

STUDENT ACADEMIC PERFORMANCE ANALYSIS USING SUPPORT VECTOR REGRESSION AND MULTILAYER PERCEPTRON: AN EDUCATIONAL DATA MINING APPROACH

Nisa Miftachurohmah^{1*}, Nasruddin²

¹ Computer Science, Faculty of Information Technology, Universitas Sembilanbelas November Kolaka, Kolaka, 93511, Indonesia

² Mathematics Education, Faculty of Teacher Training and Education, Universitas Sembilanbelas November Kolaka, Kolaka, 93511, Indonesia

Email: ^{*1}nisa.informatics@gmail.com, ²nash.matematika@gmail.com

(Received: 03 February 2026, Revised: 10 February 2026, Accepted: 12 March 2026)

Abstract

Predicting student academic performance lays a foundation for data-informed educational decisions. This work uses the public Student Performance Data set from Kaggle website which contains 10,000 records and is used for predicting 'Performance Index'. Median imputation, one hot encoding for categorical variables and feature standardization were used for preprocessing of data. The model was evaluated through 5-fold cross-validation, and the proportions of training and testing data were set at 80:20 in each fold. Two different regression models were utilized: Support Vector Regression (SVR) with RBF kernel and a Multilayer Perceptron (MLP) consisting of two hidden layers (128–64 neurons). Both models achieved excellent prediction accuracy. SVR achieved an MAE of 1.6653, RMSE of 2.0991 and $R^2=0.9881$ whereas MLP slightly performed better than SVR with a MAE 1.6596, RMSE of 2.0872 and $R^2=0.9882$. Learning curve analysis showed stable convergence with little overfitting. The results show their efficiency and they are both kernel-based and neural network-based methods to predict academic performance. Future work will need to test the models on much more diverse data sets and may incorporate further context variables to improve model robustness and interpretability.

Keywords: *Educational Data Mining, Student Performance Prediction, SVM, MLP, Machine Learning.*

This is an open access article under the [CC BY](#) license.



**Corresponding Author: Nisa Miftachurohmah*

1. INTRODUCTION

Examining students' academic performance is a crucial method of determining how well learning is taking place in schools. In order to develop effective teaching strategies that enhance academic outcomes and ultimately improve educational quality, schools, educators, and policymakers must understand the factors influencing students' academic success [1,2]. However, traditional statistical approaches often struggle to capture complex and nonlinear patterns embedded in the growing volume of educational data derived from students' learning activities, behavioral records, and cognitive assessments. Therefore, data-driven analytical approaches have become increasingly necessary in educational research [3].

Educational Data Mining (EDM) aims to extract meaningful patterns and predictive insights from educational data to better understand how students

learn and perform academically [4,5]. By identifying hidden relationships and predictive structures, EDM enables forecasting of academic outcomes and supports evidence-based decision-making in education [6]. Machine learning-based predictive models are widely adopted in EDM because of their ability to handle high-dimensional data and model nonlinear relationships effectively [7,8,9].

In previous studies have compared regression algorithms in determining the model used in creating a web-based final grade prediction system, with an MAE value of 4.35 [10], the SVR model can predict. Setiadi *et al* have investigated the predictive power of learning management system (LMS) log data for student performance in higher education and Analyzed interactions from 114 students in a sports pedagogy course, with an MAE value of 5.39 [11]. Beckham *et al* have predicted student grade using ML models to prove whether that factor actually affects student

grade, their used MLP 12-Neuron model performs the best with RMSE value of 4.32 [12].

Previous studies have investigated various machine learning techniques for predicting student academic performance. Linear regression and decision-based algorithms are frequently applied due to their interpretability and simplicity [13,14]. Predictive performance is often highly dependent on dataset characteristics, feature representation, and validation design. For this reason, this study focuses specifically on regression-based paradigms to examine differences in generalization behavior between kernel-based and neural-network-based models under controlled experimental settings.

Support Vector Regression (SVR) is widely recognized for its ability to model nonlinear relationships using kernel functions while maintaining strong generalization through structural risk minimization [15,16,17]. Its ϵ -insensitive loss function allows the model to remain robust against noise and outliers, making it suitable for structured educational datasets that contain heterogeneous behavioral attributes. On the other hand, Multilayer Perceptron (MLP), as a neural-network-based regression model, offers high representational flexibility and the capacity to capture complex feature interactions through multilayer nonlinear transformations [18,19]. This flexibility enables MLP to model intricate relationships among behavioral learning variables, although it may exhibit sensitivity to hyperparameter configuration and dataset size.

Many recent studies primarily emphasize predictive accuracy and model optimization, with growing attention to explain ability and performance enhancement techniques [20,21,22]. However, less attention has been given to systematically evaluating how structured behavioral learning attributes interact with regression-based modeling approaches under controlled experimental settings. Consequently, systematic evaluation under identical preprocessing procedures, hyperparameter tuning strategies, and validation schemes is needed to better understand model robustness and generalization stability in structured educational datasets [23,24,25].

This study compares Support Vector Regression and Multilayer Perceptron within an Educational Data Mining framework to predict student academic performance. Previous studies have explored multimodal prediction and comparative machine learning approaches in academic performance forecasting [26,27], yet systematic regression-based comparisons under fully controlled experimental settings remain limited. Rather than proposing a new algorithm, the contribution of this research lies in a controlled and reproducible comparative framework. Both models are evaluated under identical preprocessing pipelines and cross-validation protocols to ensure methodological consistency.

Additionally, the analysis explicitly incorporates structured behavioral learning features—including

prior academic performance, study habits, extracurricular participation, sleep patterns, and practice intensity. Beyond conventional evaluation metrics, learning curve analysis is employed to assess generalization stability and variance behavior. Through this approach, the study provides deeper insight into regression model dynamics in educational contexts and supports more reliable data-driven decision-making systems.

2. RESEARCH METHOD

As part of an EDM (Educational Data Mining) approach, this research follows a quantitative methodology to predict student academic achievement through regression-based machine learning techniques. The research methodology includes four major steps, namely data collection and pre processing, model construction, evaluation part and comparative study. All experiments were performed under standardized and reproducible conditions.

2.1 Data Set

In this work, we use a publicly available dataset from the Kaggle platform. The dataset consists of 10,000 student records and structured attributes from the education field which are related to academic performance. Predictors are measured by prior grades, hours of sleep, participation in clubs or activities, intensity of studying and other structured behavioral learning attributes. The goal is to predict the continuous score of academic performance, which makes it convenient to be used as a regression prediction task data set.

The dataset has been split into training and testing subsets in an 80:20 ratio to guarantee reproducibility and a fair model evaluation. In particular, 8,000 records were for the model training and 2,000 records for independent testing. Application of a publicly available data set improves maintainability of the experiment and possible replication in similar conditions.

2.2 Data Preprocessing

Data preprocessing was applied to make the model compatible and facilitate stable learning. Categorical features were converted to numerical values and subsequently used for the model calculation. Z-score standardization was performed on the numerical features to make feature scaling consistent for all variables.

Feature scaling is essential for SVR and MLP, which are sensitive to the differences of feature magnitude. Normalization helps features with large scale not to dominate in the learning, while at the same time allows stable convergence.

In order to make comparison fair, both models were developed with the same pretrained processing and hyperparameter tuning. The performance of the model was assessed by 5-fold cross-validation on the training set to evaluate generalization stability and

avoid overfitting. For this practice training data were divided into five parts, four for training and one for validating in a loop. The performance of the best model was then tested on the independent testing set.

2.3 Support Vector Regression (SVR)

Support Vector Regression aims to find a function $f(x)$ that deviates from the actual target values by at most ε , while maintaining model flatness.

The regression function is defined in Equation (1):

$$f(x) = w^T \phi(x) + b \quad (1)$$

where:

w : the weight vector,

$\phi(x)$: the bias term,

b : a nonlinear mapping to a higher-dimensional feature space.

The optimization problem of SVR is formulated in Equation (2):

$$\min_{w, \xi_i, \xi_i^*} \frac{1}{2} \|w\|^2 + C \sum_{i=0}^n \xi_i + \xi_i^* \quad (2)$$

subject to Equation (3), (4) and (5):

$$y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i \quad (3)$$

$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \quad (4)$$

$$\xi_i, \xi_i^* \geq 0 \quad (5)$$

where:

C : the regularization parameter,

ε : defines the margin of tolerance,

ξ_i, ξ_i^* : slack variables.

In this study, the Radial Basis Function (RBF) kernel is used in Equation (6):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (6)$$

where γ controls the kernel width.

2.4 Multilayer Perceptron (MLP)

The Multilayer Perceptron is a feedforward neural network composed of an input layer, hidden layer(s), and an output layer, Equation (7).

$$h_j = g \left(\sum_{i=1}^d w_{ij} x_i + b_j \right) \quad (7)$$

where:

w_{ij} : the weight connecting input neuron i to hidden neuron j ,

b_j : the bias term,

$g(\cdot)$: the activation function.

In this study, the ReLU activation function is used in Equation (8):

$$g(z) = \max(0, z) \quad (8)$$

The output layer computes the predicted academic performance as shown in Equation (9):

$$\hat{y} = \sum_{j=1}^m w_j h_j + b \quad (9)$$

The network parameters are optimized using the Mean Squared Error (MSE) loss function as defined in Equation (10):

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Model training is performed using the backpropagation algorithm with gradient-based optimization.

2.5 Model Evaluation Metrics

To evaluate and compare the performance of SVR and MLP, three standard regression metrics are employed in equation (11), (12), (13).

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

where \bar{y} is the mean of the observed values.

In k-fold cross-validation, the data is divided into k roughly equal parts, or "folds." In the case of 5 fold cross validation, data is divided into five broad groups. At iteration i, 4 folds are used for training the model and the others one is held for testing. This is repeated five times, with each fold acting as the test set exactly once. The overall evaluation performance is averaging the five iterations results to derive a more solid estimation on the generalization ability of the model.

If M_i represents the evaluation metric obtained in the i-th fold, then the overall performance is calculated as the average of all fold results, as shown in equation (13):

$$M_{avg} = \frac{1}{k} \sum_{i=1}^k M_i \quad (14)$$

For $k = 5$, the average performance metric is defined in equation (15):

$$M_{avg} = \frac{1}{5} \sum_{i=1}^5 M_i \quad (15)$$

So, the standard deviation across the five folds is computed in equation (16):

$$\sigma = \sqrt{\frac{1}{5} \sum_{i=1}^5 (M_i - M_{avg})^2} \quad (16)$$

2.6 Comparative Analysis

Both the SVR and MLP models were trained and tested with the same dataset, preprocessing steps, and evaluation metrics to guarantee fairness and reproducibility. The study consists of comparing the prospective quality estimating approaches in terms of their prediction accuracy, the error that is made on these predictions and the explanatory power which reports MAE, RMSE, and R².

3. RESULT AND DISCUSSION

The initial stage of this research began with an analysis of the input feature distribution to identify data characteristics prior to modeling. Figure 1 displays the distribution of input features before the normalization process, which includes Hours Studied, Previous Scores, Extracurricular Activities, Sleep Hours, and Sample Question Papers Practiced. There are considerable differences in the scales of features, particularly feature Previous Scores' scale is a lot bigger compared to other features. This type of phenomenon can lead to bias in model learning, particularly for models like Support Vector Regression (SVR) and Multilayer Perceptron (MLP) which are sensitive to data scale.

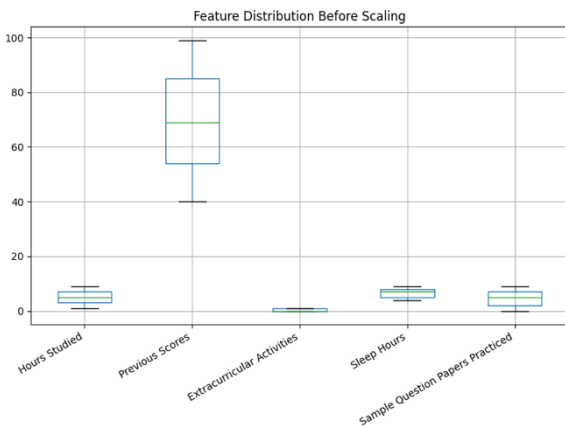


Figure 1. Feature Distribution Before Scaling

After standardization of the numeric features, the input feature distribution is displayed again in Figure 2. The results of visualization indicate that all numerical features are rescaled to a mean value approximately zero and same scale following the normalization, while the categorical feature Extracurricular Activities maintains original binary formatting. These results verify that our pipeline has removed any scale differences among the features by preprocessing, and all variables are treated as equally important during modeling.

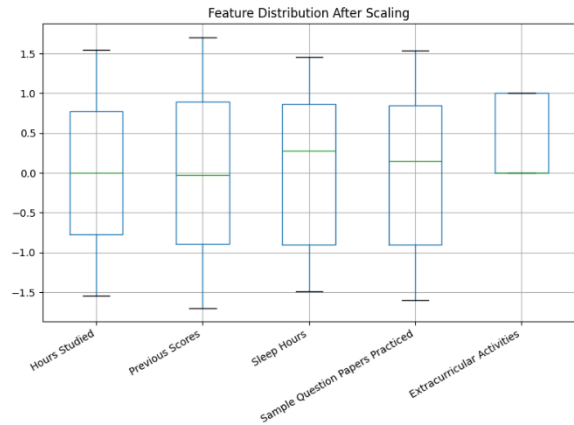


Figure 2. Feature Distribution After Scaling

3.1 Model Performance Evaluation

Model performance evaluation was conducted using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and coefficient of determination (R²) metrics. Table 1 shows the quantitative results in predicting alongshore variable with the SVR and MLP models.

Table 1. 5-Fold Cross-Validation Results of SVR and MLP Models

Model	MAE (Mean)	MAE (Std)	RMSE (Mean)	RMSE (Std)	R ² (Mean)	R ² (Std)
SVR	1.6653	0.0143	2.0991	0.0192	0.9881	0.0002
MLP	1.6596	0.0186	2.0872	0.0221	0.9882	0.0003

The results in Table 1 show that the SVR and MLP models reached quite high level of accuracy in predicting academic performance of students. The best performance for the MLP model was just a little better than SVR under all measures (MAE of 1.6596, RMSE of 2.0872 and R² value reached of 0.9882). However, the distinction between the two models is marginal; which means both appear to be very appropriate for this dataset. It can be seen that the standard deviation in each fold is very small, indicating high stability and consistency of model performance. In general, the results have shown that the kernel-based (SVR) and neural networkbased(MLP) are good methods to predict academic performance.

Table 2. Comparative Performance of SVR and MLP with Previous Studies

Model	Author	MAE	RMSE
SVR	Alboaneen, (2022)[10]	<i>et.al</i> 4.35	-
	Setiadi, <i>et.al</i> (2025)[11]	5.39	-
	Miftachurohmah, (2026)	<i>et.al</i> 1.6653	2.0991
MLP	Beckham, (2023)[12]	<i>et.al</i> -	4.32

Miftachurohmah, <i>et.al</i> (2026)	1.6596	2.0872
--	--------	--------

Table 2 Comparison results between our proposed SVR and MLP models with those reported in previous works. If we compare them with those of cross-validation in Table, it is found that the prediction ability of our models is better than such comparison. Increasing the width of both models has no practical difference: the MAE and RMSE for SVR are 1.6653 and 2.0991 on average, while those for MLP are slightly lower with MAE = 1.6596 and RMSE = 2.0872, in both cases with an $R^2 \sim 0.988$. The results are in Table 1. These results demonstrate the extreme predictability and stability of model performance across different folds.

That's much better than in previous studies. For instance, the SVR model of Alboaneen *et al.* (2022) with MAE of 4.35, and Setiadi *et al.* (2025) who found an MAE 5.39, this is statistically higher than the one obtained in the present study (1.6653). The MLP model of Beckham *et al.* (2023) obtained a RMSE of 4.32 while the present study reduced it to 2.0872. These decreases indicate that the modeling approach presented has much better prediction accuracy.

This superior result in Table 2 is supported by the stability and robustness of these approaches in Table 1, where the very low standard deviation over the same 5-fold cross-validation demonstrates that they generalize extremely well. The improvement of the performance is caused due to structured preprocessing pipeline, proper scaling of features, optimal hyperparameter settings and strict cross validation schema. However, due to variations in dataset properties, feature descriptors and educational frameworks across studies direct numerical comparisons must be interpreted with care.

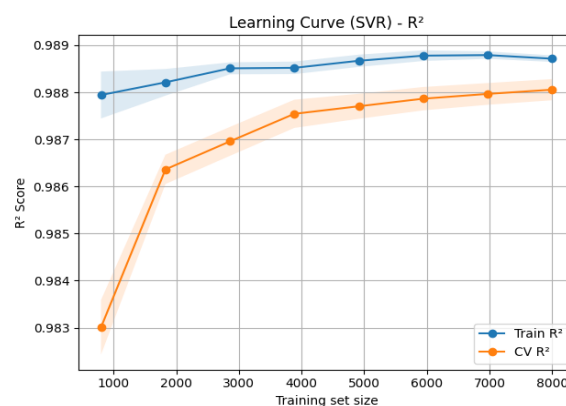
In general, matching between Table 1 and Table 2 reinforces the conclusion that the proposed SVR and MLP models are highly competitive and robust solutions for student academic performance prediction in educational data mining framework.

3.2 Analysis of Learning Curve

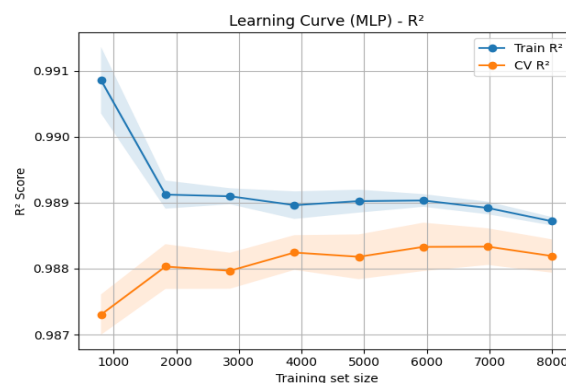
The generalization of the both models was then analyzed by means of learning curve. The learning curve of the R^2 metric of the SVR model is presented in Figure 3(a) as training set size increases. The R^2 training values are very high for all models (≈ 0.988 – 0.989), which means that the model fits the training data very well. Similarly, the cross-validation (CV) R^2 losses increase gradually with increasing training size converging to the training performance and stabilizing at around 0.988. The small distance between the training and validation curves is an indication of less likely overfitting of the SVR model, implying its generalization ability.

The learning curve of the Multilayer Perceptron (MLP) model is reported in Fig. 3(b) in terms of R^2 metric. First, the training : R^2 values are slightly better (just over > 0.990), which means that the NN is well

suiting to learn efficiently from small datasets. But as we keep increasing the training size, the training score decreases by a small margin and it becomes constant near 0.989 while the cross-validation R^2 keeps getting better and converging at around 0.988. This behaviour represents a slight decrease of the initial overfitting with more data. The small and quite stable gap of the training and validation curves illustrates that the MLP model does not suffer from high variance. In general, the convergence pattern validates that the chosen two-hidden-layer structure (128–64 neurons) is well set and converges to a good solution for predicting academic performance.



(a)



(b)

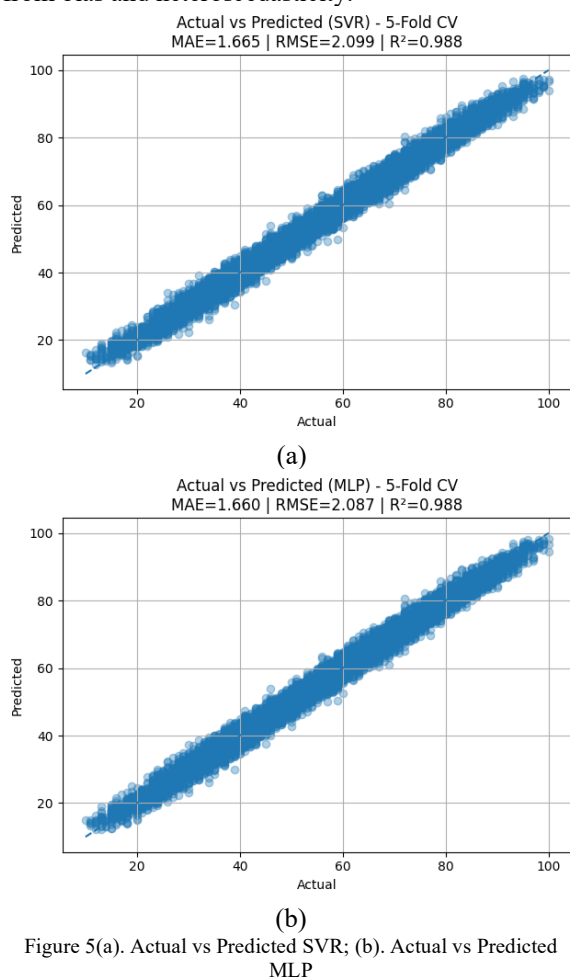
Figure 3(a). Learning of SVR; (b). Learning Curve of MLP

Both plots in Figure 3(a) and 3(b) show strong generalization ($R^2 > 0.98$). The SVR model somewhat converges more stable, whereas the MLP comes first with a slightly better initial training performance but then behaves comparably as number of samples increases. The close proximity of training and validation curves in both models verifies the lack of overfitting as well as confirms that the 5-fold cross-validation paradigm adopted in this investigation is robust.

3.3 Actual vs. Predicted Analysis

The real versus predicted value plots are depicted in Figure 4(a) for SVR and Figure 4(b) for MLP. Figure 4(a) shows the scatter plot of true vs. predicted

academic performance using SVR model in 5-fold cross-validation. The prediction and actual values are strongly correlated along the 45-degree reference line. This alignment is a consequence of the high determination coefficient ($R^2 = 0.988$), which means that the model explains around 98.8% of variance in this response variable. The low mean absolute error (MAE) and root mean squared error (RMSE) of 1.665 and 2.099, respectively, are also indicative of no significant prediction errors. Furthermore, the even distribution of points in terms of performance range also implies that SVR keeps a stable prediction accuracy regardless whether for students with low or high academic scores. The lack of systematic patterns of deviation indicates that the model is fairly free from bias and heteroscedasticity.



Actual versus predicted values obtained from MLP model with 5-fold cross validation was shown in Figure 4(b). Akin to the SVR positions, the samples are well-statistically-in line with the diagonal reference line. The predictions give an The R^2 of the model is 0.988, with lower MAE (1.660) and or RMSE (2.087) than SVR which means it slightly outperforms its accuracy precautiously terminology SVR as a benchmark. The scatter points are still being light in the whole score range as a testament to stable performance and good generalization. There is no apparent curvature or systematic underestimation/

overestimation, which means that the two-hidden-layer MLP architecture (128–64 neurons) can be considered as tuned for modeling the underlying non-linear trends in the data.

Both the Fig. 4(a) and (b), showed a very highly predicted response as their R^2 values were nearly equal to one with negligible difference between the error values too. Although the MLP model performs slightly better in error metrics, the difference is insignificant and suggests that both SVM-based (SVR) and neural network-based (MLP) models are equally useful for predicting student academic performance. The strong linearity of the 5-fold cross-validation framework in both plots provides additional evidence to the robustness, reliability of our educational data mining approach.

4. CONCLUSION

This research aimed to predict students' academic performance based on two techniques: the Support Vector Regression (SVR) and Multilayer Perceptron (MLP) using Educational Data Mining (EDM). As can be seen in Figure 7, both models performed well on predictive ability under cross-validation ($R^2 \approx 0.988$). The SVR model achieved an MAE of 1.6653 and RMSE of 2.0991, which is also less compared with the results generated by the MLP model (MAE = 1.6596 and RMSE = 2.0872). The small discrepancies between training and validation scores, and almost minimal std since the train/dev splits are reliable predictors of strong model stability and test performance. Learning curve and actual-versus-predicted analysis are used to demonstrate the stability and accuracy of both methods. Results indicate that both kernel-based (SVR) and neural networkbased (MLP) approaches are well-suited for modeling structured pattern in academic performance database. Future work could involve validation on a more diverse set of educational datasets and integration of potentially relevant context and behaviour features as well experimenting with ensemble or deep learning models.

5. REFERENCES

- [1] Z. Jinmin and F. Qi, "Relationship between learning flow and academic performance among students: a systematic evaluation and meta-analysis," *Front. Psychol.*, vol. 14, Nov. 2023, doi: 10.3389/fpsyg.2023.1270642.
- [2] I. Wickramasinghe, R. Aragon, and J. R. Valles, "Influential factors on elementary students' academic performance and identifying the appropriate performance group," *Discover Education*, vol. 3, no. 1, p. 76, Jun. 2024, doi: 10.1007/s44217-024-00167-x.
- [3] N. Miftachurohmah, N. Nasruddin, and M. Palobo, "The Impact of ICT-Integrated Learning on High School Students' Achievement in Mathematics: A Data-Driven Analysis," *RIGGS: Journal of Artificial Intelligence and Digital*

- Business*, vol. 4, no. 4, pp. 4197–4205, Dec. 2025, doi: 10.31004/riggs.v4i4.4039.
- [4] C. Simionescu, M. Danubianu, B. C. Gradinaru, and M. S. Maciucă, “Educational Data Mining in European Union – Achievements and Challenges: A Systematic Literature Review,” *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 3, 2024, doi: 10.14569/IJACSA.2024.0150386.
- [5] M. Bin Mahbub, “EXPLORING THE IMPACT OF DATA MINING IN EDUCATION THROUGH A QUALITATIVE STUDY OF LEARNING PATTERNS,” *Acta Electronica Malaysia*, vol. 8, no. 2, pp. 68–73, Oct. 2024, doi: 10.26480/aem.02.2024.68.73.
- [6] N. Nasruddin, N. Miftachurohmah, P. R. Sianturi, I. Iriawati, and I. Iwan, “Penerapan Regresi Robust dalam Educational Data Mining (EDM) untuk Klasifikasi Tingkat Penalaran Abduktif Siswa,” in *STAINS: Seminar Teknologi dan Sains*, Kediri: Program Studi Teknik Informatika - UNP Kediri, Jan. 2026, pp. 756–761.
- [7] Y. Xia, “Leveraging Statistical and Machine Learning Techniques to Uncover Determinants of English Performance in Chinese Middle School Students,” Jun. 26, 2024. doi: 10.21203/rs.3.rs-4505011/v1.
- [8] M.-R. Borna, H. Saadat, A. T. Hojjati, and E. Akbari, “Analyzing click data with AI: implications for student performance prediction and learning assessment,” *Front. Educ. (Lausanne)*, vol. 9, Dec. 2024, doi: 10.3389/educ.2024.1421479.
- [9] E. López-Meneses, L. López-Catalán, N. Pelicano-Piris, and P. C. Mellado-Moreno, “Artificial Intelligence in Educational Data Mining and Human-in-the-Loop Machine Learning and Machine Teaching: Analysis of Scientific Knowledge,” *Applied Sciences*, vol. 15, no. 2, p. 772, Jan. 2025, doi: 10.3390/app15020772.
- [10] D. Alboaneen, M. Almelih, R. Alsubaie, R. Alghamdi, L. Alshehri, and R. Alharthi, “Development of a Web-Based Prediction System for Students’ Academic Performance,” *Data (Basel)*, vol. 7, no. 2, p. 21, Jan. 2022, doi: 10.3390/data7020021.
- [11] H. Setiadi, F. A. I. Josi, Winarno, A. Doewes, and R. W. Damayanti, “Optimizing Student Performance Prediction: A Comparative Analysis of Regression Algorithms and Feature Selection Techniques on LMS Log Data,” *Vietnam Journal of Computer Science*, pp. 1–42, Dec. 2025, doi: 10.1142/S2196888825500265.
- [12] N. R. Beckham, L. J. Akeh, G. N. P. Mitaart, and J. V. Moniaga, “Determining factors that affect student performance using various machine learning methods,” *Procedia Comput. Sci.*, vol. 216, pp. 597–603, 2023, doi: 10.1016/j.procs.2022.12.174.
- [13] M. Kumar, V. Bhardwaj, D. Thakral, A. Rashid, and M. T. Ben Othman, “Ensemble Learning Based Model for Student’s Academic Performance Prediction Using Algorithms,” *Ingénierie des systèmes d’information*, vol. 29, no. 5, pp. 1925–1935, Oct. 2024, doi: 10.18280/isi.290524.
- [14] Z. Zhan and T. Shen, “Development of a prediction model for student teaching satisfaction based on 10 machine learning algorithms,” *Sci. Rep.*, vol. 15, no. 1, p. 36547, Oct. 2025, doi: 10.1038/s41598-025-19039-x.
- [15] B. Bargam, A. Boudhar, C. Kinnard, H. Bouamri, K. Nifa, and A. Chehbouni, “Evaluation of the support vector regression (SVR) and the random forest (RF) models accuracy for streamflow prediction under a data-scarce basin in Morocco,” *Discover Applied Sciences*, vol. 6, no. 6, p. 306, Jun. 2024, doi: 10.1007/s42452-024-05994-z.
- [16] Y. Liu, L. Wang, and K. Gu, “A support vector regression (SVR)-based method for dynamic load identification using heterogeneous responses under interval uncertainties,” *Appl. Soft Comput.*, vol. 110, p. 107599, Oct. 2021, doi: 10.1016/j.asoc.2021.107599.
- [17] D. Tao, Q. Ma, S. Li, Z. Xie, D. Lin, and S. Li, “Support Vector Regression for the Relationships between Ground Motion Parameters and Macroseismic Intensity in the Sichuan–Yunnan Region,” *Applied Sciences*, vol. 10, no. 9, p. 3086, Apr. 2020, doi: 10.3390/app10093086.
- [18] J. E. G. I. M. and T. J. L. Johnston, “Uncovering Student Engagement Patterns in Moodle with Interpretable Machine Learning,” 2024.
- [19] Z. Hu, “Data-driven innovative models of personalized teaching,” *Discover Education*, vol. 4, no. 1, p. 364, Sep. 2025, doi: 10.1007/s44217-025-00815-w.
- [20] T. Tong and Z. Li, “Predicting Learning Achievement Using Ensemble Learning with Result Explanation,” Jul. 26, 2024. doi: 10.21203/rs.3.rs-4674228/v1.
- [21] W. Ahmed, M. A. Wani, P. Plawiak, S. Meshoul, A. Mahmoud, and M. Hammad, “Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions,” *Sci. Rep.*, vol. 15, no. 1, p. 26879, Jul. 2025, doi: 10.1038/s41598-025-12353-4.
- [22] M. Fazil and B. M. Albahlal, “MultiFAR: Multidimensional information fusion with attention-driven representation learning for student performance prediction,” *PLoS One*, vol. 20, no. 10, p. e0333099, Oct. 2025, doi: 10.1371/journal.pone.0333099.
- [23] M. M. Hussain, S. Akbar, S. A. Hassan, M. W. Aziz, and F. Urooj, “Prediction of Student’s Academic Performance through Data Mining Approach,” *Journal of Informatics and Web*

- Engineering*, vol. 3, no. 1, pp. 241–251, Feb. 2024, doi: 10.33093/jiwe.2024.3.1.16.
- [24] M. Bellaj, A. Ben Dahmane, S. Boudra, and M. Lamarti Sefian, “Educational Data Mining: Employing Machine Learning Techniques and Hyperparameter Optimization to Improve Students’ Academic Performance,” *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 20, no. 03, pp. 55–74, Feb. 2024, doi: 10.3991/ijoe.v20i03.46287.
- [25] H. Setiawan, I. F. Tampati, I. G. M. Putra, and Girinoto, “Optimization of MLP-Regressor for Predicting Student’s Cumulative Grade Point Average (GPA),” in *2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, IEEE, Jul. 2024, pp. 628–633. doi: 10.1109/ICSINTESA62455.2024.10747977.
- [26] A. Kord, A. Aboelfetouh, and S. M. Shohieb, “Academic course planning recommendation and students’ performance prediction multi-modal based on educational data mining techniques,” *J. Comput. High. Educ.*, Jan. 2025, doi: 10.1007/s12528-024-09426-0.
- [27] I. Dawar, S. Negi, S. Lamba, and A. Kumar, “Enhancing Student Academic Performance Forecasting: A Comparative Analysis of Machine Learning Algorithms,” *SN Comput. Sci.*, vol. 5, no. 6, p. 758, Aug. 2024, doi: 10.1007/s42979-024-03118-3.