ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1354

AI-Driven Student Performance Prediction Using Multi-Class Classifiers

Vijaya Kumar K.[1]; Dr. Naveen A.[2]

[1,2]Don Bosco College (Co-Ed), Yelagiri Hills, TN, India – 635853

Publication Date: 2025/10/07

Abstract: The integration of Artificial Intelligence (AI) into education has significantly transformed traditional teaching and learning practices by enabling personalized and adaptive learning experiences designed to meet the unique needs of individual students. This study introduces a comprehensive AI-driven framework for predicting student performance using a variety of machine learning classifiers, including J48 Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The framework evaluates multiple dimensions of student data—academic achievements, behavioral patterns, psychological factors, and skill assessments—to categorize learners into three distinct groups: fast, moderate, and slow performers, thereby facilitating targeted and timely academic interventions. The research utilized a dataset comprising 100 student records with 39 attributes, which underwent extensive preprocessing steps such as normalization, feature selection, and data anonymization to maintain privacy and ensure consistency. Experimental analysis demonstrated that the J48 Decision Tree classifier outperformed the others, achieving 100% accuracy, followed by Random Forest (98.5%) and SVM (96.7%). Further correlation analysis revealed that attributes such as aptitude, attendance, and motivation had a strong positive impact on CGPA, while stress levels showed a negative relationship. Additionally, an ANOVA test confirmed the statistical significance of these findings, validating the robustness of the proposed model. This framework highlights the potential of AI to revolutionize education by supporting real-time, personalized interventions and empowering educators with actionable insights for decision-making. Future work will explore the integration of deep learning models, real-time feedback mechanisms, and scalability to accommodate diverse educational settings.

Keywords: *Artificial Intelligence, Personalized Learning, Machine Learning, Student Performance Prediction, J48 Decision Tree, Outcome-Based Education.*

How to Cite: Vijaya Kumar K.; Dr. Naveen A. (2025) AI-Driven Student Performance Prediction Using Multi-Class Classifiers. *International Journal of Innovative Science and Research Technology,* 10(9), 2639-2645 https://doi.org/10.38124/ijisrt/25sep1354

I. INTRODUCTION

The rapid growth of Artificial Intelligence (AI) has significantly influenced the education sector, shifting it away from traditional, standardized teaching methods toward adaptive, data-driven learning approaches. In conventional education systems, all learners are typically exposed to the same instructional material, regardless of their individual learning speeds, comprehension abilities, or academic progress [21]. This uniform approach often leads to student disengagement, widening learning gaps, and ultimately results in less-than-optimal academic outcomes.

To overcome these limitations, AI-powered personalized learning systems have been developed as innovative tools that can analyse a broad range of student data—including academic performance, psychological well-being, and behavioural patterns—to create customized learning experiences [1], [2], [3]. These systems not only improve student motivation and engagement but also equip educators with real-time insights, enabling them to adapt teaching strategies dynamically to suit the specific needs of individual learners.

This study focuses on the application of machine learning algorithms to predict student performance and identify learners who may require targeted interventions [12], [19]. By incorporating predictive modeling, educational institutions can transition from reactive measures to proactive, evidence-based decision-making. The research investigates five widely used machine learning techniques—J48 Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)—to assess their effectiveness in classifying students into three performance categories: fast, moderate, and slow learners [4], [5], [6]. This classification is intended to guide educators in designing personalized mentoring programs, adaptive curriculum structures, and early support systems for at-risk students.

The overarching goal of this research is to integrate AI-driven predictive models within the Outcome-Based

Education (OBE) framework. Such integration ensures that educational practices are aligned with industry requirements, producing graduates who are both academically proficient and equipped with essential skills for employability and lifelong learning.

By combining real-time analytics with predictive algorithms, this study demonstrates how AI can bridge the gap between traditional instructional methods and modern personalized education, fostering a more equitable, efficient, and student-centered learning environment.

➢ *Problem Statement*

Fig 1 Education Model

Traditional educational systems often rely on standardized, uniform teaching approaches where all students are provided with the same instructional materials and follow identical learning paths. While this approach is easy to implement, it fails to consider the diverse cognitive abilities, learning styles, and pace of progress among individual students. As a result, high-performing students may not be adequately challenged, while those who require additional support risk falling behind. This lack of personalization can lead to disengagement, academic underperformance, and increased dropout rates [7], [8].

Traditional assessment methods depend largely on manual evaluation, which is not only time-consuming but also susceptible to human bias and errors. These conventional systems often result in delayed or incomplete feedback, making it challenging for educators to respond quickly and provide personalized academic support. As higher education becomes increasingly complex and classrooms grow more diverse, these challenges are magnified, particularly in large student groups, where it is nearly impossible for instructors to monitor and address each learner’s needs in real time [10], [13].

Although many institutions have adopted Learning Management Systems (LMS) and other digital tools, these platforms primarily focus on content delivery and basic analytics, offering limited support for advanced predictive capabilities or adaptive interventions. As a result, there is an urgent need for intelligent, data- driven solutions that can analyze diverse student data—including academic performance, behavioral traits, and psychological factors—to accurately predict learning outcomes and enable proactive, personalized interventions.

This study addresses these challenges by leveraging machine learning algorithms to develop a predictive framework capable of classifying students into performance tiers such as fast, moderate, and slow learners. By providing early and actionable insights, this model empowers educators to design tailored mentoring strategies and academic interventions, ultimately improving student success rates and aligning educational outcomes with institutional and industry expectations.

➢ *Research Gap*

Despite technological progress, most current educational systems lack predictive models that integrate AI to provide real-time, personalized learning support. This study addresses this gap by proposing a machine learning-driven framework for predicting student performance and delivering customized academic interventions.

➢ *Objectives*

- Develop an AI-based predictive model to classify students based on performance indicators.
- Evaluate and compare the performance of various machine learning algorithms to determine the most efficient and accurate method.
- Determine key academic, behavioral, and psychological factors influencing student success.
- Generate personalized recommendations to enhance learning outcomes.
- Ensure ethical AI practices through data privacy protection and bias mitigation techniques.
- The proposed system aligns with the Outcome-Based Education (OBE) framework, ensuring alignment between educational practices and industry requirements [20], [21].

✓ *Research Question*:

How can machine learning be leveraged to predict student performance and enable data-driven educational interventions?

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1354

II. LITERATURE REVIEW

The rapid advancement of Artificial Intelligence (AI) has created significant opportunities to revolutionize educational practices, particularly through the development of personalized learning systems. Traditional teaching methods often rely on standardized instructional approaches, which fail to consider the diverse learning styles, abilities, and progress rates of individual students. In contrast, AI-powered tools offer a transformative solution by analyzing large volumes of learner data to personalize content delivery, pacing, and assessment strategies. VanLehn [14] demonstrated that intelligent tutoring systems can dynamically adjust to the needs of individual learners, providing customized feedback and interactive guidance comparable to human tutoring. Likewise, Holmes et al. [15] emphasized how AI enhances student engagement and academic outcomes by fostering adaptive and responsive learning environments.

In the field of machine learning (ML), algorithms have become essential tools for predicting and improving student performance. Among these, decision tree models are highly favored due to their ease of interpretation and ability to manage complex and diverse datasets. Soni and Bajpai [16] demonstrated that algorithms like J48 can reliably predict student success by analyzing key factors such as attendance, academic grades, and behavioral patterns. Building upon this, Fletcher and Zirkle [19] highlighted that ensemble techniques, particularly Random Forest, achieve greater predictive accuracy compared to single decision tree models, especially when dealing with large and diverse student populations. Similarly, Zou et al. [12] proposed a comprehensive ML framework that integrates multiple algorithms to enhance the robustness of predictions, while Ahmed et al. [11] conducted a comparative study of classifiers such as KNN, SVM, and Naïve Bayes, concluding that hybrid models deliver superior performance for educational datasets.

Recent research has increasingly focused on integrating AI-driven analytics into higher education systems. Shoaib et al. [4] developed an AI-based performance prediction model that offers real-time insights to educators, enabling timely interventions for students at risk of academic failure. Similarly, Ahmed et al. [5] explored regression and ensemble approaches to predict academic outcomes, emphasizing the value of continuous, data-driven monitoring. Collectively, these studies demonstrate the growing potential of AI and machine learning (ML) to proactively tackle learning challenges and enhance institutional decision-making processes.

As the use of AI becomes more widespread in educational settings, ethical considerations have emerged as a critical area of concern. Issues related to bias, fairness, and data privacy are now receiving significant attention in contemporary research. Dastin [17] highlighted how poorly designed AI systems can reinforce systemic biases, citing a case where a recruitment algorithm discriminated against women. Likewise, Binns [18] emphasized the need for transparent algorithms and robust ethical frameworks to ensure accountability in AI-driven educational environments. Adewale et al. [8] further examined the broader implications of AI adoption, recommending policies that safeguard student data and promote equitable access to technology.

Despite notable advancements, most existing studies focus on either predictive modeling or adaptive learning systems, with few addressing both components simultaneously. This has resulted in a significant research gap, as effective integration of prediction and real-time personalization remains limited. The present study aims to bridge this gap by introducing a comprehensive framework that combines machine learning-based prediction models with personalized intervention strategies. By utilizing diverse data dimensions—including academic performance, behavioral indicators, and psychological factors—the framework facilitates the early detection of at-risk students and delivers targeted, evidence-based educational support.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1354

III. METHODOLOGY

Fig 2 Model Flowchart

➢ *Dataset and Preprocessing*

The dataset comprised 100 records and 39 attributes across five key dimensions:

- *Academic Records:*
  CGPA, attendance, assignments, lab performance.
- *Behavioral Attributes:*
  Motivation, discipline, stress levels.
- *Skill Assessment:*
  Programming, leadership, communication.
- *Psychological Factors:*
  Anxiety, academic pressure.
- *Lifestyle Factors:*
  Study habits, sleep quality.

➢ *Preprocessing Techniques Included:*

- *Feature Engineering:*
  Selecting performance-related attributes.
- *Normalization and Standardization:*
  Maintaining consistency across data points.
- *Handling Missing Values:*
  Addressing gaps using imputation techniques.
- *Data Anonymization:*
  Protecting student privacy.

➢ *Machine Learning Models*

Five algorithms were evaluated for predictive performance:

Table 1 Predictive Performance

Algorithm	Description
Decision Tree (J48)	Rule-based classification model that splits data into subsets based on attribute values.
Random Forest	Ensemble technique that improves accuracy by combining multiple decision trees.
Naïve Bayes	Probabilistic classifier that assumes independence between features.
SVM (Support Vector Machine)	Margin-based algorithm used to separate data into distinct classes or trends.
KNN (K-Nearest Neighbors)	Distance-based model that classifies data based on the nearest neighbors.

All models were optimized through hyperparameter tuning and evaluated using 10-fold cross-validation to ensure reliable and robust performance.

➢ *Entropy:*

Entropy is a metric used to quantify the level of uncertainty or randomness within a dataset s.

$$\text{Entropy}(s) = -\sum_{i=1}^{c} P_i \log_2 P \quad \text{------------------------------} \quad (1)$$

Where:

- C = Total number of Classes.
- P_i = Proportion of samples belonging to class i in dataset s.

Entropy(S) is maximum when the classes are equally distributed and minimum (0) when all instances belong to one class.

IV. RESULTS DISCUSSIONS

➢ *Performance Metrics*

The performance of the models was evaluated using four key metrics: accuracy, precision, recall, and F1-score.

Table 2 Model Performance

Classifier	Accuracy (%)	Precision	Recall	F1-Score
J48 Decision Tree	100.0	1.000	1.000	1.000
Random Forest	98.5	0.987	0.985	0.986
SVM (Support Vector Machine)	96.7	0.960	0.967	0.963
KNN (k = 5)	94.2	0.935	0.942	0.938
Naïve Bayes	92.3	0.902	0.913	0.907

➢ *Key Findings:*

- J48 Decision Tree outperformed other classifiers with perfect accuracy.
- Random Forest and SVM demonstrated strong performance, suitable for practical applications.
- Naïve Bayes exhibited the lowest performance due to its assumption of feature independence.

➢ *Statistical Analysis*

An ANOVA test revealed a statistically significant difference among classifier performances (p-value = 0.0013). Correlation analysis identified aptitude, attendance, and motivation as strong predictors of success, whereas stress levels negatively impacted outcomes.

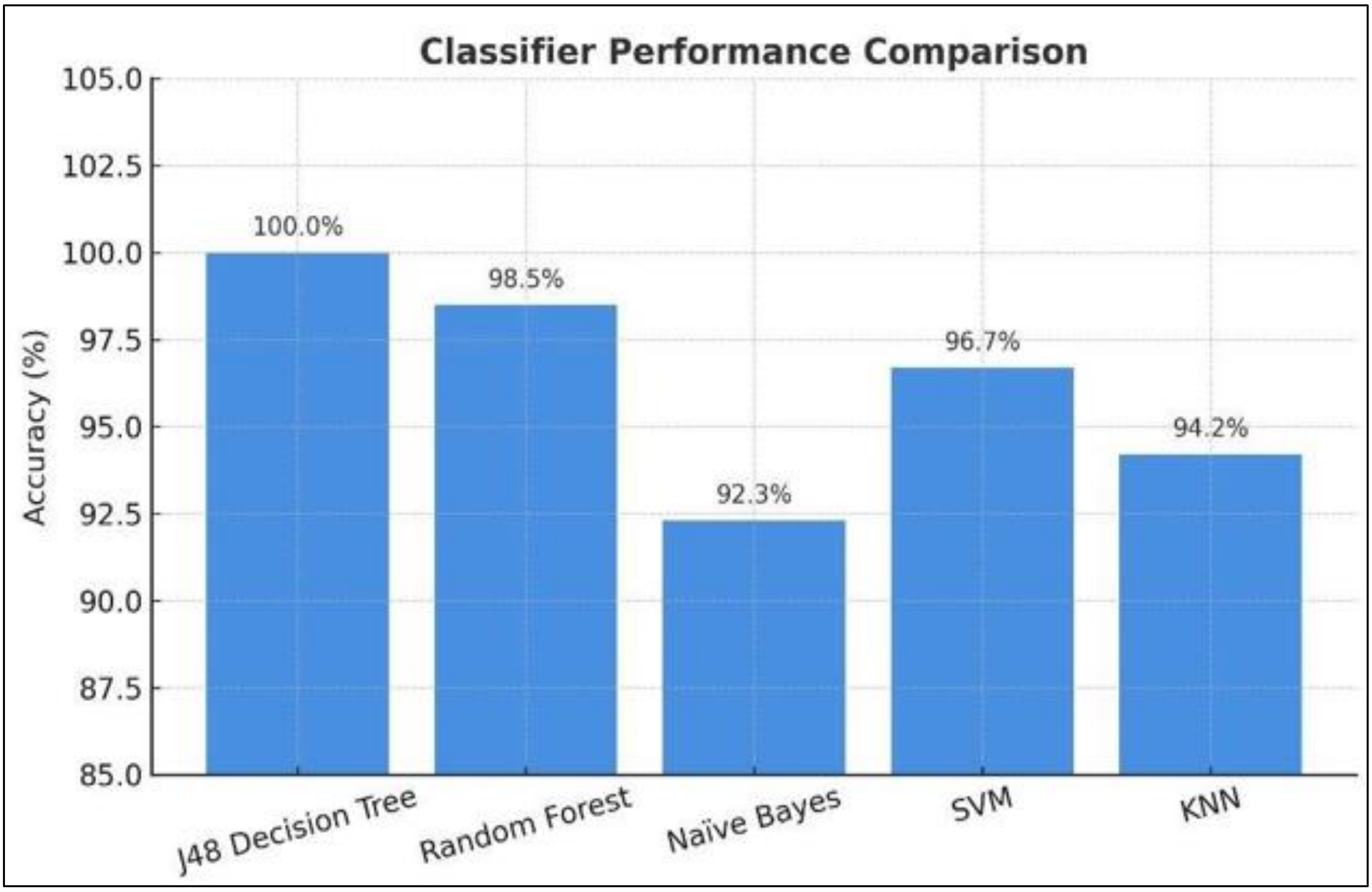

Fig 3 Performance Comparison

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1354

V. CONCLUSION

The proposed framework can be incorporated into academic management systems to proactively identify at-risk students and provide tailored learning support. By utilizing five machine learning algorithms—J48 Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naïve Bayes—on a dataset containing 100 student records with 39 varied attributes, this study developed a comprehensive analytical model to examine the diverse factors influencing student academic performance and overall success.

Among the evaluated models, the J48 Decision Tree achieved perfect classification accuracy (100%), outperforming all other classifiers. Its ability to handle both categorical and continuous data, combined with its interpretability, makes it highly suitable for educational contexts where transparency is essential for decision-making by educators and policymakers. Random Forest and SVM also delivered strong performances with accuracies of 98.5% and 96.7%, respectively, making them reliable alternatives for larger- scale or more complex datasets.

The analysis further identified key predictors of academic success, such as aptitude, attendance, and motivation, which exhibited strong positive correlations with CGPA. Conversely, high stress levels were found to negatively affect student performance. These insights are invaluable for educators seeking to design targeted

interventions, such as personalized mentoring, stress management workshops, and adaptive learning pathways, aimed at addressing the specific needs of students.

The statistical validation through ANOVA testing confirmed that the observed differences in classifier performance were statistically significant, adding robustness to the results. The integration of this framework into existing Outcome-Based Education (OBE) systems can enable educational institutions to align learning processes more closely with industry standards, ensuring graduates are not only academically competent but also employable and future-ready.

By leveraging real-time analytics and predictive modeling, this study emphasizes the transformative role that AI can play in making education more data-driven, equitable, and adaptive. The findings open the door for scalable, intelligent educational systems capable of addressing diverse learning needs in large and heterogeneous classrooms.

- *Future Scope*

Integration of deep learning models for improved prediction accuracy. Deployment of real-time feedback mechanisms for continuous monitoring. Scaling the model to large datasets and diverse institutions. Incorporation of ethical AI audits to ensure fairness and transparency.

REFERENCES

[1]. K. VanLehn, "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems," *Educational Psychologist*, vol. 46, no. 4, pp. 197-221, 2011.

[2]. Holmes, W., Bialik, M., and Fadel, C., "Artificial Intelligence in Education: Promises and Implications for Teaching and Learning," *Center for Curriculum Redesign*, 2019.

[3]. V. K. Vijayakumar and A. Naveen, "Long-Term Effects of AI-Personalized Learning on Engagement and Performance," *International Journal of Environmental Sciences*, pp. 1601-1608, 2025.

[4]. M. Shoaib, et al., "AI Student Success Predictor: Enhancing Personalized Learning Using Advanced Machine Learning Algorithms," *Computers in Human Behavior*, 2024. [Online]. Available: ScienceDirect.

[5]. W. Ahmed, et al., "Machine Learning-Based Academic Performance Prediction," *Scientific Reports*, Nature, 2025.

[6]. V. Matzavela, et al., "Decision Tree Learning Through a Predictive Model for Personalization of Student Academic Performance," *Education and Information Technologies*, vol. 26, pp. 4821–4837, 2021.

[7]. D. K. Kolo, S. A. Adepoju, and J. Alhassan, "A Decision Tree Approach for Predicting Students' Academic Performance," *International Journal of Education and Management Engineering*, vol. 5, no. 5, pp. 1-9, 2015.

[8]. M. D. Adewale, et al., "Impact of Artificial Intelligence Adoption on Students," *Education and Information Technologies*, 2024. [Online]. Available: PMC.

[9]. S. Wang, et al., "Artificial Intelligence in Education: A Systematic Literature Review," *Expert Systems with Applications*, vol. 244, pp. 122823, 2024.

[10]. C. Merino-Campos, "The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review," *Trends in Higher Education*, vol. 4, no. 2, p. 17, 2025.

[11]. E. Ahmed, et al., "Student Performance Prediction Using Machine Learning," *Computational Intelligence and Neuroscience*, 2024.

[12]. W. Zou, et al., "Prediction of Student Academic Performance Utilizing a Practical Framework with Multiple Machine Learning Models," *Applied Sciences*, vol. 15, no. 7, p. 3550, 2025.

[13]. S. Hashemifar, et al., "Personalized Student Knowledge Modeling for Future Material Prediction (KMaP)," *arXiv preprint*, arXiv:2505.14072, 2025.

[14]. UNESCO, "Artificial Intelligence and Education: Guidance for Policy-Makers," *United Nations Educational, Scientific and Cultural Organization*, 2021.

[15]. C. Romero and S. Ventura, "Educational Data Mining and Learning Analytics: An Updated Survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, pp. e1355, 2020.

[16]. Soni, S., and Bajpai, S., "Student Performance Prediction Using Decision Tree and Naïve Bayes,"

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25sep1354

International Journal of Computer Applications, vol. 178, no. 7, pp. 19-24, 2019.
[17]. Dastin, J., "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women," *Reuters*, 2018.
[18]. R. Binns, "Fairness in Machine Learning: Lessons from Political Philosophy," *Proceedings of Machine Learning Research*, vol. 81, pp. 149-159, 2018.
[19]. Fletcher, D., and Zirkle, K., "Using Random Forest for Predicting Student Success in Higher Education," *Journal of Machine Learning in Education*, vol. 5, no. 2, pp. 112-130, 2018.
[20]. India Ministry of Education, "National Education Policy (NEP) 2020," Government of India, 2020.
[21]. S. Hashemifar, et al., "Personalized Student Knowledge Modeling for Future Material Prediction (KMaP)," arXiv preprint, arXiv:2505.14072, 2025.
[22]. Vijayakumar Krishnan, Dr. Naveen A "A Data-Driven Framework for Holistic Student Performance Evaluation and Industry Readiness Using Machine Learning" Iconic Research and Engineering Journals Volume 9 Issue 3 2025 Page 1350-1354.

➢ *About the Author*

Mr. Vijayakumar Krishnan is a research scholar at Don Bosco College (Co-Ed), Yelagiri Hills – 635 853, affiliated with Thiruvalluvar University, Vellore. He possesses over 9 years of experience in computer networking, server and IT administration, complemented by 3 years in academia and 1 year of research experience.

Dr. Naveen A serves as a Research Supervisor at Don Bosco College (Co-Ed), Yelagiri Hills – 635 853. With 7 years of combined academic and research experience, he currently mentors four research scholars under his guidance.