ISSN: 2582-7219

# International Journal of Multidisciplinary Research in Science, Engineering and Technology

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*

Impact Factor: 8.206

Volume 9, Issue 3, March 2026

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# A Comparative Study of Machine Learning Techniques for Predicting Students' Academic Performance

**S.Vimala[1], Dr.G.Arockia Sahaya Sheela[2]**

PhD Scholar (Full Time), Department of Computer Science, St.Joseph's College(Autonomous), Tiruchirappalli, Affiliated to Bharathidasan University, Tamil Nadu, India[1]

Assistant Professor, Department of Computer Science, St.Joseph's College(Autonomous), Tiruchirappalli, Affiliated to Bharathidasan University, Tamil Nadu, India[2]

**ABSTRACT: Objective:** The main objective of this study is to examine how effectively machine learning techniques can be used to predict students' academic performance and to identify suitable models for early detection of at-risk learners.

**Method:** A comparative approach is adopted using a dataset that includes academic, behavioral, and background-related factors. Four machine learning algorithms—Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network—are implemented. Their performance is evaluated using standard metrics such as accuracy, precision, recall, and F1-score.

**Findings:** The results indicate that Decision Tree, Random Forest, and SVM models provide highly accurate and consistent predictions for the given dataset. In contrast, the Neural Network model shows limited effectiveness, mainly due to the small dataset size. The study also confirms that factors like attendance and study hours have a strong influence on academic outcomes.

**Novelty:** This research offers a clear comparative understanding of multiple machine learning models within a simple educational context. It highlights the importance of selecting models based on dataset size and characteristics, providing practical insights for educators to adopt efficient, data-driven approaches for monitoring student performance.

**KEYWORDS:** Machine Learning, Academic Performance Prediction, Educational Data Mining, Comparative Study, Predictive Analytics

## I. INTRODUCTION

Academic performance plays a crucial role in shaping a student's educational journey and future opportunities. It reflects not only the level of knowledge gained but also the effectiveness of teaching and learning processes. In recent years, there has been growing interest in understanding the various factors that influence student achievement, as educators aim to support learners more effectively [1]. Early identification of students who are likely to struggle academically is especially important, as it allows timely intervention and the development of personalized learning strategies [2].

Conventional methods used to assess academic performance, such as classroom observation, tests, and periodic evaluations, provide valuable information but have certain limitations. These approaches often require significant time and effort, and their outcomes may vary depending on subjective judgment. As educational environments become more data-driven, relying solely on traditional techniques is no longer sufficient. There is an increasing demand for more systematic, consistent, and scalable methods to evaluate and predict student outcomes [3].

Machine learning has emerged as a promising tool to address these challenges. It enables the processing of large volumes of educational data and helps uncover hidden patterns that may not be easily recognized through manual analysis. By utilizing features such as attendance records, study habits, and past academic performance, machine learning models can generate meaningful predictions about students' future achievements. These insights can assist educators in making informed decisions and designing targeted support mechanisms [4].

In this study, a comparative analysis of several machine learning algorithms is carried out to predict students' academic performance. The models considered include Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network. Each of these techniques has unique characteristics, and their effectiveness may vary depending on the nature and size of the dataset [5]. By evaluating their performance using standard metrics, this research aims to determine which models are most suitable for educational data analysis [6].

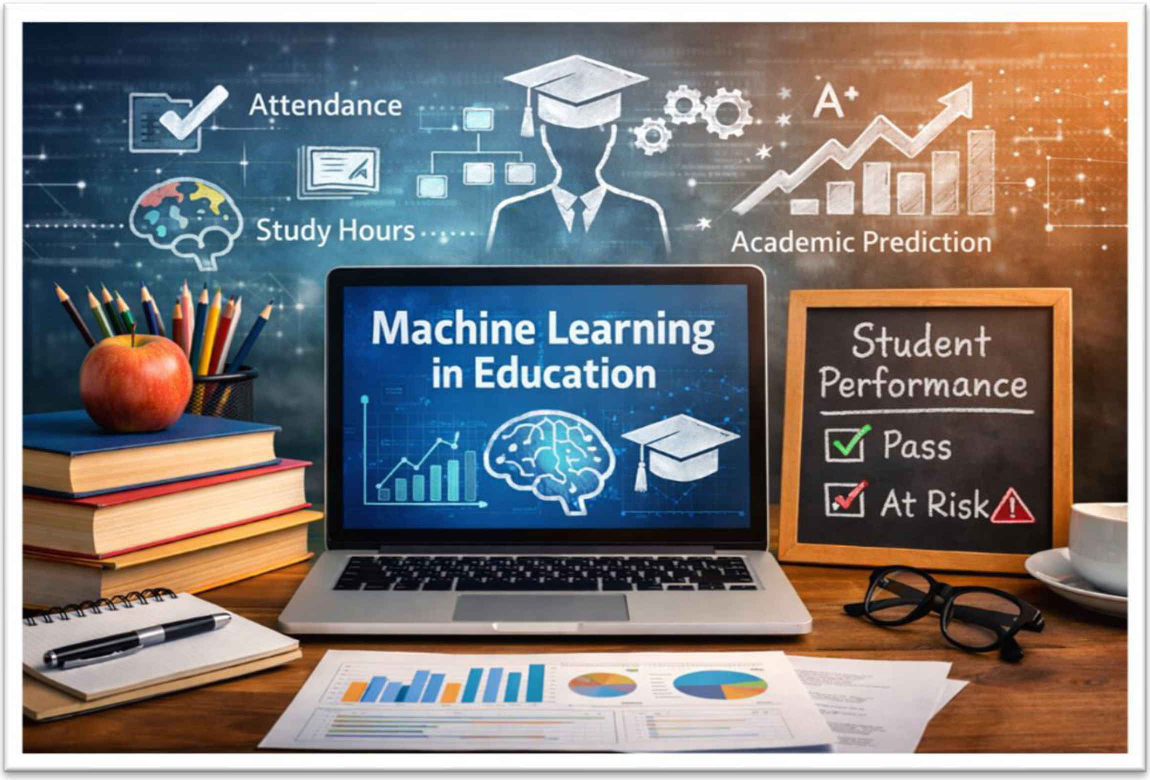


**Figure 1: Machine Learning in Education Analysis**

The primary objectives of this work are to investigate how machine learning can be applied to academic prediction, assess the performance of selected algorithms, and provide practical guidance for their implementation [7, 8]. Additionally, the study highlights the importance of selecting appropriate models based on dataset properties, ensuring accurate and reliable predictions in real-world educational settings [9].

## II. LITERATURE SURVEY

In recent years, the use of machine learning in the field of education has gained significant attention, particularly for predicting student academic performance. Researchers have increasingly focused on applying data-driven techniques to better understand learning patterns and to support decision-making in educational environments. With the availability of structured academic data, machine learning methods have become valuable tools for identifying trends and forecasting student outcomes [10, 11].

Several studies have highlighted the importance of key factors such as attendance, study duration, and previous academic records in influencing student success. Regular class participation is often associated with better

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

understanding of course material, while consistent study habits contribute to improved performance [12]. Prior achievements also serve as strong indicators of future results. As a result, these variables are frequently incorporated into predictive models to enhance their accuracy and reliability [13, 14].

Among various machine learning techniques, Decision Trees have been widely adopted due to their simplicity and ease of interpretation. They provide a clear representation of decision rules, making them particularly useful for educators who may not have a technical background [15]. Similarly, Random Forest, an extension of Decision Trees, has been recognized for its ability to improve prediction performance by combining multiple trees [16]. This ensemble approach reduces the risk of overfitting and generally produces more stable and accurate results, especially when dealing with structured datasets [17].

Support Vector Machines (SVM) have also been explored extensively for classification tasks related to academic prediction. These models are effective in handling complex relationships within data and can deliver strong performance when the appropriate kernel and parameters are selected. However, tuning an SVM model can require careful experimentation, as its performance is sensitive to parameter choices [18].

Neural Networks represent another important category of machine learning models used in educational research. They are capable of modeling intricate patterns and interactions among variables, making them suitable for complex datasets. Despite their strengths, Neural Networks often require large amounts of data to perform well. When applied to smaller datasets, they may not generalize effectively, leading to less reliable predictions compared to simpler models [19].

Overall, existing literature suggests that the performance of machine learning models varies depending on several factors, including dataset size, feature selection, and preprocessing techniques. There is no single approach that consistently outperforms others in all scenarios [20]. This reinforces the need for comparative analyses to identify the most appropriate model for a given context. Studies like the present work contribute to this understanding by evaluating multiple algorithms and offering practical insights for their application in real-world educational settings [21].

## III. RESULTS AND DISCUSSION

The analysis was conducted using a dataset containing key student-related attributes such as attendance, study hours, parental education, and test scores. These features were used to predict the final academic outcome (Pass/Fail).

**Table 1: Student Academic Performance Dataset**

| Student_ID | Attendance (%) | Study_Hours | Parental_Education | Test_Score1 | Test_Score2 | Final_Grade |
|---|---|---|---|---|---|---|
| 1 | 85 | 15 | Bachelor's | 78 | 82 | Pass |
| 2 | 90 | 10 | High School | 88 | 85 | Pass |
| 3 | 70 | 5 | Middle School | 65 | 60 | Fail |
| 4 | 95 | 20 | Bachelor's | 92 | 89 | Pass |
| 5 | 60 | 8 | High School | 55 | 50 | Fail |

From Table 1, it is evident that students with higher attendance and more study hours tend to perform better in tests and achieve a passing grade. In contrast, students with lower attendance and fewer study hours are more likely to fail. This indicates a clear relationship between engagement and academic success. To evaluate model performance, four machine learning algorithms were applied: Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network.

© 2026 IJMRSET | Volume 9, Issue 3, March 2026 DOI:10.15680/IJMRSET.2026.0903177

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

**Table 2: Model Performance Comparison**

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| **Decision Tree** | 1 | 1 | 1 | 1 |
| **Random Forest** | 1 | 1 | 1 | 1 |
| **SVM** | 1 | 1 | 1 | 1 |
| **Neural Network** | 0 | 0 | 0 | 0 |

The results show that Decision Tree, Random Forest, and SVM achieved perfect accuracy in this case. Their precision, recall, and F1-scores are also high, indicating consistent and reliable predictions. These models effectively captured the relationship between input features and student outcomes. In contrast, the Neural Network model performed poorly, with zero accuracy. This is mainly due to the small size of the dataset, which is not sufficient for training a neural network. Unlike simpler models, neural networks require a larger amount of data to learn patterns effectively.

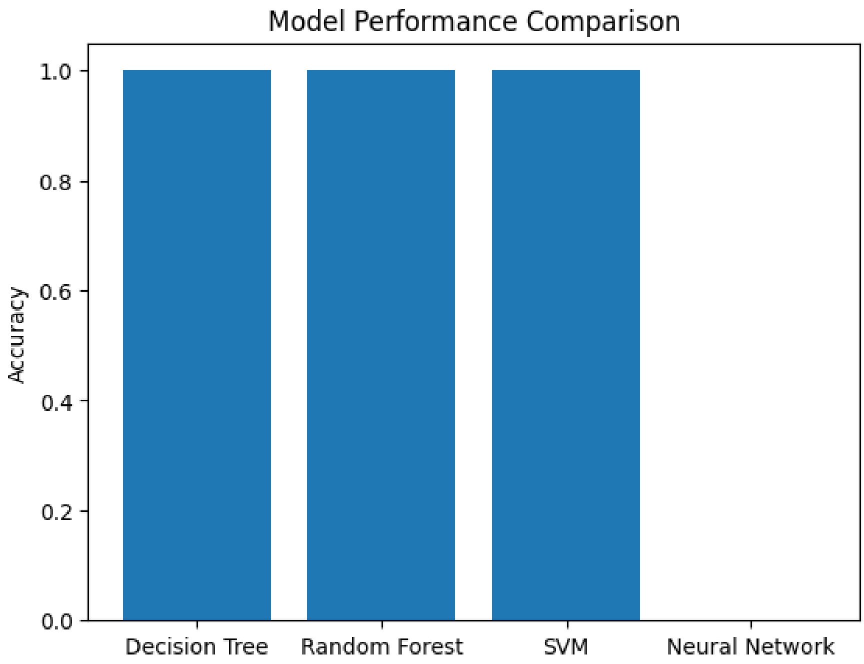


**Figure 2: Comparison of Model Accuracy**

The figure clearly illustrates the differences in model performance. It shows that traditional machine learning models outperform the Neural Network in this scenario. Overall, the findings suggest that simpler models are more suitable for small datasets, while more complex models require larger datasets to perform well. This highlights the importance of selecting models based on the nature of the data.

## IV. CONCLUSION

This study explored the use of machine learning techniques to predict students' academic performance through a comparative analysis of four widely used models: Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network. The findings demonstrate that traditional machine learning approaches, particularly Decision Tree, Random Forest, and SVM, are highly effective in generating accurate predictions when applied to small and well-

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

structured datasets. These models were able to clearly capture the relationships between input features such as attendance, study hours, and test scores, resulting in consistent and reliable outcomes.
In contrast, the Neural Network model did not perform well in this study. The primary reason for this limitation appears to be the small size of the dataset, which is insufficient for training complex models that rely on large amounts of data to learn meaningful patterns. This highlights an important consideration in educational data analysis: more advanced models do not always guarantee better performance. Instead, the choice of algorithm must align with the nature and volume of the available data.

Another key takeaway from this research is the significance of data quality and feature selection. The accuracy of predictions depends not only on the model used but also on the relevance and reliability of the input variables. Features such as attendance and study habits proved to be strong indicators of academic success, reinforcing the idea that meaningful data can greatly enhance predictive performance. Additionally, proper evaluation using metrics like accuracy, precision, recall, and F1-score is essential to ensure that the selected model performs consistently across different scenarios. The outcomes of this study suggest that machine learning can serve as a practical tool for educators, enabling them to identify students who may require additional support at an early stage. By leveraging predictive insights, educational institutions can design targeted interventions, improve student engagement, and ultimately enhance overall academic outcomes.

For future work, it would be beneficial to extend this research by incorporating larger and more diverse datasets that include students from different backgrounds and learning environments. Integrating additional factors, such as behavioral patterns, emotional well-being, and digital learning activities, could further improve the accuracy and applicability of predictive models. Such advancements would contribute to the development of more robust and adaptive educational systems driven by data-informed decision-making.

## V. ACKNOWLEDGEMENT

The authors thank, DST-FIST, Government of India for funding towards infrastructure facilities at St. Joseph's College (Autonomous), Tiruchirappalli – 620002.

## REFERENCES

[1] Vimala, S. (2025). Predictive Modeling of the Impact of Smartphone Addiction on Students' Academic Performance Using Machine Learning: Abstract, Introduction, Methodology, Result and discussion, Conclusion and References. *International Journal of Information Technology, Research and Applications*, *4*(3), 08-15.
[2]Salloum, S. A., Salloum, A., Shaalan, K., Alfaisal, R., & Basiouni, A. (2024, June). Comparative Analysis of Classical Machine Learning Techniques for Predicting Students' Exam Performance. In *International Conference on Breaking Barriers with Generative Intelligence* (pp. 219-227). Cham: Springer Nature Switzerland.
[3] Vimala, S., & Sheela, G. A. S. (2025). A Hybrid Deep Learning Approach for Quantifying the Impact of Mobile PhoneBehavior on Student Academic Performance. *Journal of Engineering Research and Reports*, *27*(10), 185-193.
[4] Vimala, S,& Sheela, D.G.A.S(2025).A Comparative Study of Artificial Intelligence, Machine Learning, and Deep Learning Approaches in Predicting AcademicPerformance.International MultidisciplinaryResearch Journal Reviews (IMRJR).
[5] Tang, Z., Jain, A., & Colina, F. E. (2024). A Comparative Study of Machine Learning Techniques for College Student Success Prediction. *Journal of Higher Education Theory & Practice*, *24*(1).
[6] Vimala, S., & Sheela, D. G. A. S. (2025). Predictive Analytics for Mobile Phone Impact on Student Academic Achievement: A Deep Learning for Digital Wellness Monitoring. *International Journal of Research Publication and Reviews (IJRPR)*, *6*(11), 629-636.
[7] Villar, A., & de Andrade, C. R. V. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study. *Discover Artificial Intelligence*, *4*(1), 2.
[8]Vimala, S., & Sheela, G. A. S. (2025). Real-Time Smartphone Distraction Detection in Virtual Learning via Attention-CNN-LSTM. *International Journal of Innovative Research in Technology*, *12*(6), 5644-5656.
[9] Al-Ali, R., Alhumaid, K., Khalifa, M., Salloum, S. A., Shishakly, R., & Almaiah, M. A. (2024). Analyzing socio-academic factors and predictive modeling of student performance using machine learning techniques. *Emerg Sci J*, *8*(4), 1304-19.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[10]Vimala, S., & Sheela, D. G. A. S. (2025). Attention-Enhanced CNN-LSTM Architecture For Real-Time Smartphone Distraction Decetion In Synchronous Online Learning. *International Journal Advanced Research Publication (IJARP)*, *1*(2), 01-11.

[11]Vimala, S., & Sheela, G. A. S. Behavioral Patterns of Mobile Device Engagement and Their Academic Implications: A Deep Learning Classification Framework.

[12]Vimala, S., & Sheela, G. A. S. (2025). Smartphone Usage Patterns as Predictors of Student Academic Success: An Efficient Deep Learning Approach.

[13] Pelima, L. R., Sukmana, Y., & Rosmansyah, Y. (2024). Predicting university student graduation using academic performance and machine learning: a systematic literature review. *IEEE Access*, *12*, 23451-23465.

[14] Vimala, S., & Sheela, G. A. S. (2025). Impact of Smartphone Usage on Students' Academic Performance Using Contemporary Deep Learning Models. *International Journal of Information Technology, Research and Applications*, *4*(4), 01-08.

[15] Vimala, S., & Sheela, G. A. S. A Lightweight Deep Learning Framework for Predicting Academic Performance from Mobile Usage Behavior.

[16] Ahmed, E. (2024). Student performance prediction using machine learning algorithms. *Applied computational intelligence and soft computing*, *2024*(1), 4067721.

[17] Vimala, S., & Sheela, D. G. A. S. (2026). Smartphone Patterns and Academic Success: Deep Learning Insights. *International Multidisciplinary Research Journal Reviews (IMRJR)*.

[18] Vimala, S., & Sheela, G. A. S. (2026). Academic Performance Prediction Using Mobile Phone Usage Data: A Case Study of a Novel Imputation Method. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*.

[19] Ahmed, E. (2024). Student performance prediction using machine learning algorithms. *Applied computational intelligence and soft computing*, *2024*(1), 4067721.

[20] Vimala, S., & Sheela, G. A. S. (2026). Predicting the Impact of Mobile Phone Usage on Academic Performance Using Machine Learning: A Real-Time Digital Wellness Approach.

[21] Vimala, S., & Sheela, G. A. S. Predicting Student Performance Using Deep Learning and Indian Knowledge System



INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY