data. Following this, SelectKBest with the F-statistic is used for feature selection. This method ranks features based on their capability to distinguish between the "Pass" and "Fail" classes, and the top 20 features are retained for further model training.

3.3. Model Development

A wide range of machine learning models is tested to ensure a comprehensive evaluation of predictive performance. Each model brought unique characteristics in terms of handling different data patterns and relationships. The chosen models are tuned and evaluated using random grid search with cross-validation to identify optimal hyperparameters for performance enhancement. The classifiers include Logistic Regression (LR), Decision Tree (DT), and Gaussian Naive Bayes (GNB), as well as more advanced ensemble techniques like Random Forest (RF), AdaBoost, and Gradient Boosting Machines such as LightGBM (LGBM) and XGBoost. The regularization strength parameter for LR and DT and also the learning rates for the LGBM, are systematically varied within certain range. The CatBoost model and the multilayer perceptron models are also considered as a result of its capabilities in capturing complex and complicated pattern in the data.

3.4. Model Evaluation Metrics

To checkmate the model performance, the dataset is splitted into an 80% for training set and 20% for the testing set to maintain the results generalizability. Performance evaluation metrics including Accuracy, Recall, F1-Score, and Specificity are utilized to measure the précised classification performance. Confusion matrix is employed to provide detailed insight into the number of correctly and incorrectly predicted score for both classes.

3.5. Polyshap Explainability

To improve the interpretability of this study, PolySHAP values are computed, which provide an underlining of the individual attribute’s contribution to its predictions, providing a clear and justifiable model decision-making process. The PolySHAP explainer is employed to the Random Forest model, and visualization plots such as force, bar, summary, waterfall and dependence plots are provided. These plots permit for a well detailed examination of how each features affected the prediction tasks, which is vital for understanding model behavior and necessitating public policy-related decision-making process.

3.6. Model Interpretability

Model interpretability, including accuracy, precision, sensitivity, F1-score and specificity as seen respectively in equation (1) to equation (5), are employed for each model to investigate and compare performance efficiently and effectively. The results are put together into a comprehensive table, summarizing all the performance metrics to recognize the model offering the best balance between interpretability and prediction accuracy. The selected classifier is then utilized to provide actionable insights from the data, helping to propose potential policy mitigations for educational success.

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

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

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

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

$$Specificity = \frac{TN}{TP + FN} \quad (5)$$

$where\ TP: True\ Positive, TN: True\ Negative, FP: False\ Positive, FN: False\ Negative$

3.7. Enviromental Setup

The model implementation of this study is carried out using the Python programming language and ML libraries, with a hardware configuration of 64GB RAM, 8GB GPU, and an 11th Generation Intel Core i7-11800H processor at 2.30GHz running on a Windows 11 operating system.

3.8. Hyperparameter Tweaking

The hyperparameter optimization of each classifier, based on the prediction of student performance outcomes in this study, enables us to refine the models for improved prediction accuracy and reliability. A randomized search cross-validation approach is employed to fine-tune the model parameters, ensuring the best possible fit for accurate and consistent predictions. Table 1 provides a comprehensive overview of the primary hyperparameters used in the prediction of student success and failure, offering insights into the tuning strategies that enhance model performance in this educational context.

Table 1. Random Grid Search Hyperparameter for PolySHAP for Student Academic Performance Outcomes.

Algorithm	Hyperparameter Explanation	Randomized Grid Search Range
Logistic Regression	C (Inverse regularization strength)	C: [0.01, 0.1, 1, 10, 100]
SVM (Support Vector Machine)	C (Regularization strength), kernel (Type of kernel), gamma (Kernel coefficient), probability (Probability estimates)	C: [0.1, 1, 10], kernel: ['rbf', 'linear'], gamma: [0.01, 0.1, 1], probability: [True]
Gaussian Naive Bayes	var_smoothing (Stability added to variance)	var_smoothing: [1e-09, 1e-08, 1e-07, 1e-06]
Random Forest	n_estimators (Number of trees), max_features (Max features for split), min_samples_split (Min samples to split), min_samples_leaf (Min samples at leaf), max_depth (Tree depth)	n_estimators: [50, 100, 200], max_features: ['sqrt', 'log2'], min_samples_split: [2, 4, 6], min_samples_leaf: [1, 2], max_depth: [None, 10, 20, 30]
Decision Tree	criterion (Split quality measure), min_samples_split (Min samples to split), max_depth (Tree depth)	criterion: ['gini', 'entropy'], min_samples_split: [2, 6, 10], max_depth: [None, 10, 20, 30]
AdaBoost Classifier	n_estimators (Number of boosting stages), learning_rate (Boosting learning rate)	n_estimators: [50, 100, 200], learning_rate: [0.01, 0.1, 1]
LightGBM	n_estimators (Number of boosting stages), max_depth (Tree depth), learning_rate (Boosting learning rate)	n_estimators: [50, 100, 200], max_depth: [-1, 10, 20], learning_rate: [0.01, 0.1, 1]
XGBoost	n_estimators (Number of boosting rounds), learning_rate (Boosting learning rate), max_depth (Tree depth), subsample (Fraction of samples used for fitting), colsample_bytree (Fraction of features used per tree)	n_estimators: [50, 100, 200], learning_rate: [0.01, 0.1, 0.2], max_depth: [3, 6, 9], subsample: [0.6, 0.8, 1], colsample_bytree: [0.6, 0.8, 1]

CatBoost	iterations (Number of boosting rounds), learning_rate (Boosting learning rate), depth (Tree depth), l2_leaf_reg (L2 regularization term), bagging_temperature (Controls overfitting)	iterations: [50, 100, 200], learning_rate: [0.01, 0.1, 0.2], depth: [4, 6, 10], l2_leaf_reg: [1, 3, 5], bagging_temperature: [0, 1, 5]
KNN	n_neighbors (Number of neighbors to use), weights (Weight function used in prediction), p (Power parameter for Minkowski distance), algorithm (Algorithm for computing nearest neighbors)	n_neighbors: [3, 5, 7, 9], weights: ['uniform', 'distance'], p: [1, 2], algorithm: ['auto', 'ball_tree', 'kd_tree', 'brute']
ANN	hidden_layer_sizes (Neurons in hidden layers), activation (Activation function), solver (Weight optimization algorithm), alpha (L2 penalty), learning_rate (Learning rate schedule)	hidden_layer_sizes: [(50, 50), (100,)], activation: ['logistic', 'tanh', 'relu'], solver: ['adam', 'sgd'], alpha: [0.0001, 0.001, 0.01], learning_rate: ['constant', 'adaptive']

4. RESULTS AND ANALYSIS

4.1. Polyshap Explainability

The provided SHAP plots in Figure 1 to Figure 5 offer comprehensive insights into the interpretability of the machine learning model's predictions in the context of student performance. The summary plot and bar plot further break down the feature importance across all predictions. The summary plot in Figure 1 uses a beeswarm format to display the SHAP value distribution for each feature, revealing the importance and effect (positive or negative) on the model's predictions. The most important attributes, such as "age goout" and "age failures," constantly show high values for SHAP values, showcasing their strong impact on academic success prediction. Nevertheless, the bar plot in Figure 2 sets attributes based on their mean absolute SHAP value, pinpointing "age goout" and "age failures" as the top contributors to the prediction of the model. The dependence plot in Figure 3 shows how certain attribute interacts with its SHAP value, depicting interactions with "age goout." This extensive visual analysis permits for transparent, data-driven insights, supporting actionable strategies in educational scenarios. Also, the waterfall plot in Figure 4 offers a detailed flow of how each attributes cumulatively lead to a specific prediction, demonstrating the additive nature of attribute contributions. Finally, the force plot in Figure 5 visualizes the impact of each feature on an individual prediction. In this case, features like "failures Pstatus_T", "failures goout," and "failures freetime," either push the prediction higher or lower from the base value, aiding to recognizing the main factors contributing to the prediction outcome. This offers a clear and understandable characteristic of how individual attributes influence the model's decision-making, offering personalized explanations for each student.

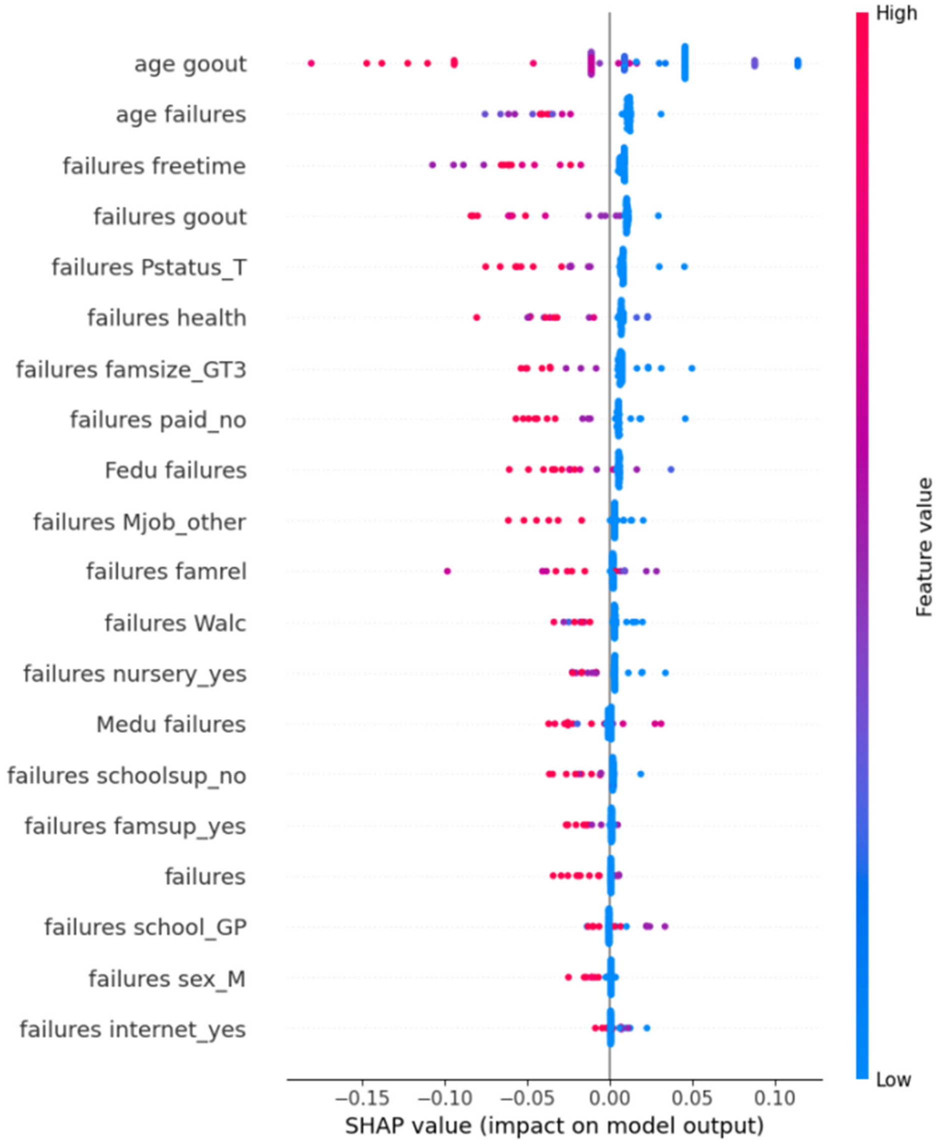

Figure 1. Summary Plot of the PolySHAP.

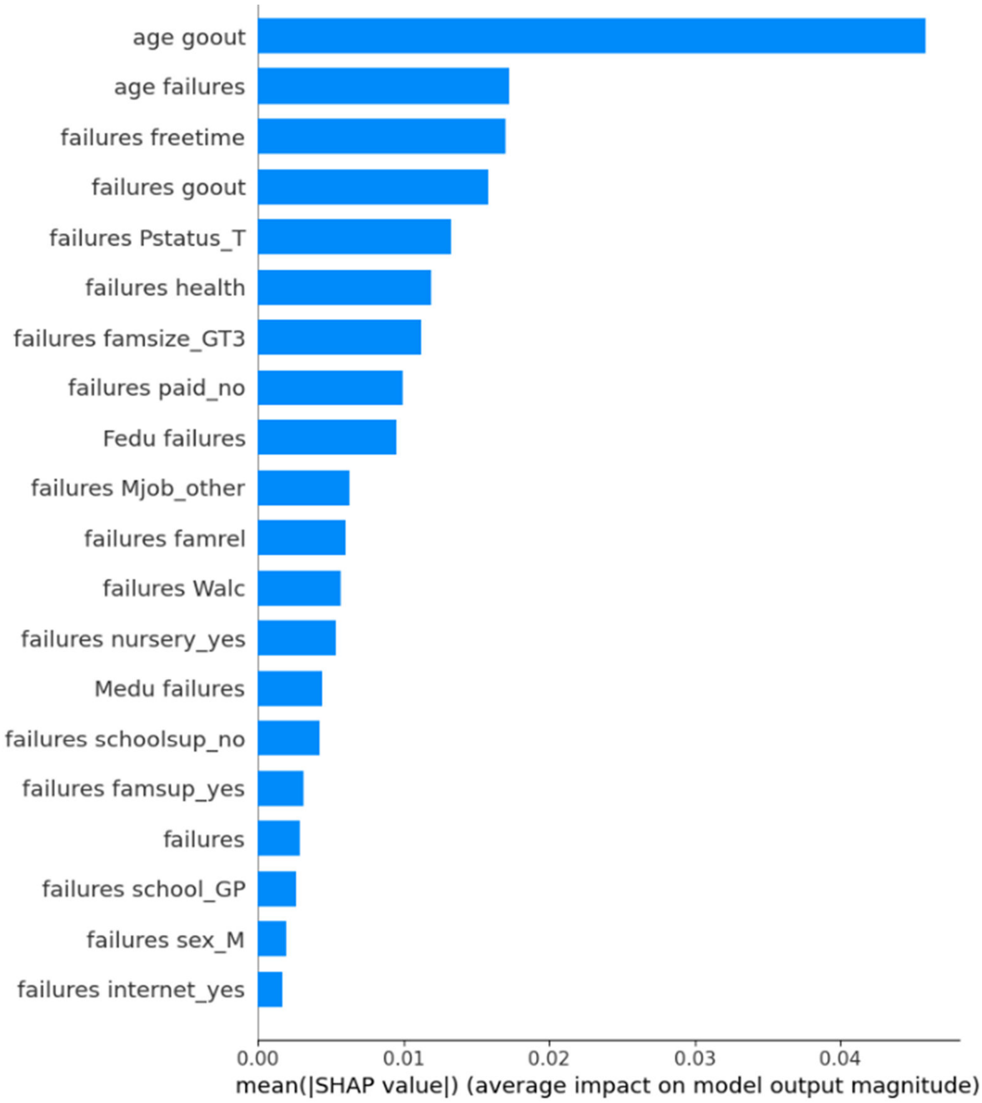

Figure 2. Bar Plot of the PolySHAP.

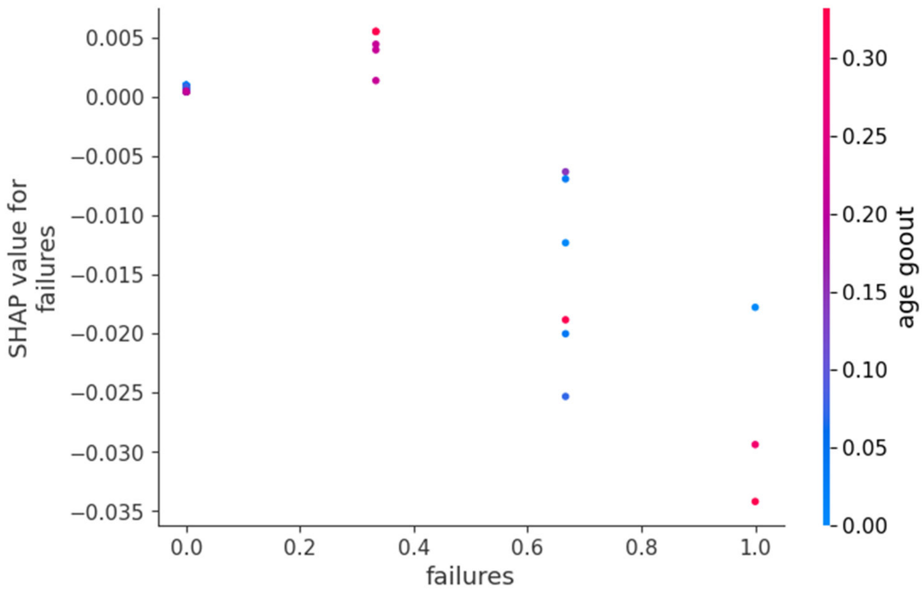

Figure 3. Dependence Plot of the PolySHAP.

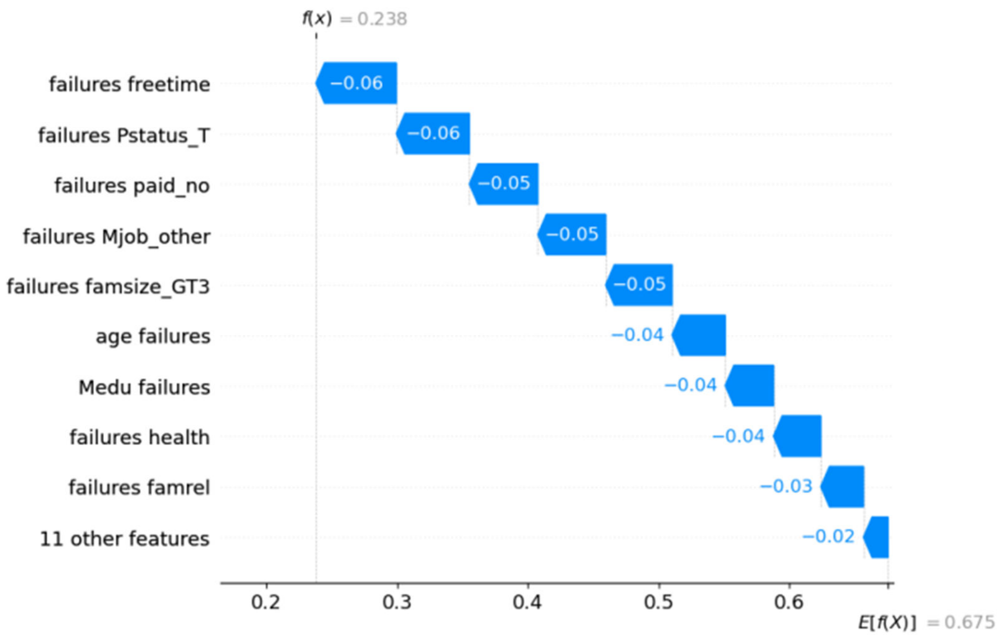

Figure 4. Waterfall Plot of the PolySHAP.

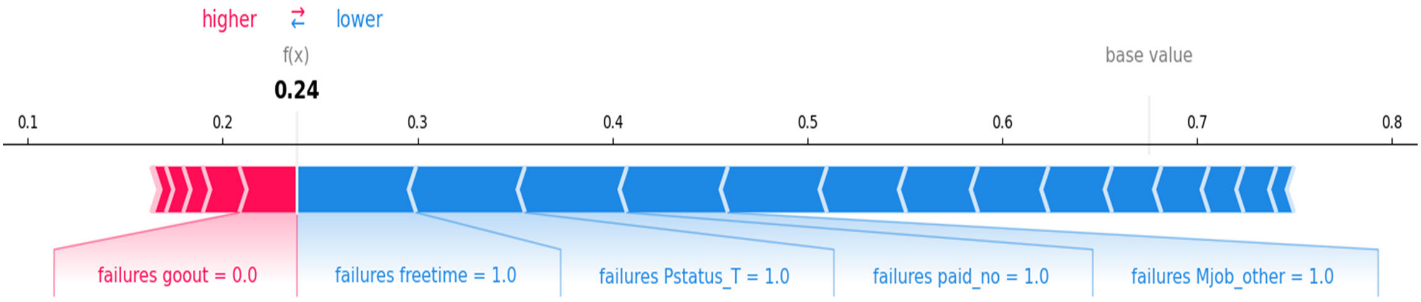

Figure 5. Force Plot of the PolySHAP.

Table 2 offers a comparative analysis of numerous ML classifiers utilized to evaluate the performance of the proposed POLYSHAP model across metrics including accuracy, precision, recall, F1-score, and training time. RF stands out with the highest accuracy of 97.22% and F1-score of 94.75%, indicating strong classification capabilities and recall of 96.15%. The ANN depicts excellent specificity of 97.04% and ROC-AUC of 98.91%, indicating a strong capability to differentiate between classes, particularly in recognizing true negatives. Other classifiers such as GNB, CatBoost, LGBM, and KNN display balanced performances amongst all evaluation metrics, with accuracies around 95.95%. However, DT underperforms in accuracy of 89.62%, F1-score of 90.33%, and precision of 90.00%, posing issues in attribute interaction capture and potential overfitting, while SVM ensures balanced recall of 94.23% and specificity of 95.93% but has lower accuracy of 90.89%. The utilization of polyfeatures in the POLYSHAP model improves performance by capturing complicated feature relationships, boosting classification accuracy, and enhancing specificity, specifically meriting RF and ANN. Nevertheless, for simple models like SVM and DT, the more complexity from polyfeatures has led to performance degradation. Training time varies significantly, with RF and ANN needing more time due to their complex frameworks, while models like GNB, SVM, and DT train faster but do not always achieve high performance. Generally, the utilization use of polynomial features improves the model performance, especially benefiting RF and ANN, while models like DT and SVM need more tweaking to fully leverage the polyfeatures technique.

Table 2. Performance Metrics of the Proposed POLYSHAP model with ML classifiers.

ML Model	ACC (%)	PRE (%)	REC (%)	F1-s (%)	SPE (%)	ROC-AUC (%)	Training Time (seconds)
LR	94.68	94.24	94.23	93.05	95.04	92.04	1.99
SVM	90.89	91.01	94.23	90.99	95.93	92.40	0.07
GNB	95.95	95.38	94.23	93.76	90.74	92.26	0.01
RF	97.22	95.76	96.15	94.75	90.74	92.61	0.88
DT	89.62	90.00	94.23	90.33	92.22	93.88	0.02
ADA	93.42	93.13	94.23	92.35	93.33	94.81	0.36
LGBM	95.95	95.38	94.23	93.76	90.74	91.47	0.79
XGBM	90.89	91.01	94.23	90.99	95.93	91.58	0.67
CatBoost	95.95	95.38	94.23	93.76	90.74	92.40	3.63
KNN	95.95	95.38	94.23	93.76	90.74	92.72	0.20
ANN	94.68	94.24	94.23	93.05	97.04	98.91	1.39

5. CONCLUSION

This study proposes a novel PolySHAP model for the extensive analysis of predictive classifiers for student academic performance, integrating various ML models. By incorporating the novel PolySHAP values, the study covers the gap between interpretability and model accuracy, permitting stakeholders like policymakers to achieve actionable insights from complex and complicated algorithms. The results depict those ensemble classifiers, such as RF, exhibit superior performance in terms of accuracy score, F1-score, and specificity, particularly when using polyfeature to capture higher-order relationships within the attributes. The utilization of PolySHAP improves the clarity of the model predictions, making it easier to understand and interpret how each attribute contribute to predicted outcomes. This transparency aids data-driven strategies aimed at enhancing student success rate and also assists inform public policy decisions on educational interventions and resource allocation. The investigations depict the potential of advanced ML approaches integrated with explainable AI, to improve predictive abilities and offer interpretable outcomes in educational domain.

Aside the promising results, the study encountered few limitations. Firstly, the dataset mainly focus on particular academic and demographic attributes, which fully did not encompass all the necessary factors influencing student academic performance. This limitation proffers that integrating more features, like psychological, and socio-emotional variables, most like will enhance the model predictions. More so, classifiers like DT and SVM, showed lower prediction performance when using the polyfeatures, showcasing that complicated feature selection might not be universally beneficial amongst all the models. Additionally, while PolySHAP values offer interpretability, their computational cost is expensive, particularly, for larger dataset and complicated datasets, which may pose issues in real-world applications.

Future research should aim to address these issues by extending the attribute set to include a broader range of attributes that could affect student academic performance. More so, optimizing polyfeature selection process could yield more tailored attributes that benefit a wider set of classifiers. The construction of hybrid models integrating numerous ML approaches may improve predictive performance while ensuring interpretability. Furthermore, applying real-time prediction and intervention systems in educational sectors will be a valuable expansion, ensuring timely support for student at-risk. Integrating cross-cultural datasets and checkmating algorithms evaluation performance among numerous educational scenarios will assist in maintaining generalizability of the proposed model. Lastly, future trend will focus on enhancing the computational efficiency and searching alternative explainability approaches to decrease computational overhead while ensuring interpretability.

Conflict of Interest

Authors declare no conflicts of interest

Data Availability

https://archive.ics.uci.edu/dataset/320/student+performance

References

[1] Ahmar, A. S., & Rahman, A. (2018). Development of prediction model of student academic performance. Journal of Physics: Conference Series_, 1028(1), 012162.

[2] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In _Advances in neural information processing systems_ (pp. 4765-4774).

[3] Fan, X., & Sun, W. (2017). Research on improved K-means clustering algorithm for college students. Journal of Physics: Conference Series_, 887(1), 012022.

[4] Gray, K. E., & Perkins, M. C. (2019). Early engagement data for predicting student performance and risk of failure. Computers & Education_, 136, 60-68.

[5] Francis, J., & Babu, M. S. (2019). A survey on predictive analytics in higher education to enhance student performance. _Journal of Physics: Conference Series_, 1228(1), 012011.

[6] Mason, C. M., McFarland, J. W., & Shorter, S. H. (2018). Predicting student attrition using neural networks and course grades. _Journal of College Student Retention: Research, Theory & Practice_, 20(2), 172-186.

[7] Fan, J., & Sun, M. (2021). Academic achievement prediction in higher education through interpretable modeling. PLOS ONE, 16(3), e0247590. https://doi.org/10.1371/journal.pone.0247590

[8] Serrano, R., & Casillas, J. (2024). Predicting Academic Success of College Students Using Machine Learning Techniques. Data Journal, 9(4), 60. https://doi.org/10.3390/data9040060

[9] Albreiki, I., Zaki, N., & Zubi, M. D. (2021). Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. Applied Sciences, 11(19), 9023. https://doi.org/10.3390/app11199023

[10] Albreiki, I. M., Zaki, N., & Zubi, M. D. (2021). Educational Data Mining Techniques and Applications: A Review. Journal of Educational Technology Systems, 50(1), 41-73. https://doi.org/10.1177/00472395211014205

[11] Coussement, K., & De Caigny, A. (2020). Predicting Student Dropouts in Online Learning Using a Logit Leaf Model: Decision Support for Student Retention. Decision Support Systems, 138, 113382. https://doi.org/10.1016/j.dss.2020.113382

[12] Gray, J., & Perkins, D. (2019). Predicting Academic Challenges Using Engagement Data and Machine Learning Models. Journal of Educational Data Mining, 11(3), 1-19. https://jedm.educationaldatamining.org/index.php/JEDM/article/view/392

[13] Alelyani, S., & Zubi, M. (2021). Machine Learning Approaches for Predicting Academic Performance: A Review. International Journal of Advanced Computer Science and Applications, 12(1), 23-29. https://doi.org/10.14569/IJACSA.2021.0120146