Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 1, January 2025, pp. 223~231
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i1.pp223-231

Predicting student performance using Moodle data and machine learning with feature importance

Jamal Kay Rogers[1,3], Tamara Cher Mercado[1], Remelyn Cheng[2,3]
[1]College of Information and Computing, University of Southeastern Philippines, Davao City, Philippines
[2]College of Arts and Sciences, Mapua Malayan Colleges Mindanao, Davao City, Philippines
[3]Graduate School Department, University of the Immaculate Conception, Davao City, Philippines

Article Info

Article history:

Received Jul 7, 2024
Revised Sep 3, 2024
Accepted Sep 7, 2024

Keywords:

Algorithm
Educational data mining
Feature engineering
Learning management system
Online learning
Predictive analytics

ABSTRACT

Despite the growing technological advancement in education, poor academic performance of students remains challenging for educational institutions worldwide. The study aimed to predict students' academic performance through modular object-oriented dynamic learning environment (Moodle) data and tree-based machine learning algorithms with feature importance. While previous studies aimed at increasing model performance, this study trained a model with multiple data sets and generic features for improved generalizability. Through a comparative analysis of random forest (RF), XGBoost, and C5.0 decision tree (DT) algorithms, the trained RF model emerged as the best model, achieving a good ROC-AUC score of 0.77 and 0.73 in training and testing sets, respectively. The feature importance aspect of the study identified the submission actions as the most crucial predictor of student performance while the delete actions as the least. The Moodle data used in the study was limited to 2-degree programs from the University of Southeastern Philippines (USeP). The 22 courses still resulted in a small sample size of 1,007. Future research should broaden its focus to increase generalizability. Overall, the findings highlight the potential of machine learning techniques to inform intervention strategies and enhance student support mechanisms in online education settings, contributing to the intersection of data science and education literature.

Corresponding Author:

Jamal Kay Roger
College of Information and Computing, University of Southeastern Philippines
Iñigo Street, Obrero, Davao City, Philippines
Email: jamalkay.rogers@usep.edu.ph

1. INTRODUCTION

Despite technological advancements in education, poor academic performance remains one of educational institutions' most challenging aspects [1]. Factors such as the lack of family support, financial issues, motivation, learning facilities, and teaching techniques contribute to this concern. Although the ever-evolving growth of technology provided great opportunities to improve academic performance, challenges remain due to these factors [2]. Many institutions have employed management techniques, web-based learning, and emerging technologies such as data analytics, the internet of things (IoT), and data mining techniques to address poor academic performance challenges. Machine learning algorithms and deep learning models have been used to analyze student data and predict student success [3], [4]. These student data can be accessed with the advent of online learning through learning management systems (LMS).

The growing popularity of online learning with LMS has paved the way for institutions to collect user data through databases and logs. Studies have shown that resources and activities from LMS, such as files, links, and forum participation, can positively impact academic performance [5]. Using LMS data provides institutions with data-driven approaches that can help identify students needing assistance and provide interventions, ultimately improving educational outcomes [6], [7].

LMS has different platforms like modular object-oriented dynamic learning environment (Moodle), Blackboard, Canvas, iSpring Learn, Sakai, and Google Classroom. The Moodle is one of the most widely used LMSs [8]. Moodle is user-friendly and engages students, improving academic performance [9]. Moodle offers course creation, content delivery, assessment, and interactive tools [10], [11]. Moodle's role in education is crucial because it can facilitate distance learning, increase student engagement, and enable personalized learning. Moodle has shown positive results in enhancing the teaching and learning process of institutions using it [12]. With Moodle, educational data mining (EDM) can be enhanced. Educators and administrators can use Moodle logs to extract data summaries and visualizations, generating relevant insights on how students interact with the LMS. EDM provides techniques for predicting students at risk of failing a course. Through EDM, Moodle data can be analyzed to identify low-performing students and the most influencing factors of academic performance, aiding in early interventions [13].

EDM uses machine learning techniques to predict student performance and provide insights for educators and administrators, giving them opportunities for early interventions [14]. The application of EDM supports informed decision-making and improves educational outcomes [15]. By analyzing Moodle logs and using various machine learning and other data analysis techniques, EDM improves the reliability of academic performance predictions. Machine learning algorithms, such as random forests (RF) and decision trees (DT), are commonly used in EDM to achieve reliable results. Additionally, EDM contributes to developing systems that understand student interactions and learning within online environments [16]. Overall, using machine learning techniques enhances the ability to predict student performance, offering opportunities for intervention. This strategy can benefit both students and the institution.

Multiple studies [17]-[24] have used EDM to predict student performance using Moodle logs. However, studies that used machine learning focused more on increasing model performance in the training and testing environments without considering generalizability. Other studies used only quantitative analysis. These previous studies used Moodle data from only a single to a few courses or courses with the same instructor without considering heterogeneity. This concern can negatively impact findings when applied to other settings. Furthermore, these studies differ in the features used and may not be Moodle generic; some may not apply to other settings. Overall, the previous studies were limited in terms of the training data, which can negatively impact generalizability [25].

Generalizability remains an issue in machine learning, where models perform poorly on new data [26]-[28]. Generalizability ensures model deploy ability to other Moodle courses. Therefore, this study aimed to predict students' performance using Moodle logs with machine learning using generalized features and multiple course data, enabling model deploy ability to other Moodle courses.

The study focused on developing a predictive model by comparing various machine learning algorithms trained on multiple Moodle course data with generic features, aiming to be applied to various courses. The study also aimed at identifying the factors (features) that contribute to academic performance. Institutions can use the study's outcome as a predictive model to inform academic-related decisions and future researchers for further development and potential applications. The conduct of the study contributes to the growing literature on academic performance prediction through EDM.

2. METHOD

Using data-wrangled Moodle logs, the study used supervised binary classification machine learning to predict student performance. Data collected from a Philippine university's Moodle logs were used to feature-engineer Moodle-generic predictors of student performance. The feature engineering focused on creating theoretically aligned features, ensuring applicability to other Moodle courses while aiming at generalizability. The Moodle data were used to train multiple machine learning algorithms and compare their performance. The best-performing algorithm was chosen for the final model. The final model was tested using a testing set from the data collected to determine its performance on unseen data. The data preprocessing aspect of the study was performed using the R programming language. The machine learning aspect was performed using the same language through the tidy model's package. The following subsections detail the steps and techniques used in the study.

2.1. Moodle data set and data preprocessing

The study's data collection and preprocessing aspect was based on the study of [29] as detailed in the data collection and feature engineering subsections of the Method section. The data was collected from the Moodle logs of 22 different courses coming from the Bachelor of Science in Information Technology

(BSIT) and the Bachelor of Science in Computer Science (BSCS) programs of the University of Southeastern Philippines (USeP), a state university in the Davao region of the Philippines. These courses were from 3 academic years: 2021-2022, 2022-2023, and 2023-2024. Multiple courses with different instructors were used to increase the representation of different types of courses. The data set after data wrangling is structured data containing rows and columns. The rows represent the count of student's logged actions based on column variables or features. These features are the Moodle event names from the Moodle logs.

Feature engineering was based on existing theories. The most used Moodle activities and resources (submissions, quizzes, and forums) were considered to ensure applicability to other Moodle courses. These represent necessary tools for communication, collaboration, and assessment [30]. The create, read, update, and delete actions (CRUD), which represent a standard log record (database query types) in information systems, were also considered. These features were engineered by adding the values of the appropriate Moodle event names logged by students. The variable days logged in represents a time-based feature that counts the number of days students have logged in to their course. The eight (8) variables are the generalized features of the study, ensuring applications across multiple Moodle platforms. These features are the predictors or independent variables of the study. The study engineered another feature, Performance, to label students' performance as high or low based on course grades. This feature is the response or dependent variable. A course grade of 86 and above denotes high performance; else, low performance. Table 1 shows the data set summary used in the study after data preprocessing.

Figure 1 shows the data distribution based on performance, where students labeled high are 615 while those labeled low are 392. This data is considered imbalanced data. Figure 2 shows the correlations between the numeric features of the data, some indicating high correlations. The highest Pearson R correlation coefficient is 0.95 between the submission actions and create actions. High correlation coefficients of 0.76 and 0.74 exist between update actions and create actions, and read actions and forum days logged in, respectively. Another high correlation coefficient of 0.69 exists between read actions and quiz actions. High correlations in the data set render some machine learning algorithms perform poorly, especially parametric algorithms for classification modeling [31].

Table 1. The summary of the data set used in the study

Variable (feature) name	Variable class	Variable type
Submission actions	Numeric (double)	Predictor (independent)
Quiz actions	Numeric (double)	Predictor (independent)
Forum actions	Numeric (double)	Predictor (independent)
Create actions	Numeric (double)	Predictor (independent)
Read actions	Numeric (double)	Predictor (independent)
Update actions	Numeric (double)	Predictor (independent)
Delete actions	Numeric (double)	Predictor (independent)
Days logged in	Numeric (double)	Predictor (independent)
Performance	Character (factor levels: high, low)	Response (dependent)
No. of rows (observations)	: 1007	
No. of columns (features)	: 9	

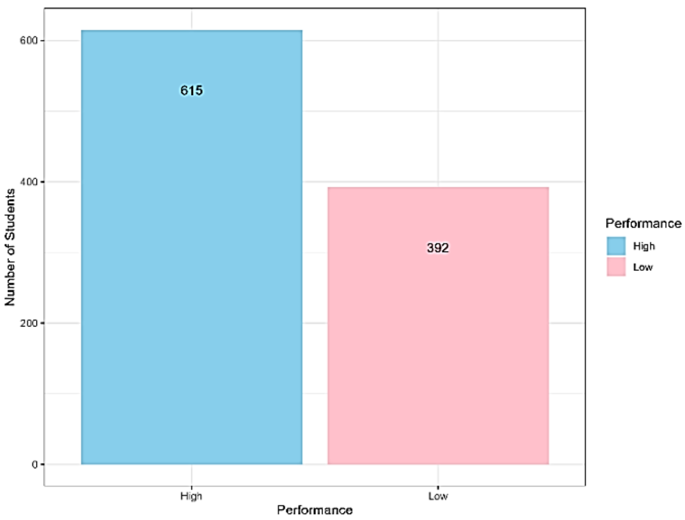

Figure 1. Performance distribution of the Moodle data based on performance

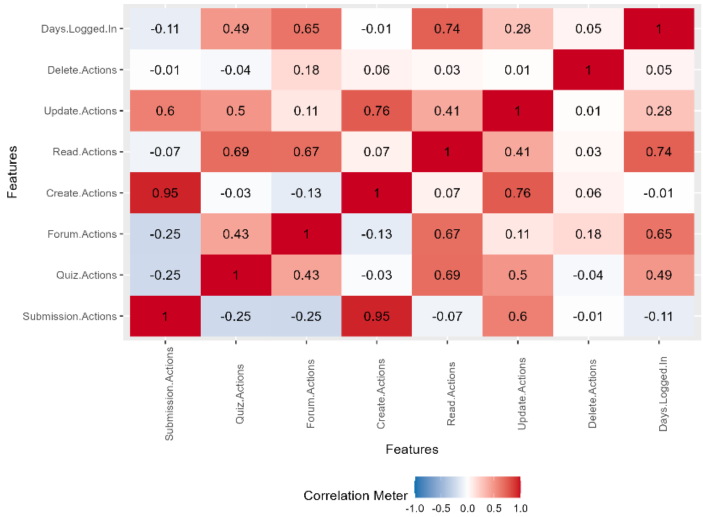

Figure 2. Correlation plot of the features of the data set showing Pearson r coefficients

2.2. Machine learning

Due to their non-parametric nature, multiple tree-based machine learning algorithms were used to train for predictive analytics. The modeling framework in Figure 3 was used as the guiding framework of the study. The data set was split into training and testing sets. Resampling was performed on the training set, and tree-based models were compared. The best-performing model was selected and used to fit the entire training data. The best model's performance on training was verified on the testing set.

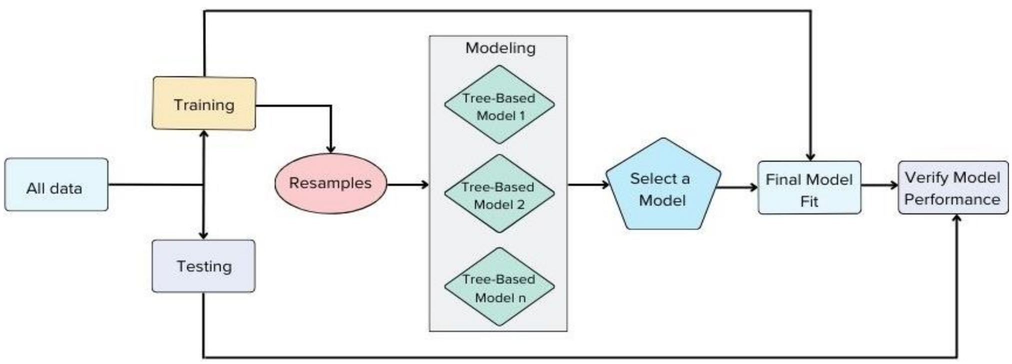

Figure 3. The modeling framework used in the study

3 metrics were used to measure model performance: ROC-AUC (receiver operating characteristics – area under the curve), sensitivity, and specificity. The ROC-AUC is a graphical representation of the trade-off between the true positive rate (TPR or sensitivity) on the x-axis and the false positive rate (FPR or 1-specificity) on the y-axis at various settings of the model. It quantifies a model's ability to distinguish between positive and negative cases across all possible thresholds. A perfect model will give a ROC-AUC score of 1, and a random classifier model gives a score of 0.5, a straight diagonal line. A score above 0.5 or close to 1 is considered a good score. Sensitivity or true positive rate (TPR) measures the proportion of actual positive cases classified correctly by the model. Specificity or true negative rate (TNR) measures the proportion of actual negative cases correctly classified by the model.

$$\text{ROC-AUC} = \int_0^1 TPR\ d(FPR) \tag{1}$$

$$\text{TPR} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{TNR} = \frac{TN}{TN + FP} \tag{3}$$

In the study, the high-performance label is the positive class, while the low-performance label is the negative class. True positives (TP) are the cases where a model correctly predicts the positive class. Moreover, false negatives (FN) are the cases where the model incorrectly predicts the negative class when the true class is positive. Furthermore, true negatives (TN) are the cases where the model correctly predicts the negative class. Lastly, false positives (FP) are the cases where the model incorrectly predicts the positive class when the true class is negative.

2.2.1. Data splitting and resampling

The whole data set was split into training and testing sets. 80% (805 observations) of the data went into the training set, and the remaining 20% (202 observations) was reserved for testing the final model. Stratification ensured a balanced representation of the outcome variables, reducing biases [32]. The training set contains 492 high-performing and 313 low-performing students, while the testing set contains 123 and 79, respectively.

Following the framework in Figure 2, the training set goes into a resampling process using 10-fold cross-validation. 10-fold cross-validation splits the training data into 10 folds, each acting as a testing set, while the remaining 9 are training sets. This type of validation ensures no data leakage during the training [33]. Resampling provides more reliable model performance, reducing the variance during training. In the resampling, stratification was also applied. Table 2 shows the resamples from the training set through 10-fold cross-validation.

Table 2. Resamples from the training set using 10-fold cross-validation

Fold	Training split	Testing split
Fold01	723	82
Fold02	723	82
Fold03	724	81
Fold04	725	80
Fold05	725	80
Fold06	725	80
Fold07	725	80
Fold08	725	80
Fold09	725	80
Fold10	725	80

2.2.2. Modeling and variable importance

The study compared 3 tree-based machine learning algorithms for training: XGBoost, RF, and C5.0 DT. These algorithms were chosen because of their explainability and interpretability [34] and ability to perform well with limited data [35], making them suitable for dealing with highly correlated features. In the comparison, the hyperparameters of the algorithms were tuned to get the best parameter combinations from each. Hyperparameter tuning generated multiple models based on each combination. The tune race ANOVA method from tidymodels was used during training to shorten training time by eliminating poor-performing models. A recipe preprocessor from the tidymodels package was set to normalize the training data and oversample the minority class (low) using the adaptive synthetic (ADASYN) sampling approach. ADASYN was used to improve the algorithms' performance on the imbalanced data set during training [36], [37].

The importance of the features of the modeling performed was determined with the help of the vip package in R. The vip package visualizes the strength of the relationship between each feature and the response predicted while considering all other features used in modeling. The package calculates variable importance (VI) scores using model-specific or model-agnostic approaches such as variance-based, permutation, and Shapley methods [38].

3. RESULTS AND DISCUSSION

3.1. Comparision and selection of best model

The study compared and ranked the tuned algorithms' performance on training based on 3 metrics: sensitivity, specificity, and ROC-AUC. This approach is also evident in the studies of [17], [20]-[23], wherein they compared multiple machine learning algorithms across multiple metrics. The modeling conducted in the study generated multiple models, and good-performing models were ranked as determined by the tune race ANOVA method. Based on the ROC-AUC score, the RF model performed best at 0.77 with a sensitivity of 0.73 and specificity of 0.66. The C5.0 DT comes in second place with an ROC-AUC score of 0.74 and sensitivity and specificity of 0.67 and 0.72, respectively. Last place is the XGBoost with ROC-ACU of 0.71, sensitivity of 0.70, and specificity of 0.57. Table 3 summarizes the results of these models based on the 3 metrics with preference to the ROC-AUC metric. The ROC-AUC was chosen as the appropriate reference metric because of the imbalanced nature of the data, and it is a popular metric for binary classification [39]. The RF model was chosen as the overall best model, complementing the studies of [20], [22], [23], but with different sets of metrics. The RF model has the highest value in distinguishing high- and low-performing students. The model is also best at predicting high-performing students while in second place at predicting low-performing students.

Table 3. Summary results of the best tree-based models across 3 metrics

Model	ROC-AUC	Sensitivity	Specificity
RF	0.77	0.73	0.66
C5.0 DT	0.74	0.67	0.72
XGBoost	0.71	0.70	0.57

3.2. Final model fit and verification of model performance

The chosen best model was fitted on the testing set, unseen data during data training, to validate the model's overall performance. The chosen model's performance on the testing set across all metrics is shown in Table 4. Its ROC-AUC score on the testing set is 0.73. The model classified high-performing students with a sensitivity score of 0.69. The model also classified low-performing students with a specificity score of 0.66.

Figure 4 shows the plot of the ROC-AUC. The plot indicates that the model has a moderate discriminatory ability, better than random guessing but not highly accurate. The model captures a decent portion of high-performing student cases. However, there is room for improvement in increasing the TPR. The model is also somewhat effective at avoiding false positives. However, there is also room for improvement in increasing the TNR. Increasing the model's performance may be achieved with increased data size [40].

Table 4. The chosen model's performance on the testing set

Model	ROC-AUC	Sensitivity	Specificity
RF	0.73	0.69	0.66

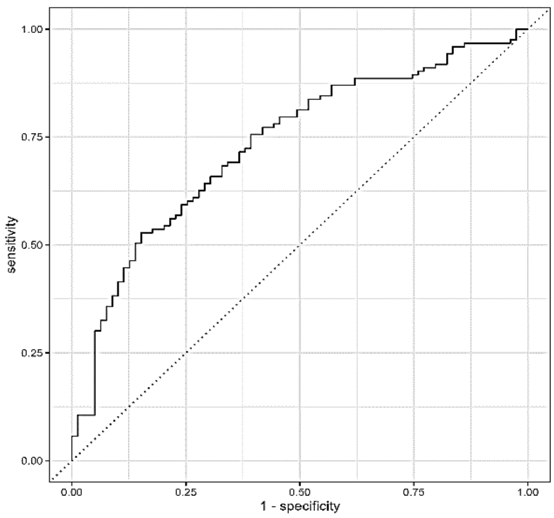

Figure 4. Plot of the ROC-AUC performance of the model across different threshold settings

3.3. Feature importance interpretation

The study of [41] showed different patterns of how students interact with Moodle based on content, activities, and assessment. Therefore, selecting the most important features in modeling can significantly help improve the model's performance [42]. The features were ranked from highest to lowest in importance, where the most crucial feature is listed first and the least important is listed last. Table 5 shows the VI scores of each feature. The submission actions have been identified as the most crucial feature, with a VI score of 72.73. This suggests that actions related to submissions are the strongest predictor of academic performance. This finding complements the studies of [21], [22], referring to submitting assignments within Moodle as the best or one of the best predictors of academic performance.

The read actions, days logged in, and create actions are close in second, third, and fourth places, respectively. With a VI score of 57.53, read actions signifies that reading or accessing content is crucial to academic performance. The days logged in, having a VI score of 55.80, suggests that consistent engagement with the Moodle LMS is notable in predicting academic performance, consistent with the studies of [21], [23]. The create actions implies that activities related to creating content within the Moodle LMS also strongly influence academic performance, with a VI score of 55.10.

Quiz actions comes in at fifth place with a VI score of 48.13, implying that interactions within quizzes or assessments within Moodle can still influence academic performance. The forum actions, interactions within forums and discussions can also influence academic performance with a VI score of 45.5 and ranked sixth. Ranked in seventh place is the update actions, indicating that student updates within the LMS have a relatively lower impact on predicting academic performance, with a VI score of 37.39. While still relevant, update actions' impact on academic performance appears to be less significant compared to the features that are higher ranked. Delete actions are the least important, with a VI score of 4,22. This suggests that removing content from the LMS has the weakest influence on academic performance.

Table 5. Variable importance score of each feature

Feature (predictor)	VI score
Submission Actions	72.73
Read Actions	57.53
Days Logged In	55.80
Create Actions	55.10
Quiz Actions	48.13
Forum Actions	45.53
Update Actions	37.39
Delete Actions	4.22

4. CONCLUSION

The study predicted student performance using Moodle data from various courses and tree-based machine learning algorithms. In comparing 3 tree-based machine learning algorithms, namely, RF, XGBoost, and C5.0 DT, the RF yielded the highest ROC-AUC score of 0.77 with a sensitivity score of 0.73 and a specificity score of 0.66 in the training set. In the testing set, the RF got 0.73 ROC-AUC with sensitivity and specificity scores of 0.69 and 0.66, respectively. These values indicate that RF can discriminate between high-performing and low-performing students reasonably well but with room for improvement.

The feature importance aspect of the study feature engineered 9 Moodle-generic variables from the LMS data collected from USeP: submission actions, quiz actions, forum actions, create actions, read actions, update actions, delete actions, days logged in, and performance as the label feature (response). These features are aimed at generalizability and can be applied to any Moodle course. The submission actions are the most crucial feature in predicting student performance among the predictor features, while the delete actions are the least important. By highlighting the importance of features in predicting outcomes, the study laid the groundwork for future investigations to refine predictive models and inform targeted interventions tailored to individual student needs. The study was limited to just 2-degree programs, BSIT and BSCS, and 22 courses, resulting in a relatively small sample size of 1,007 observations. Nevertheless, the study's data size is more significant in number compared to previous studies. However, the findings show that there is much room for improvement. Future research could broaden its scope to encompass additional programs and other features for improved generalizability. Future research could also extend the study to other universities and consider other LMS platforms outside Moodle. Future studies can consider model deployment, allowing educators to leverage the model through information systems and dashboards.

Overall, the study's findings underscore the potential of machine learning techniques in enhancing educational intervention strategies, which can pave the way for effective support mechanisms for students. Educators can foster a more inclusive learning environment by implementing tailored support mechanisms. The study contributes to the growing body of literature on the intersection of data science and education, offering pathways for improving student academic performance in online learning environments.

ACKNOWLEDGMENTS

The authors thank the University of Southeastern Philippines (USeP) for allowing us to access and use their Moodle data, which was instrumental to this study. The collaboration with USeP enabled us to delve into valuable insights and further understand the research subject. The support from USeP significantly contributed to the success of this endeavor.

REFERENCES

[1] R. A. N. Al-Tameemi *et al.*, "Determinants of poor academic performance among undergraduate students—A systematic literature review," *International Journal of Educational Research Open*, vol. 4, Jan. 2023, doi: 10.1016/j.ijedro.2023.100232.

[2] M. S. Benoza and T. Palaoag, "Exploring college academic performance of K to12 IT and Non-IT-related strands to reduce academic deficiencies," *International Journal of Advanced Computer Science and Applications,* 2023 [Online]. Available: www.ijacsa.thesai.org

[3] K. Kanatti, J. Jayashankar, P. Gayathri, and A. Kumar, "Student's performance and difficulties prediction," *International Research Journal on Advanced Science Hub*, vol. 5, no. Issue 05S, pp. 185–189, May 2023, doi: 10.47392/irjash.2023.s024.

[4] A. Bodkhe, "Machine learning based result analysis prediction of student's academic performance," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, no. 5, pp. 5058–5062, May 2023, doi: 10.22214/ijraset.2023.52851.

[5] L. Aguagallo, F. Salazar-Fierro, J. García-Santillán, M. Posso-Yépez, P. Landeta-López, and I. García-Santillán, "Analysis of student performance applying data mining techniques in a virtual learning environment," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 11, pp. 175–195, 2023, doi: 10.3991/ijet.v18i11.37309.

[6] R. A. N. Al-Tameemi *et al.*, "Determinants of poor academic performance among undergraduate students—A systematic literature review," *International Journal of Educational Research Open*, vol. 4, Jan. 2023, doi: 10.1016/j.ijedro.2023.100232.

[7] W. Gordillo-Tenorio, C. Meléndez-Flores, F. Sierra-Liñan, and M. Cabanillas-Carbonell, "Information technologies that help improve academic performance, a review of the literature," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 4, pp. 262–279, 2023, doi: 10.3991/ijet.v18i04.34821.

[8] J. Bojiah, "Effectiveness of moodle in teaching and learning," *Journal of Hunan University Natural Sciences*, vol. 49, no. 12, pp. 320–328, Dec. 2022, doi: 10.55463/issn.1674-2974.49.12.33.

[9] S. Qin, L. Orchakova, Z. Y. Liu, Y. Smirnova, and E. Tokareva, "Using the learning management system 'modular object-oriented dynamic learning environment' in multilingual education," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 3, pp. 173–191, 2022, doi: 10.3991/ijet.v17i03.25851.

[10] K. Sevnarayan, "Moodle as a tool to reduce transactional distance at an open distance and e-learning University," *Studies in Learning and Teaching,* vol. 4, no. 4, 2023, doi: 10.46627/silet.

[11] A. Nasrum, M. Subawo, and U. Hidayati, "The role of moodle in education: a bibliometric review," *JME (Journal of Mathematics Education)*, vol. 8, no. 1, Jun. 2023, doi: 10.31327/jme.v8i1.1903.

[12] J. Bojiah, "Effectiveness of moodle in teaching and learning," *Journal of Hunan University Natural Sciences*, vol. 49, no. 12, pp. 320–328, Dec. 2022, doi: 10.55463/issn.1674-2974.49.12.33.

[13] A. Akhtar, M. Serajuddin, and H. Zafrul, "Introduction of educational data mining by using a variety of techniques in order to achieve the goal from the moodle LMS," *European Journal of Open Education and E-learning Studies*, vol. 6, no. 1, Jun. 2021.

[14] S. Begum and S. S. Padmannavar, "Student performance analysis using Bayesian optimized random forest classifier and KNN," *International Journal of Engineering Trends and Technology*, vol. 71, no. 5, pp. 132–140, May 2023, doi: 10.14445/22315381/IJETT-V71I5P213.

[15] S. A. A. Kharis *et al.*, "Design of student success prediction application in online learning using fuzzy-KNN," *Barekeng: Jurnal Ilmu Matematika dan Terapan,* vol. 17, no. 2, pp. 0969–0978, Jun. 2023, doi: 10.30598/barekengvol17iss2pp0969-0978.

[16] S. K. Gupta, I. Kumar, M. Kumar, S. Kumari, and S. Devi, "Student performance analysis using machine learning," *International Journal of Progressive Research in Engineering Management and Science*, May 2023, doi: 10.58257/ijprems31181.

[17] S. Shrestha and M. Pokharel, "Educational data mining in moodle data," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 10, no. 1, p. 9, Apr. 2021, doi: 10.11591/ijict.v10i1.pp9-18.

[18] A. T. D. Nguyen, "Using machine learning to predict the low grade risk for students based on log file in moodle learning management system," *International Journal of Computing and Digital Systems*, vol. 11, no. 1, pp. 1133–1140, 2022, doi: 10.12785/ijcds/110191.

[19] H. Sokout and T. Usagawa, "Using access log data to predict failure-prone students in Moodle using a small dataset," *SHS Web of Conferences*, vol. 102, p. 04001, 2021, doi: 10.1051/shsconf/202110204001.

[20] M. M. Tamada, R. Giusti, and J. F. de M. Netto, "Predicting students at risk of dropout in technical course using LMS logs," *Electronics (Switzerland)*, vol. 11, no. 3, Feb. 2022, doi: 10.3390/electronics11030468.

[21] C. Kaensar and W. Wongnin, "Analysis and prediction of student performance based on moodle log data using machine learning techniques," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 10, pp. 184–203, 2023, doi: 10.3991/ijet.v18i10.35841.

[22] R. J. Quinn, G. Gray, R. Quinn, and G. Gray, "Prediction of student academic performance using Moodle data from a further education setting," *Irish Journal of Technology Enhanced Learning,* 2020, [Online]. Available: http://www.ilta.ie/

[23] S. Gaftandzhieva *et al.*, "Exploring online activities to predict the final grade of student," *Mathematics,* vol. 10, no. 20, Oct. 2022, doi: 10.3390/math10203758.

[24] Y. Zhang, A. Ghandour, and V. Shestak, "Using learning analytics to predict students performance in moodle LMS," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 20, pp. 102–114, 2020, doi: 10.3991/ijet.v15i20.15915.

[25] J. Y. Bo, H.-W. Huang, A. Chan, and G. Traverso, "Pretraining ECG data with adversarial masking improves model generalizability for data-scarce tasks," *arXiv preprint arXiv:2211.07889*, Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.07889.

[26] K. Sytwu, L. R. DaCosta, and M. C. Scott, "Generalization across experimental parameters in machine learning analysis of high-resolution transmission electron microscopy datasets," *arXiv preprint arXiv:2306.11853,* Jun. 2023, [Online]. Available: http://arxiv.org/abs/2306.11853

[27] A. Nikolikj *et al.*, "Assessing the generalizability of a performance predictive model," *In Proceedings of the Companion Conference on Genetic and Evolutionary Computation*, May 2023, [Online]. Available: http://arxiv.org/abs/2306.00040

[28] C. Zang *et al.*, "Accuracy and generalizability of machine learning models for adolescent suicide prediction with longitudinal clinical records," *Translational psychiatry*, 2023, doi: 10.21203/rs.3.rs-2790491/v1.

[29] J. K. B. Rogers, T. C. R. Mercado, and R. S. Decano, "Moodle interactions and academic performance: educational data mining in a Philippine university," *Journal of Education and Learning (EduLearn)*, vol. 19, no. 1, pp. 542–550, Feb. 2025, doi: 10.11591/edulearn.v19i1.21549.

[30] N. H. S. Simanullang and J. Rajagukguk, "Learning management system (LMS) based on moodle to improve students learning activity," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Mar. 2020, doi: 10.1088/1742-6596/1462/1/012067.

[31] S. Mishra and R. K. Pradhan, "Analyzing the impact of feature correlation on classification acuracy of machine learning model," *2023 International Conference on Artificial Intelligence and Smart Communication, AISC 2023*, pp. 879–883, 2023, doi: 10.1109/AISC56616.2023.10085542.

[32] A. Linden, "A comparison of approaches for stratifying on the propensity score to reduce bias," *Journal of Evaluation in Clinical Practice*, vol. 23, no. 4, pp. 690–696, Aug. 2017, doi: 10.1111/jep.12701.

[33] B. Weck and X. Serra, "Data leakage in cross-modal retrieval training: a case study," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2023-June, 2023, doi: 10.1109/ICASSP49357.2023.10094617.

[34] K. Filom, A. Miroshnikov, K. Kotsiopoulos, and A. R. Kannan, "On marginal feature attributions of tree-based models," *arXiv preprint arXiv:2302.08434*, Feb. 2023, [Online]. Available: http://arxiv.org/abs/2302.08434

[35] G. Pedretti, S. Serebryakov, J. P. Strachan, and C. E. Graves, "A general tree-based machine learning accelerator with memristive analog CAM," *Proceedings - IEEE International Symposium on Circuits and Systems*, vol. 2022-May, pp. 220–224, 2022, doi: 10.1109/ISCAS48785.2022.9937772.

[36] S. Karthikeyan and T. Kathirvalavakumar, "Genetic algorithm-based over-sampling with DNN in classifying the imbalanced data distribution problem," *Indian Journal of Science and Technology*, vol. 16, no. 8, pp. 547–556, Feb. 2023, doi: 10.17485/IJST/v16i8.863.

[37] A. M. Halim, M. Dwifebri, and F. Nhita, "Handling imbalanced data sets using SMOTE and ADASYN to improve classification performance of ecoli data sets," *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, Jun. 2023, doi: 10.47065/bits.v5i1.3647.

[38] D. Yu and H. Wu, "Variable importance evaluation with personalized odds ratio for machine learning model interpretability with applications to electronic health records-based mortality prediction," *Statistical Medicine*, vol. 42, no. 6, pp. 761–780, Mar. 2023, doi: 10.1002/SIM.9642.

[39] L. Lavazza, S. Morasca, and G. Rotoloni, "On the reliability of the area under the ROC curve in empirical software engineering," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jun. 2023, pp. 93–100, doi: 10.1145/3593434.3593456.

[40] M. Jebbari, B. Cherradi, S. Hamida, M. A. Ouassil, T. E. Harrouti, and A. Raihani, "Enhancing learner performance prediction on online platforms using machine learning algorithms," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 35, no. 1, pp. 343–353, Jul. 2024, doi: 10.11591/ijeecs.v35.i1.pp343-353.

[41] I. R. Al-Kindi, Z. Al-Khanjari, and Y. Jamoussi, "Extracting student patterns from log file Moodle course: A case study," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 11, no. 2, pp. 917–926, Jun. 2022, doi: 10.11591/ijere.v11i2.23242.

[42] S. Hanumanthappa and C. Prakash, "Machine learning based education data mining through student session streams," *International Journal of Reconfigurable and Embedded Systems (IJRES)*, vol. 13, no. 2, pp. 383–394, Jul. 2024, doi: 10.11591/ijres.v13.i2.pp383-394.

BIOGRAPHY OF AUTHORS

Jamal Kay Rogers is an associate professor at the College of Information and Computing (CIC) of the University of Southeastern Philippines (USeP) in Davao City, Philippines. Engr. Rogers is registered electronics engineer with a Bachelor's in Electronics and Communication Engineering and a Master's in Engineering, majoring in Electronics and Communication Engineering. He is taking a Doctor of Philosophy in Education major in Information Technology Integration. His research interests are the IoT, data science, machine learning, and educational data mining (EDM). He can be contacted at email: jamalkay.rogers@usep.edu.ph.

Tamara Cher Mercado is a professor at the College of Information and Computing of the University of Southeastern Philippines (USeP) in Davao City, Philippines. She is the Vice President for Planning and Quality Assurance (VPPQuA). Dr. Mercado holds a Doctor of Philosophy in Information Technology, a Master of Science in Information Science, and a Bachelor of Science in Computer Engineering. Her research interests include information systems, knowledge management, and offline web services. She can be contacted at email: tammy@usep.edu.ph.

Remelyn Cheng is an associate professor at the College of Arts and Sciences of the Mapua Malayan Colleges Mindanao (MMCM) in Davao City, Philippines. Dr. Cheng holds a Doctor of Philosophy in Mathematical Sciences, a Master of Arts in Education in Teaching Mathematics, and a Bachelor of Science in Mathematics. Her research interests include phenomenon-based learning and all things mathematics. She can be contacted at email: remyasahid@gmail.com.