INNOVATIVE: Journal Of Social Science Research
Volume 4 Nomor 1 Tahun 2024 Page 3780-3792
E-ISSN 2807-4238 and P-ISSN 2807-4246
Website: https://j-innovative.org/index.php/Innovative

Predicting Science Learning Outcomes Of Elementary Students In Rural Area Using Machine Learning Algorithms

Edward Harefa[1✉]
Department of Elementary Education, Universitas Nias
Email: edwardharefa77@gmail.com[1✉]

Abstrak

Machine learning (ML) telah membuka jalan bagi penelitian yang berfokus pada peningkatan pengalaman belajar peserta didik dan membantu mengatasi tantangan yang dihadapi oleh sistem pendidikan. Teknologi ML menganalisis data untuk mengenali pola dan menggunakannya untuk membuat prediksi. Penelitian ini memperkenalkan model ML yang mengklasifikasikan dan memprediksi keberhasilan akademik peserta didik dengan memanfaatkan algoritma ML seperti *random forest*, *support vector machines, gradient boosting*, *decision tree*, *logistic regression*, *Extreme gradient boosting* (XGBoost), dan *deep learning*. Penelitian ini bertujuan untuk memprediksi hasil belajar sains peserta didik berdasarkan data historis dan mengidentifikasi faktor-faktor kunci yang mempengaruhi keberhasilan akademik peserta didik. Dengan demikian, pendekatan yang diusulkan menawarkan solusi untuk memprediksi kinerja akademik peserta didik secara efisien dan akurat dengan membandingkan beberapa model ML dengan model Deep Learning. Hasil penelitian menunjukkan bahwa *Extreme Gradient Boosting (XGBoost)* dapat memprediksi kinerja akademik peserta didik dengan akurasi 97,12%. Selain itu, hasil penelitian ini menunjukkan adanya fitur sosial dan demografis yang signifikan yang mempengaruhi keberhasilan akademik peserta ini. Studi ini menyimpulkan bahwa penerapan algoritma ML di kelas akan membantu para pendidik mengidentifikasi kesenjangan dalam pembelajaran siswa dan memungkinkan deteksi dini terhadap peserta didik yang berkinerja buruk sehingga memberdayakan para pendidik dalam pengambilan keputusan.

Kata Kunci: *Hasil belajar sains, machine learning, sekolah dasar*

Abstract

Machine learning has paved the way for research focused on improving learners' learning experiences and helping to address challenges faced by the education system. ML technologies analyze data to recognize patterns and use them to make predictions. This research introduces ML models that classify and predict learners' academic success by utilizing ML algorithms such as random forest, support vector machines, gradient boosting, decision tree, logistic regression, Extreme gradient boosting (XGBoost), and deep learning. This research aims to predict students' academic success based on historical data and identify key factors that affect students' academic success. Thus, the proposed approach offers a solution to predict learners' academic performance efficiently and accurately by comparing several ML models with Deep Learning models. The results show that Extreme Gradient Boosting (XGBoost) can predict students' academic performance with 97.12% accuracy. In addition, the results of this study indicate the presence of significant social and demographic features that affect the academic success of these participants. The study concludes that the application of ML algorithms in the classroom will help educators identify gaps in student learning and enable early detection of underperforming students, thus empowering educators in decision-making.

Keyword: *Elementary school, machine learning, science learning outcomes*

INTRODUCTION

Learning analytics is the measurement, collection, analysis and reporting of data about learners and their context for the purpose of understanding and optimizing learning and the environment in which it occurs (Mukred et al., 2024; Paolucci et al., 2024). Analyzing this data provides insight into the learning process and learner achievement. Further analysis can identify academic, demographic and social factors that influence primary school learners' academic success. Learners' academic success can be measured by assessing final results across subjects. Teachers measure learners' academic performance in primary schools from summative to formative assessments. According to a report from the United States of America Department of Education and the National Assessment of Educational Progress (NAEP), the education system in the is experiencing several challenges such as low academic achievement of learners, increasing dropout rates, graduation delays, and lack of preparedness of learners for global competition. Over the years, the academic success of learners has continued to decline, even more so among minority learners (Rimadana et al., 2019). Furthermore Johnson et al. (2021) revealed that compared to other regions, learners living in disadvantaged areas have unique challenges that can hinder academic achievement and growth, for example, often traveling long distances to get to school, causing fatigue and reduced learning time and inadequate family economic factors. Elementary school education is one of the Government of Indonesia's compulsory education programs (Harefa,

2023) but in practice there are still obstacles in its implementation, especially in frontier, rural and disadvantaged areas. Rural area is a regions characterized by a lack of complete facilities and infrastructure from infrastructure, both public and health infrastructure (Rosmana et al., 2022). Based on Presidential Regulation Number 30 of 2020 concerning the Determination of Underdeveloped Regions for 2020-2024, Nias Regency is one of the underdeveloped regions. Nias Regency is one of the regencies in North Sumatra located on Nias Island. The main occupations of the people of Nias Regency are fishermen and farmers, with a poverty rate of 15.69% (22,100 people), based on 2020 Statistics Indonesia data. There are 163 primary schools in Nias Regency spread across 10 sub-districts (Kemdikburistek, 2023). The average education system is still minimal, the infrastructure is just, the lack of adequate educators and the quality of teachers is still average, which can lead to low academic achievement of primary school students.

Advances in educational technologies such as artificial intelligence, virtual reality, 3D printing, smart multimedia devices, internet of things, and machine learning are beginning to improve the process and management of learner learning (Harefa et al., 2019; Harefa & Zhou, 2021; Yu et al., 2021). Machine learning (ML) analyzes data to recognize patterns and uses those patterns to make predictions. Applying ML in the classroom will enable educators to identify important factors that influence the success of learners' academic performance. In addition, ML will enable educators to identify underperforming learners, thus helping educators in decision-making (Costa-Mendes et al., 2021; Sandra et al., 2021; Wang et al., 2022). Several tools such as software like RStudio, Python Scikit-learn, and TensorFlow, are available to predict learners' academic performance. Xu et al. (2019) applied decision tree, neural network, and support vector machine (SVM) classification ML algorithms to predict academic performance from learners' internet usage behavior. Their results showed that learners' internet usage behavior effectively predicted academic performance with 71%-76% accuracy; however, the authors only considered accuracy as a performance metric. In their research, Hasan et al. (2019) proposed a system that uses ML algorithms trained to predict students' academic performance by classifying them as poor or good. The model was trained on data collected from university sources and implemented using k-nearest neighbor and decision tree classifiers. The results show that the decision tree classifier has an accuracy of 94.44%, but the authors only consider accuracy as its performance metric.

Bhutto et al. (2020) proposed a classification ML model using SVM classifier and logistic regression to predict learners' academic performance. The model extracts features from a pre-processed dataset obtained from an online education platform to classify learners' academic performance in poor, average, or good categories. The results show that SVM yields an accuracy of 79%, which is higher than logistic regression. The authors considered accuracy, recall, precision, and F1-score using confusion matrix to evaluate the system performance. Jayaprakash et al. (2020) used naïve bayes, random forest and ensemble learners to predict learners' academic performance using a dataset consisting of 887 examples of 19 attributes of first-year learners. The random forest classifier outperformed the other models with 93% accuracy. The recall, precision, and F1-score evaluation metrics used confusion matrix to evaluate model performance. Research on ML in education is still in its early stages, there are still many challenges such as prediction accuracy, overfitting, underfitting, model spread that need to be considered.

Science literacy is a scientific knowledge and skills that can be used to make decisions based on facts, research, and scientific phenomena (Chen et al., 2020). Based on the results of the Program for International Student Assessment (PISA), the science literacy score of Indonesian students in PISA 2018 was 396 and decreased to 383 in PISA 2022 (OECD, 2023). The decline in the science literacy of Indonesian students according to the PISA score signals the worst possibility if this condition occurs is not taken seriously. The goal of improving this literacy can be achieved if science learning involves students directly through interaction with the environment. Science in elementary school is the basis for children to be able to accept science and technology concepts at the next higher level. Furthermore, all the science knowledge they acquire will become the basis of knowledge in modernizing themselves, and living in the age of sophisticated technology. Based on the results of the researcher's observations, primary schools in the district have not anticipated and detected students who have the potential to experience obstacles or underachievement in their learning. The impact of the lack of anticipation or early prevention of students or potentially experiencing obstacles in learning or low mastery of science materials

In this study, we propose an efficient and accurate prediction analysis of primary school students' science learning outcomes by comparing several ML models with deep learning models. In general, deep learning models have better accuracy because they extract features from the dataset gradually. ML algorithms are applied to the dataset to analyze and identify features that significantly affect students' academic performance. This research is concerned with predicting the factors that influence the academic achievement of primary school students in 3T areas using demographic, social, and academic data. This research aims to see the comparison between the sophistication of ML in processing the input data. For this purpose, the classification algorithms with the highest performance in predicting learners' academic achievement are determined using ML classification algorithms such as random forest, support vector machine, stochastic gradient descent, decision tree, adaptive boosting, logistic regression, and deep learning. Finally, by utilizing these features, several ML models were trained to classify and predict academic performance categories of academic achievement and researchers also compared the performance of the models based on accuracy scores and cross-validation scores. Furthermore, research regarding the application of ML to the prediction of elementary school learners' learning for Nias Regency has not yet been researched so this research opens new insights in the wider application of technology in the field of education.

RESEARCH METHOD

The study used a quantitative research approach. Quantitative research is a systematic investigation of phenomena by collecting quantifiable data and performing statistical, mathematical, or computational techniques (Creswell, 2017). Quantitative research collects information from existing and potential customers by using sampling methods and sending online surveys, online polls, and questionnaires. Population is the number of individuals or objects that have the same characteristics. The sample is part of the elements that exist in the population. Thomas, (2021) states that the sample is part or part of the whole population. The target population of this study were elementary students in 8 schools in Nias Regency. The sampling method used in this study was purposive sampling. Purposive sampling is one type of non-probability sampling where the researcher determines the characteristics based on the research objectives to be achieved to answer research problems so that the sample used in this study was 1044 elementary students. This dataset consists of reports of learners' science learning outcomes collected through questionnaires and interviews. The data attributes include demographic, social, and science academic achievement features. Table 1 shows a summary of the attributes of this research dataset.

Table 1. Dataset attributes

Feature category	Name of the attributes	Description	Attribute type
Demographical features	School	Student's school	Categorical
	Sex	Student's sex	Categorical
	Age	Student's age	Numeric
	Address	Student's home address type	Categorical

	Famsize	Family size	Categorical
	Pstatus	Parent's cohabitation status	Categorical
	Medu	Mother's education	Numeric
	Fedu	Fedu - father's education	Numeric
	Mjob	Mother's job	Categorical
	Fjob	Father's job	Categorical
	Reason	Reason to choose this school	Categorical
	Guardian	Guardian - student's guardian	Categorical
Social features	Internet	Internet access at home	Categorical
	Romantic	With a romantic relationship	Categorical
	Famrel	Quality of family relationships	Numeric
	Freetime	Free time after school	Numeric
	Goout	Going out with friends	Numeric
	Dalc	Workday alcohol consumption	Numeric
	Walc	Weekend alcohol consumption	Numeric
	Health	Current health status	Numeric
Academic related features	Absences	Number of school absences	Numeric
	Traveltime	Home to school travel time	Numeric
	Studytime	Weekly study time	Numeric
	Failures	Number of past class failures	Numeric
	Schoolsup	Extra educational support	Categorical
	Famsup	Family educational support	Categorical
	Paid	Number of past class failures	Numeric
	Activities	Extra-curricular activities	Categorical
	Nursery	Attended nursery school	Categorical
	Higher	Wants to take higher education	Categorical
	Final grade	Final grade	Numeric

Problem solving with ML is categorized into supervised and unsupervised learning. Unsupervised ML works with unstructured data, whereas supervised ML works with structured datasets where input variables are mapped to output variables. Supervised ML problems are grouped into regression and classification problems. Regression problems involve predicting continuous discrete values, for example, predicting students' final grades. ML classification refers to the process of predicting a category from input data points. The category output can be a binary classification - “fail” or “pass” or a multiclass classification - “excellent, good, satisfactory, poor, and fail”. Classification ML is supervised ML where the input data is labeled and mapped with the output data; the ML model is trained to predict the output from the input. Implementing an ML classifier requires importing the required ML module package, then loading the dataset (Suhaimi et al., 2019). Preprocessing and data cleaning are performed on the dataset to check for null values, duplicates, invalid values, and encoding non-numeric and categorical data attribute types.

After successful data pre-processing, feature engineering techniques explored the dataset to understand the correlation relationships between variables to identify features that significantly impact the output variable. This allowed us to improve the accuracy of the model by removing attributes that significantly impact the output variable (students' final grades) but were not important features in predicting students' academic performance. The refined dataset was then divided into training and testing sets.The training dataset trains the model, and the testing dataset measures the performance of the model based on accuracy and cross-validation. Figure 1 shows the flow chart of the ML model of this research. This research builds and trains ML classification algorithms such as random forest, support vector machine, stochastic gradient descent, decision tree, adaptive boosting, logistic regression, and deep learning. Deep learning is a technique that uses neural network concepts to build and train ML models. Deep learning consists of input layer (receiving input data), hidden layer (extracting important features gradually), and output layer (Lye et al., 2010). Deep learning consists of a convolutional neural network (CNN) model with four hidden layers in accordance with the research objectives.

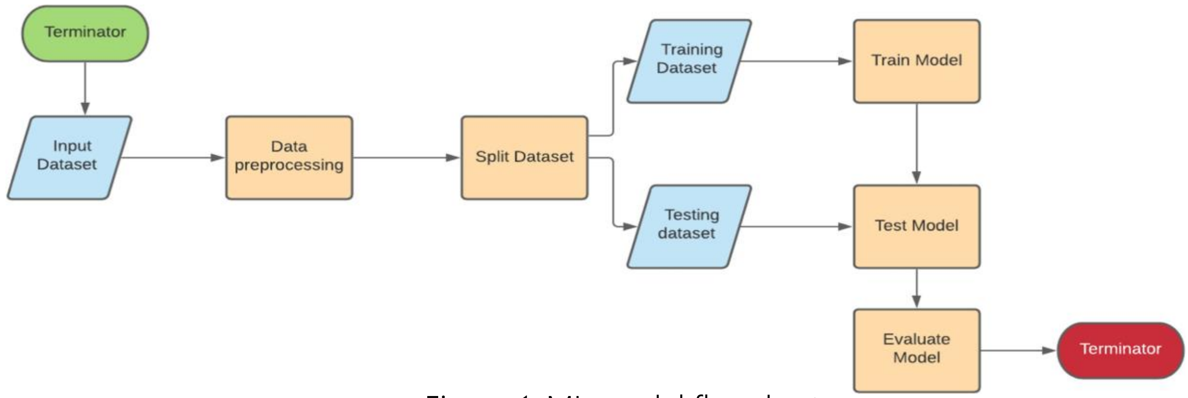

Figure 1. ML model flowchart

ML used test datasets to measure model performance. Accuracy, cross-validation, precision, recall, F1-score, confusion matrix, log loss, receiver operating characteristic (ROC), and area under curve (AUC) are some of the performance metrics used to evaluate ML classification models (Harefa et al., 2022; Vijayalakshmi & Venkatachalapathy, 2019). This study uses accuracy and cross-validation as performance metrics to evaluate ML classification models. The performance of the CNN model is evaluated using confusion matrix to calculate the accuracy, precision, and sensitivity of the model. Accuracy is the total number of correct predictions out of the total number of predictions. Cross-validation assesses how effectively the model will perform on new datasets. Confusion matrix is an error matrix that visualizes the performance of the ML model. Confusion matrix is used to calculate the accuracy, precision, and sensitivity of the model. Precision is the ratio of correctly predicted values to the total predicted values. Sensitivity evaluates the proportion of correct predictions of correct predictions from the model (Harefa & Zhou, 2022; Rimadana et al., 2019).

RESULTS AND DISCUSSION

The “plot_importance” function in the Scikit-learn library helps in plotting the important features that affect students' final grades. In predicting students' academic performance, the order of importance of the features and their scores can be seen in Figure 2. The number of school absences has the highest importance score. This indicates that learners who frequently miss school tend to have poor academic performance. Current health status, going out with friends, leisure time after school, quality of family relationships are the main social features that affect students' academic performance. Mother's occupation, father's occupation, status of living with parents, type of student's home address, and reasons for choosing this school are the least features that affect academic performance.

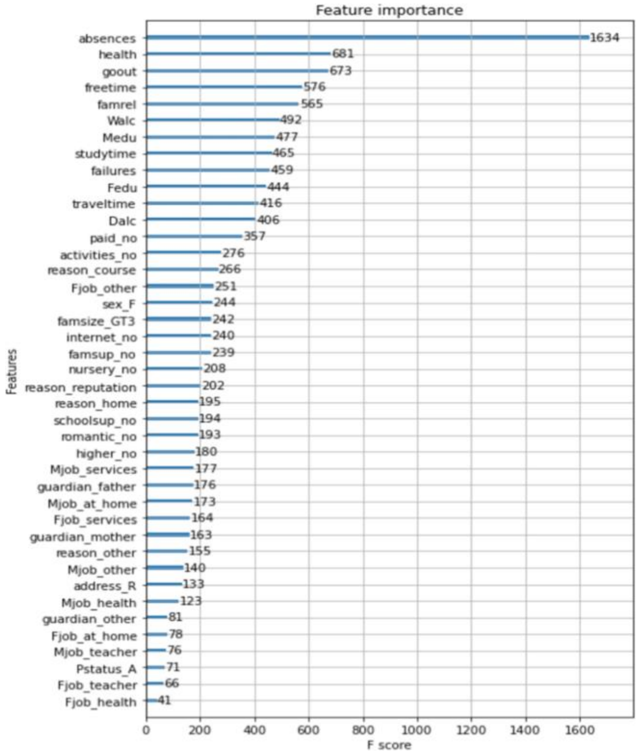

Figure 2. Important features and its score.

Table 2. Comparison of machine learning models.

ML classifier	Accuracy (%)	Cross validation (%)
Decision Tree Model	47.95	30.89
Random Forest Model	92.60	35.66
Support Vector Classifier Model	42.88	34.39
Logistic Regression Model	40.96	36.62
Ada Boost Model	35.75	32.48
Stochastic Gradient Descent	33.69	33.121
XGBoost Model	97.12	35.67
Deep Learning (CNN)	72.22	Precision = 30.31 Sensitivity = 31.38

To get an accurate evaluation of our model, the dataset containing 1044 students was divided into training and test datasets at a ratio of 70% to 30% using the 'train_test_split' function in Scikit learn. After building and training the ML model, the cross-validation function 'cross_val_score' helped calculate the average accuracy of the model on the test dataset. The cross validation function splits the test dataset into smaller subsets. The subsets are then fed into the model and the accuracy score is calculated five times with different subsets each time. After applying various classification models to the dataset, different accuracy and cross-validation scores were obtained for each model. Table 2 shows the accuracy and cross-validation values for each model. The Deep Learning model provides an accuracy of 72.74%, precision of 30.31%, and sensitivity of 31.38%. Figure 3 shows the confusion matrix used in calculating the performance matrix. The extreme gradient boosting (XGBoost) model outperformed the other models in predicting academic performance of academic achievement. The XGBoost model provides 97.12% accuracy and 35.67% cross-validation. Since the XGBoost model provides the best accuracy, this indicates that the XGBoost ML model is the ML model that best suits the nature of the data used. dataset and research objectives.

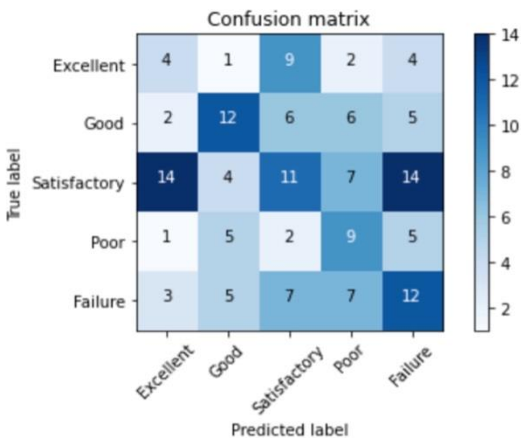

Figure 3. Deep learning confusion matrix.

CONLUSION

This research has reinforced and explored how machine learning can empower educators in decision-making in rural area. Predicting students' science learning academic achievement is an important concept in overcoming the crisis of students' academic success. This research used several ML classification models to predict students' academic achievement in science learning. The results showed an accuracy range from 33% to 98% and a cross-validation range from 30% to 37%. The XGBoost model is the most suitable ML model by achieving 97.12% accuracy at 35.67% cross-validation. In addition, the results showed that the number of school absences, current health status, going out with friends, free time after school, quality of family relationships are significant features that affect learners' academic performance. The study concludes that this research can help educators living in rural area, particularly Nias Regency, to identify gaps in learners' learning and enable early detection of underperforming learners, thus empowering educators in making decisions, which in turn can improve learners' academic success in the learning process.

REFERENCES

Bhutto, E. S., Siddiqui, I. F., Arain, Q. A., & Anwar, M. (2020). Predicting Students' Academic Performance Through Supervised Machine Learning. *2020 International Conference on Information Science and Communication Technology (ICISCT)*, 1–6. https://doi.org/10.1109/ICISCT49550.2020.9080033

Chen, S., Jamiatul Husnaini, S., & Chen, J.-J. (2020). Effects of games on students' emotions of learning science and achievement in chemistry. *International Journal of Science Education*, *42*(13), 2224–2245. https://doi.org/10.1080/09500693.2020.1817607

Costa-Mendes, R., Oliveira, T., Castelli, M., & Cruz-Jesus, F. (2021). A machine learning approximation of the 2015 Portuguese high school student grades: A hybrid approach. *Education and Information Technologies*, *26*(2), 1527–1547. https://doi.org/10.1007/s10639-020-10316-y

Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th Editio). SAGE Publications, Inc.

Harefa, E. (2023). Implementation of Scientific Inquiry Approach for Enhancing Scientific Literacy among Elementary Students. *IRAONO: Journal of Elementary and Childhood Education*, *1*(1), 32–38. https://doi.org/https://doi.org/10.56207/iraono.v1i1.13

Harefa, E., Kristiyanto, W. H., & Rondonuwu, F. S. (2019). Visualization of Conduction Heat Transfer using Augmented Reality Technology. *Indian Journal of Science and Technology*, *12*(21), 1–6. https://doi.org/10.17485/ijst/2019/v12i21/139551

Harefa, E., Li, N., & Zhou, W. (2022). Application of laser-induced breakdown spectroscopy with a generalized regression neural network and LASSO-type methods for estimation of arsenic and chromium in soil. *Journal of Analytical Atomic Spectrometry*, *37*(6), 1340–1349. https://doi.org/10.1039/D2JA00060A

Harefa, E., & Zhou, W. (2021). Performing sequential forward selection and variational autoencoder techniques in soil classification based on laser-induced breakdown spectroscopy. *Analytical Methods*, *13*(41), 4926–4933. https://doi.org/10.1039/D1AY01257F

Harefa, E., & Zhou, W. (2022). Laser-Induced Breakdown Spectroscopy Combined with Nonlinear Manifold Learning for Improvement Aluminum Alloy Classification Accuracy. *Sensors*, *22*(9), 3129. https://doi.org/10.3390/s22093129

Hasan, H. M. R., Rabby, A. S. A., Islam, M. T., & Hossain, S. A. (2019). Machine Learning

Algorithm for Student's Performance Prediction. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–7. https://doi.org/10.1109/ICCCNT45670.2019.8944629

Jayaprakash, S., Krishnan, S., & Jaiganesh, V. (2020). Predicting Students Academic Performance using an Improved Random Forest Classifier. *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 238–243. https://doi.org/10.1109/ESCI48226.2020.9167547

Johnson, A., Kuhfeld, M., & Soland, J. (2021). The Forgotten 20%: Achievement and Growth in Rural Schools Across the Nation. *AERA Open*, *7*, 233285842110520. https://doi.org/10.1177/23328584211052046

Kemdikburistek. (2023). *Data Sekolah*. https://dapo.kemdikbud.go.id/sp/2/071100

Lye, C.-T., Ng, L.-N., Hassan, M. D., Goh, W.-W., Law, C.-Y., & Ismail, N. (2010). Predicting Pre-university Student's Mathematics Achievement. *Procedia - Social and Behavioral Sciences*, *8*, 299–306. https://doi.org/10.1016/j.sbspro.2010.12.041

Mukred, M., Mokhtar, U. A., Hawash, B., AlSalman, H., & Zohaib, M. (2024). The adoption and use of learning analytics tools to improve decision making in higher learning institutions: An extension of technology acceptance model. *Heliyon*, *10*(4), e26315. https://doi.org/10.1016/j.heliyon.2024.e26315

OECD. (2023). *PISA 2022 Results (Volume I): The state of learning and qquity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Paolucci, C., Vancini, S., Bex II, R. T., Cavanaugh, C., Salama, C., & de Araujo, Z. (2024). A review of learning analytics opportunities and challenges for K-12 education. *Heliyon*, *10*(4), e25767. https://doi.org/10.1016/j.heliyon.2024.e25767

Rimadana, M. R., Kusumawardani, S. S., Santosa, P. I., & Erwianda, M. S. F. (2019). Predicting Student Academic Performance using Machine Learning and Time Management Skill Data. *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 511–515. https://doi.org/10.1109/ISRITI48646.2019.9034585

Rosmana, P. S., Iskandar, S., Fadilah, N., Azhar, N., Oktavini, D., & Munte, A. C. (2022). Upaya Pemerataan Pendidikan Berkelanjutan Di Daerah 3T. *Attadib: Journal of Elementary Education*, *6*(2), 405–418.

Sandra, L., Lumbangaol, F., & Matsuo, T. (2021). Machine Learning Algorithm to Predict Student's Performance: A Systematic Literature Review. *TEM Journal*, *10*(4), 1919–1927. https://doi.org/10.18421/TEM104-56

Suhaimi, N. M., Rahman, S. A., Mutalib, S., Hamid, N. H. A., & Mallik, A. M. A. (2019). Review on Predicting Students' Graduation Time Using Machine Learning Algorithms.

International Journal of Modern Education and Computer Science, *11*(7), 1–13. https://doi.org/10.5815/ijmecs.2019.07.01

Vijayalakshmi, V., & Venkatachalapathy, K. (2019). Comparison of Predicting Student's Performance using Machine Learning Algorithms. *International Journal of Intelligent Systems and Applications*, *11*(12), 34–45. https://doi.org/10.5815/ijisa.2019.12.04

Wang, D., Lian, D., Xing, Y., Dong, S., Sun, X., & Yu, J. (2022). Analysis and Prediction of Influencing Factors of College Student Achievement Based on Machine Learning. *Frontiers in Psychology*, *13*, 1–11. https://doi.org/10.3389/fpsyg.2022.881859

Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. *Computers in Human Behavior*, *98*, 166–173. https://doi.org/10.1016/j.chb.2019.04.015

Yu, S.-J., Hsueh, Y.-L., Sun, J. C.-Y., & Liu, H.-Z. (2021). Developing an intelligent virtual reality interactive system based on the ADDIE model for learning pour-over coffee brewing. *Computers and Education: Artificial Intelligence*, *2*, 100030. https://doi.org/10.1016/j.caeai.2021.100030