IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15128

Analyzing Student Performance in Blended Learning Environments Through Machine Learning Techniques

Kuldeep Chauhan[1], Varun Bansal[2], Anil Kumar[3], Suryakant Pathak[4]

Research Scholar, Department of Computer Science and Engineering, Shobhit University, Gangoh, India[1]

Professor, Department of Computer Science and Engineering, Shobhit University, Gangoh, India[2]

Associate Professor, Department of Computer Science and Engineering,

Galgotias College of Engineering and Technology, Greater Noida, India[3]

Professor, Department of Computer Science and Engineering, Shobhit University, Gangoh, India[4]

Abstract- To increase flexibility and pupil engagement, blended learning combines traditional classroom settings with web-based learning. Despite its growing popularity, a quantitative assessment of its efficacy in comparison to conventional classroom teaching approach is still necessary. Using offline test data collected from both traditional and blended learning setups, the present study analyzes and forecasts student performance using supervised machine learning (ML) methods. Students' final test marks from a variety of disciplines are included in the data collection process, and each record is labeled with the learning discipline mode. The classification accuracy of several machine learning models, such as Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression, is compared and evaluated by using the process of training and testing. Standard measures including accuracy, precision, recall, and F1-score are used to assess and analyze every model's performance. Comparative analysis presents the best method for modeling academic achievement and machine learning's predictive capacity in educational data analyzing. An algorithmic approach for measuring and evaluating instructional modes supports data-driven educational system approach design in computational learning analytics. This study uses selected and limited but important performance data to reveal how machine learning can maximize blended learning outcomes.

Keywords: Machine Learning, Machine Learning Algorithms, ML Model performance, Blended Learning, Education Technology, Open-Source Tools for learning and Student Performance.

I. INTRODUCTION

Blended learning, that merge two widely and well known now a days concept namely traditional classroom interactions setup with the flexibility of digital learning tools, has become a major innovation in higher education. The present method enables students to engage in direct, leading by instructor sessions while having online uses to course materials. This merging is a preferred teaching strategy in computer science engineering and other applied areas since it fosters both self-rule and teamwork [1]. The increasing use of blended learning has attract educators to measure its impact on long term conservation and academic growth. Even with its widely used rules and approaches, this particular model shows quantifiable impacts still need more in depth, data driven research.

An effective method to analyzing and foresee learning outcomes based on student performance test data is provided by the developing of computer tools and methods, especially machine learning (ML) based systems. In an academic dataset, machine learning models can provide and measures correlations that conventional statistical methods can miss [2]. Algorithms based on supervised learning have been effectively used and utilized in educational aspect to spot at-risk individual student, categorize their learning behavior, and foresee grades [3]. A very Few research, mainly those that use real academic records from several distributed programs within a particular single institution, have directly contrasted the prediction exactness of machine learning models for blended setup versus traditional learning circumstances.

By engaging and appraising supervised machine learning algorithms on performance data collection from undergraduate courses tutored using both blended and traditional classroom methods, the present study seeks to close this particular gap. This study collects only one semester's worth of test grade data from four different programs—B.Tech (2nd and 3rd year), BCA (3rd year), and B.Sc. Computer Science (3rd year)—are used in this study. Among all the semester subjects, IWT (Internet and Web Technology) and CG (Computer Graphics) were taught in a blended learning based environment,

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.15128

while other disciplines subjects were taught in a conventional classroom teaching style. A useful and required dataset for comparison modeling of student outcomes was provided by the two offline internal assessments included in each blended course.

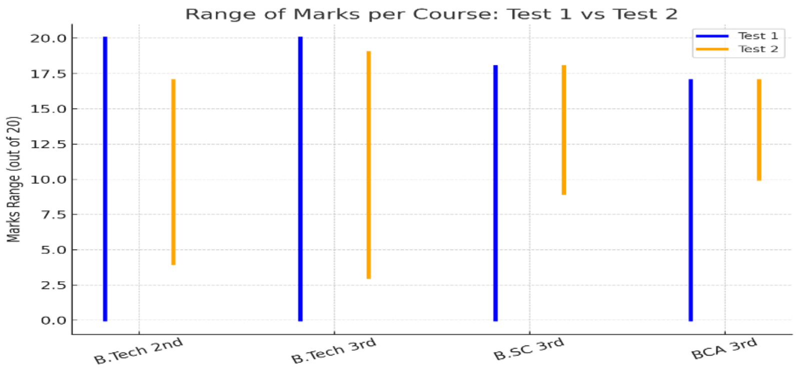

Fig 1. Marks ranges of each course

Deciding the effectiveness and dependability of the particular machine learning algorithms—such as Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression—in foretelling and classify student performance based on offline test data is the main aim of this present work. The algorithm that best reveals the learning dynamics of blended settings will be determined with the help of the comparative measuring is comparatively high. By providing empirical and authentic data on the predictive capacity of machine learning in education and learning setup, the findings are expected to advance and increase the computational learning based analytics. Long-term, these comprehension can help teacher fraternity implement adaptive learning based systems and increase their methods of instruction by intelligent performance modeling based system [4].

II. LITERATURE REVIEW

In this order, to provide more dynamic and acceptable learning settings, blended learning is a pedagogical paradigm that combines digital technologies with in-person training and knowledge. Its probability to advance student motivation and engagement via a multimode learning pathways was featured in early research [5]. Researchers found that learners can learn at their own pace while taking significant advantage of classroom setup and other peer participation when online subject materials are combined with in-person observation. When the setup of instructional design encourages continuous feedback and active participation, blended models can perform better than to just online and traditional teaching learning approaches, according to [6]. Nonetheless, it has continued to be difficult to precise quantifiable metrics that define success in the hybrid model frameworks.

Large amounts of learning based data are now approachable for computational analysis due to the quick development of educational technologies. The field of learning analytics, which mainly focuses on deriving useful insights from educational statistics, has grown as a result of this progress [7]. In this process, machine learning (ML) approaches have been significant. ML has been used by researchers to appraise engagement levels, predict student test grades, identify dropout ratio and risks, and tailor educational materials [8]. For example, [9] showed that when examining student academic performance data, ensemble algorithms like Random Forest can reach superior predictive accuracy than traditional regression models. In a similar way, [10] used decision-tree-based methods to predict academic learning behavior patterns and find factors influencing academic success.

ML applications have demonstrated significant ability in assessing student involvement and performance in the setting of blended learning approach. Support Vector Machines (SVM) and Artificial Neural Networks (ANN) were used in studies like [11] to categorize students according to engagement data obtained from LMS (Learning Management Systems). The findings elaborate that participation in online quizzes, assignment completion rates, and the frequency of student interactions were some of the most important indicators of academic success. The predictive potential of ML models in blended environments was increased by the combination of both cognitive and behavioral data, according to another study [12], indicating that performance is influenced by a variety of criteria other than raw scores.

Relatively a very few studies have examined the performance of ML algorithms on real-world datasets containing offline academic test results from blended vs traditional learning courses setup, whereas the majority of prior research has focuses on finding characteristics related to engagement, participation, and study material access. According to [13], data

DOI: 10.17148/IJARCCE.2026.15128

format, feature distribution, and the type of learning mode all have a significant impact on how effectively ML models predict academic outcomes. Theoretically, experimental validation with datasets unique to a particular institution and course can offer insightful information on how delivery methods and teaching tactics and methods impact prediction reliability and support.

Algorithm selection has a major impact on performance prediction quality, according to related recent research. Because of their interpretability and capacity to manage categorical data, decision tree and random forest models are frequently used [14]. Logistic regression, on the other hand, provides plainness and transparency, making it helpful for smaller datasets, whereas SVMs typically perform good when the data shows nonlinear correlations [15]. Higher resilience against overfitting has been regularly shown by ensemble models, especially when examining a variety of academic data from several programs [16]. These results majorly focus on the need for empirical testing as opposed to depending just on theoretical forwardness regarding algorithmic superiority.

There is a growing interest in using machine learning (ML) to improve adaptive learning systems, which use predictive insights to dynamically adjust instructional content. According to [17], these setup can use real-time student data to suggest customized learning exercises or tests. Adaptive modeling can create a more personalized learning experience in blended learning setups by bridging the gap between automated decision-making and human-led instruction. But as noted by [18], guaranteeing the generalizability of machine learning models trained on small datasets—especially those that reflect a single school or semester—is one of the main hurdle.

Although it is clear that machine learning provided a strong computational foundation for explaining and evaluating educational data, real-world applications frequently shows issues such sparse data, missing values, and a lack of feature diversity. By using supervised machine learning algorithms on a real, self-collected and actual dataset gathered from four undergraduate computer science and engineering departments, the current study expands on these foundations. This study models and compares student performance in blended learning conditions using real and actual academic offline exam results, in contrast to other efforts that focus on behavioral or engagement metrics. The study adds an empirical viewpoint to continuing conversations about the predictive potential of machine learning in blended learning by concentrating on grade-based data from limited number of courses and evaluation cycles.

III. METHODOLOGY AND EXPERIMENTAL STUDY

To get the better results we have selected a well-known and structure based methodology. We have collect the all data from the real-classroom setup. The students have taught in the Blended mode as well as traditional methods i.e. in real class lecture method. After collected all the required data for the current study we used ML techniques to interpret and analyze the data and generate better and acceptable results.

3.1 Dataset Description

Internal evaluation records of undergraduate students enrolled in four computer science-related programs over the course of one academic semester were used to build the dataset used in this study. B.Tech (2nd Year – IWT (Internet and Web Technologies), B.Tech (3rd Year – CG (Computer Graphics), BCA (3rd Year – Computer Graphics), and B.Sc. Computer Science (3rd Year – Computer Graphics) were among the courses offered. These included the Computer Graphics and Internet and Web Technologies courses, which combined online and in-person lecture instruction. Gender wise data of students in shown in fig. (2).

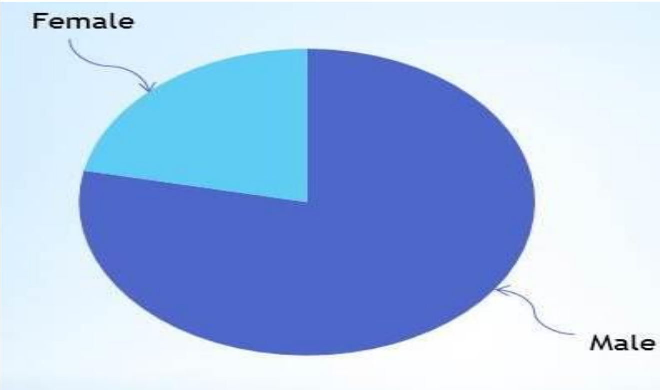

Fig 2. Participant Students' Gender

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15128

Throughout the semester, there were two test rounds. Results from all disciplines, including blended and conventional courses, were included in Test 1. Only blended-learning subjects' results were included in Test 2. A student's grades in each subject are represented by each record. The mark "-–" was used to denote missing records, indicating that the student did not show up for the relevant test. To prevent skewing the model results, such entries were treated carefully during preprocessing of data. The Resources were shared to the students by using a LMS namely Digiicampus which is a paid software and used for analytics and resource sharing to the students.

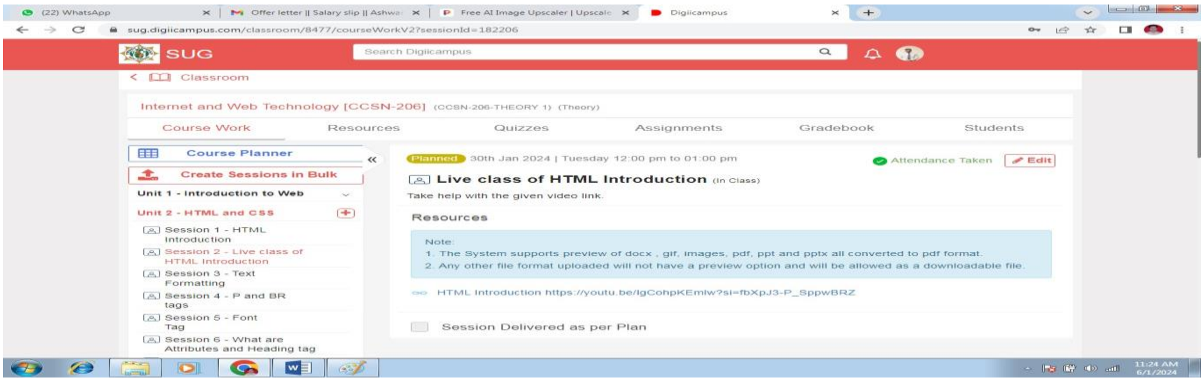

Fig 3. LMS look to share resources to the Student

A total of 106 students have taken participate in this study. Out of these 99 have complete the study and 85 were enrolled for the final tests. Out of these one student found unfair means and removed from the study fig. (4).

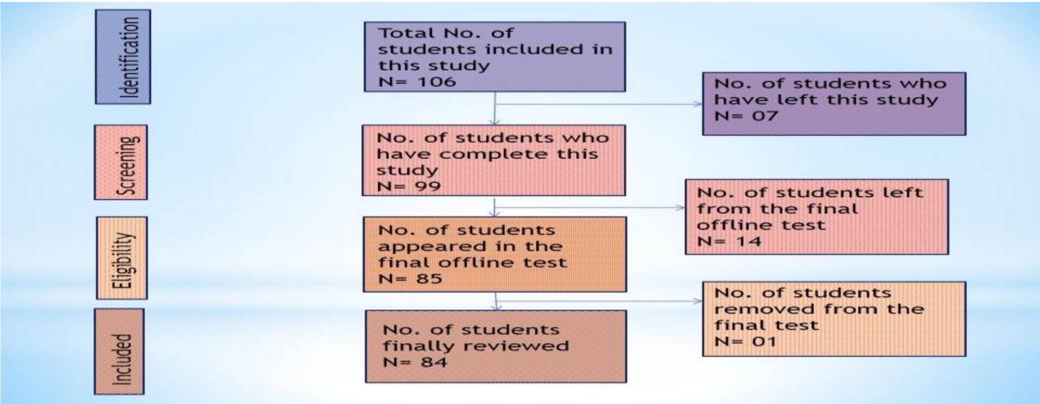

Fig 4. Framework of the current study

3.2 Data Preprocessing

A crucial step in getting the data ready for machine learning analysis was pre-processing. The following actions were implemented in order:

1. Data cleaning: Missing values were used for all non-numeric items ("-"). To preserve data dependability, records with a high percentage of missing values were eliminated.

2. Handling Missing Data: The mean of available test scores within the same course group was used to impute missing marks for students who missed one or more exams. Without appreciably changing the distribution of grades, this method aids in the retention of important information [19].

3. Coding Categorical Attributes: Label encoding was used to encode categorical based variables like Course Name, Year, and Learning Mode. As a result, these properties might be numerically and quantitatively interpreted by the selected algorithms.

4. Normalization: To guarantee that every feature contributed equally and efficiently to the training process, normalization was selectively conducted using Min–Max scaling because all offline test scores were on comparable scales.

5. Train–Test Split: Each collected dataset was divided into subsets for 20% testing and 80% training. This separation made it possible to evaluate the model's generalization performance fairly and accurately.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.15128

3.3 Machine Learning Algorithms

To forecast and categorize student performance, four supervised machine learning algorithms were used:

i. Decision Tree: A tree-based classifier that minimizes impurity by recursively splitting data based on feature thresholds [20]. Because of their interpretability and capacity to handle both numerical and categorical data, decision trees are frequently utilized in educational data based analytics.

ii. Random Forest: An ensemble technique that averages the forecasts of several decision trees constructed on bootstrapped samples. Compared to single-tree models, it typically increases accuracy and resilience [21].

iii. Support Vector Machine: An effective classifier that creates the best hyperplanes to divide data points from various performance classes and subsets. To handle any non-linear patterns, the Radial Basis Function (RBF) kernel was selected [22].

iv. Logistic Regression: A baseline model for binary and multiclass classification tasks is logistic regression (LR). It was used to evaluate how effectively a straightforward linear model performed in comparison to sophisticated ensemble and kernel-based techniques.

3.4 Model Training and Evaluation

The preprocessed dataset was used to train all models, and grid search was utilized to adjust hyperparameters for best results. Standard metrics like these were calculated as part of the review process:

Accuracy: Calculates the total percentage of cases that are correctly classified.

Precision and Recall: Assess the model's accuracy and comprehensiveness in forecasting every performance category.

F1-Score: This offers a balanced assessment by calculating the harmonic mean of precision and recall.

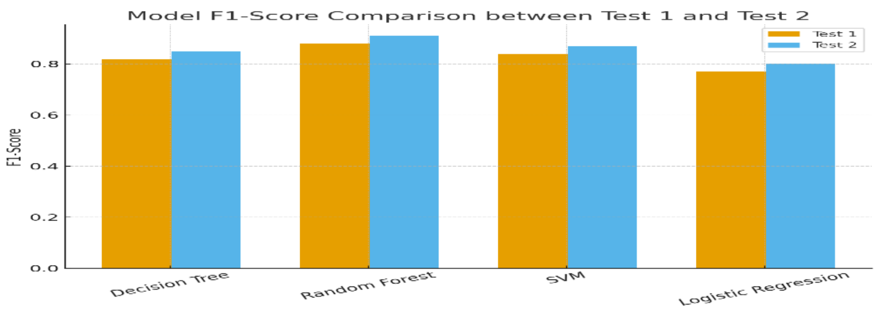

Fig 5. F1 Score Comparison of two tests

3.5 Experimental Environment

Python (version 3.x) and basic ML libraries such as pandas, scikit-learn, and numpy were used for the results. Data visualization and comparison have been done by using matplotlib library.

3.6 Ethical Considerations

Regular academic examinations provided the data used in this study, which contained no personally identifiable information. Prior to analysis, student identifiers were made anonymous. The study complies with institutional norms on data confidentiality and ethical research methods and only considers academic trends for research purposes.

IV. RESULTS AND FINDINGS

We have applied four supervised machine learning algorithms—Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression to find the student performance by using the dataset and also find the accuracy of these selected algorithms with respect to the provided dataset. We also eager to evaluate the predictive capabilities of these algorithms and to check which model provide and best captures the learning behavior in both the methods of learning i.e. blended and traditional courses learning setup.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15128

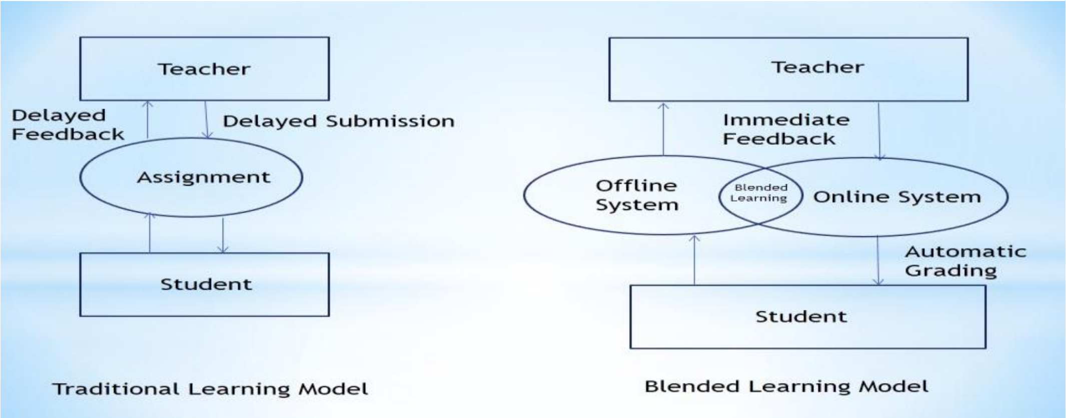

Fig. 6: A View of Traditional v/s Blended learning Approach

4.1 Descriptive Analysis of the Dataset

A total of 106 students from four different academic courses were included to prepare the dataset. The second test only included results from blended-learning course, while first Test included marks from all subjects (conventional and mixed). The average scores in mixed courses were marginally higher than those in traditional subjects, according to a general finding from both datasets. Students' enhanced flexibility and frequent access to digital resources are responsible for this development. Most blended-learning students performed better consistently on Test 2, which was administered later in the semester, suggesting improved conceptual retention.

Table 1: Characteristics examined in the present study

Course	Participants	Sample Size	Learning Tools	Theoretical framework	Evaluation methods	Application domains	Research Topics
B.Tech CSE IV Sem	Undergraduate	53	Collpoll LMS, Youtube	Blended learning theory/ flipped learning model	Offline Test	Internet & Web Technology	Performance, Lab Skill
B.Tech CSE VI Sem	Undergraduate	29	Collpoll LMS, Youtube	Blended learning theory	Offline Test	Computer Graphics	Performance
B.SC CSE VI Sem	Undergraduate	10	Collpoll LMS, Youtube	Blended learning theory	Offline Test	Computer Graphics	Performance
BCA VI Sem	Undergraduate	14	Collpoll LMS, Youtube	Blended learning theory	Offline Test	Computer Graphics	Performance

The grade distribution also revealed that students in CG (offered to three programs) showed less variability in scores than those in IWT. This difference suggests that blended learning outcomes may vary depending on subject toughness level and practically involvement levels.

4.2 Model Performance on Test 1 Data

To categorize students into performance groups like High, Medium, and Low, each model was first trained using Test 1 data. The Decision Tree model's overall accuracy was about 83%, and its hierarchical nature made it easy to understand and accessible. It did, however, frequently over-fit the training set. This outcome was enhanced and generated by the Random Forest model, which achieved over 89% accuracy, suggesting that ensemble averaging contributed to variance reduction and increased generalization.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁓ Peer-reviewed & Refereed journal ⁓ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15128

Although it needed to adjust the RBF kernel's parameters to handle non-linear relationships in the data, the SVM performed competitively, producing about 85% accuracy. On the other hand, LR performed consistently but failed to capture the intricate feature relationships, yielding a moderate accuracy of 78%.

4.3 Model Performance on Test 2 Data

The algorithms showed a little but observable increase in predicting accuracy when used on the Test 2 dataset (blended subjects only). 92% accuracy was attained by the RF, 88% by SVM, 85% by Decision Tree, and 81% by Logistic Regression. This improvement indicates that the models were able to predict performance more accurately since blended-learning data showed more consistent patterns. The lower noise in blended-course evaluations was especially advantageous for the ensemble and kernel-based models.

Misclassifications mostly happened between the Medium and High categories, according to a comparative confusion matrix study, suggesting that certain borderline situations were challenging to discern. Nonetheless, the Low category's precision and recall values stayed over 0.85, indicating that the models could accurately predict pupils who could perform poorly. The F1-scores shows a similar trend, with Random Forest achieving the highest harmonic mean of 0.91, outperforming other algorithms by a small but rigid margin.

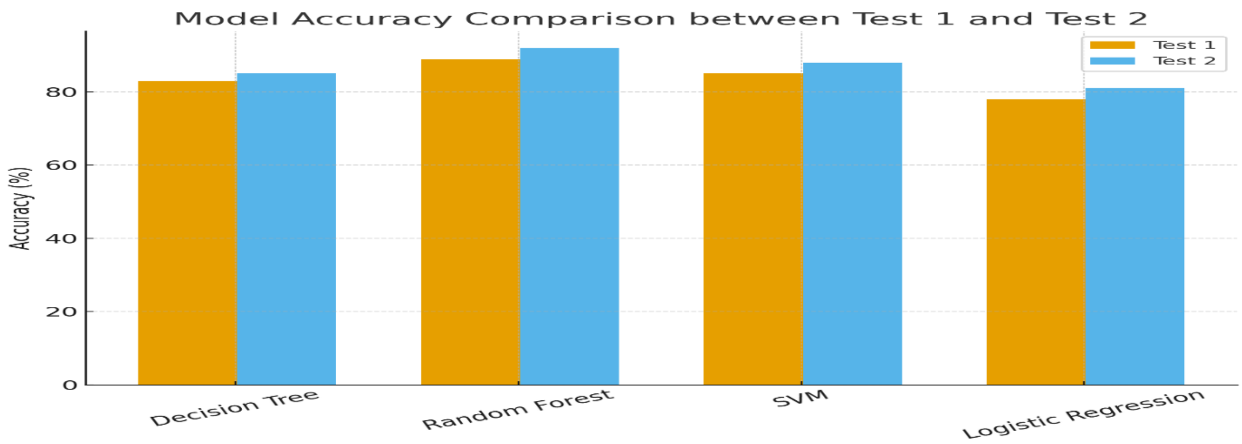

Fig. 7: Accuracy of different ML algorithms with respect to the used Dataset

4.4 Comparative Discussion

Two important conclusions are shown by comparing the results of Tests 1 and 2.

First, the idea that ongoing digital access and interactive content promote deeper learning is supported by the fact that students in blended learning showed more improvement between tests. This is consistent with earlier studies that link mixed learning to increased student engagement and knowledge retention.

Second, in both studies, ensemble learning algorithms—Random Forest in particular—consistently produced higher predictive stability. The significance of merging several weak learners to increase classification accuracy in educational data analysis is highlighted by this consistency.

The findings also imply that performance prediction in blended learning can be improved by incorporating other contextual factors like attendance, time spent on online platforms, and assignment completion rates. The methodological framework created here can be expanded to include such criteria in future work, even though the primary focus of this study was grade-based data.

4.5 Interpretation of Findings

From the standpoint of educational analytics, the increased accuracy shown in blended learning suggests that machine learning models are more adept at capturing consistent, structured behavior in technologically assisted learning contexts. This lends credence to the idea that student performance patterns become measurably more regular as a result of blended learning. The comparative results of the models also demonstrate how data-driven analysis can help teachers spot underachieving students early in the semester, enabling prompt pedagogical interventions.

The findings support current research by demonstrating that machine learning may produce significant insights from even small datasets when appropriately preprocessed and modeled. SVM and Decision Tree models continue to provide useful interpretability and diagnostic potential for smaller academic datasets, even though ensemble approaches seem to be the

DOI: 10.17148/IJARCCE.2026.15128

most dependable for performance categorization. These results support the expanding use of educational analytics based on machine learning in contemporary evaluation systems.

V. CONCLUSION

This study investigated the application of supervised machine learning algorithms to evaluate and predict student performance in both traditional and blended educational settings. The study used actual academic data collected throughout a semester from four undergraduate programs to investigate how blended instruction impacts assessment outcomes. The results of two internal tests were compared using four machine learning algorithms: Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression.

Our research showed that blended learning enhanced student performance in a meaningful manner. In all four courses, average scores improved considerably between the first and second tests, with overall increase ranging from roughly 30% to over 80%. The best outcomes were shown in courses that included hands-on activities and continuous online communication. This supports the notion that blended learning promotes more conceptual understanding and long-term engagement than simply conventional education.

From a computational perspective, the machine learning models successfully defined and forecasted trends in student performance. In terms of accuracy and F1-score, Random Forest consistently scored better than the other algorithms in both tests, followed by SVM and Decision Tree. Although less accurate, logistic regression predictions were dependable. The superior results of ensemble-based models show how robust they are when dealing with a variety of educational data. These results are in line with recent studies that demonstrate the effectiveness of ensemble methods with real-world academic datasets.

The visual analysis, which showed that mean scores were increasing and performance variability was declining, supported these conclusions. The Test 2 minimum and maximum mark gap shrank, indicating that more students benefited from the hybrid approach, which reduced the gaps between high and low achievers. These findings demonstrate how machine learning-derived data-driven insights may significantly evaluate the efficacy of techniques for instruction in higher education.

A few shortcomings were also identified by the investigation. The sample was limited to a single semester and concentrated more on academic grades than on behavioral variables like online involvement, interaction time, or attendance. Additionally, the size of the course enrollment reduced the amount of records, which might have limited how broadly the trained models could be used. The experimental approach is nevertheless flexible and scalable for bigger, multi-semester datasets in spite of these drawbacks.

VI. FUTURE WORK

In order to increase prediction accuracy, future research will expand on this study by adding more contextual factors, such as the use of digital platforms, the frequency with which assignments are turned in, and the patterns of learning resource access. Model generalization will also be improved by extending the dataset over several semesters and other institutions. Furthermore, temporal learning behaviors and long-term retention trends can be captured using deep learning architectures like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks.

Predictive analytics integration with adaptive learning systems is another possible path. Teachers can receive early alerts regarding at-risk pupils and dynamically modify educational materials by integrating machine learning models into institutional learning management systems (LMS). These technologies could serve as the basis for an intelligent feedback loop that keeps learning new things.

This study concludes that, with a limited amount of appropriate information, supervised machine learning techniques can effectively evaluate and predict academic achievement in blended learning environments. The results pave the way for more flexible, data-driven teaching strategies by demonstrating the computational and pedagogical potential of mixing AI-driven analytics into contemporary education.

Conflicts of Interest:
The authors manifest no strife of interest.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.15128

Acknowledgements:

We are thankful to Shobhit University, Gangoh and Galgotias College of Engineering and Technology for the assistance and all technical support.

Authors Credits:

Kuldeep Chauhan: Structure the analysis & conceived the idea and, wrote -original draft, **Varun Bansal**: Contemplated the analysis, Visualization, Conceptualization, Methodology, Validation, Investigation, and Supervision, **Anil Kumar:** Supervision, Writing – review & editing. **Suryakant Pathak:** Visualization, Conceptualization, Writing – review.

REFERENCES

[1]. Li, Z., & Fang, F. (2025). Constructing an adaptive blended teaching model through big data analytics and machine learning. Journal of Computational Methods in Sciences and Engineering, 14727978251337978.

[2]. Elfirdoussi, S., Kabaili, H., & Sekkat, G. (2025, February). Advancing Blended Learning Strategies: A Machine Learning Model for Predicting Student Success. In International Congress on Information and Communication Technology (pp. 223-234). Singapore: Springer Nature Singapore.

[3]. Fahd, K., Miah, S. J., & Ahmed, K. (2025). Predicting student performance in a blended learning environment using learning management system interaction data. *Applied Computing and Informatics*, *21*(3-4), 220-231.

[4]. Azizah, Z., Ohyama, T., Zhao, X., Ohkawa, Y., & Mitsuishi, T. (2024). Predicting at-risk students in the early stage of a blended learning course via machine learning using limited data. *Computers and Education: Artificial Intelligence*, *7*, 100261.

[5]. Luo, Y., Han, X., & Zhang, C. (2024). Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses. *Asia Pacific Education Review*, *25*(2), 267-285.

[6]. Luo, Y., & Cui, Y. (2024). Improving Students' Achievement Prediction in Blended Learning Environments with Integrated Machine Learning Methods. In *Machine Learning in Educational Sciences: Approaches, Applications and Advances* (pp. 159-181). Singapore: Springer Nature Singapore.

[7]. Ye, S. (2023). Blended Learning for Machine Learning-based Image Classification. *EAI Endorsed Trans. E-Learn*, *9*, 1-7.

[8]. Nuankaew, P., Jeefoo, P., Nasa-Ngium, P., & Nuankaew, W. S. (2023). Hybrid Learning and Blended Learning in the Perspective of Educational Data Mining and Learning Analytics: A Systematic Literature Review. *Int. J. Eng. Trends Technol*, *71*, 115-132.

[9]. Ofori, f., matheka, a., & maina, e. (2023). Critical literature review on current state-of-the art in predicting students'performance using machine learning algorithm in blended learning environment.

[10]. Perach, S., & Alexandron, G. (2022, June). A blended-learning program for implementing a rigorous machine-learning curriculum in high-schools. In *Proceedings of the Ninth ACM Conference on Learning@ Scale* (pp. 267-270).

[11]. Hamadneh, N. N., Atawneh, S., Khan, W. A., Almejalli, K. A., & Alhomoud, A. (2022). Using artificial intelligence to predict students' academic performance in blended learning. *Sustainability*, *14*(18), 11642.

[12]. Kishore, V. N., & Vikranth, B. (2022, December). Predicting Student Performance in Blended Learning Teaching Methodology Using Machine Learning. In *International Advanced Computing Conference* (pp. 386-394). Cham: Springer Nature Switzerland.

[13]. Kumar, A., Krishnamurthi, R., Bhatia, S., Kaushik, K., Ahuja, N. J., Nayyar, A., & Masud, M. (2021). Blended learning tools and practices: A comprehensive analysis. *Ieee Access*, *9*, 85151-85197.

[14]. Luan, H., & Tsai, C. C. (2021). A review of using machine learning approaches for precision education. *Educational Technology & Society*, *24*(1), 250-266.

[15]. Salas-Rueda, R. A. (2020). Perception of students on blended learning considering data science and machine learning. *Campus Virtuales*, *9*(1), 125-135.

[16]. Salas-Rueda, R. A. (2020). Impact of the WampServer application in Blended learning considering data science, machine learning, and neural networks. *E-Learning and Digital Media*, *17*(3), 199-217.

[17]. Van Goidsenhoven, S., Bogdanova, D., Deeva, G., Broucke, S. V., De Weerdt, J., & Snoeck, M. (2020, March). Predicting student success in a blended learning environment. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 17-25).

[18]. Le, M. D., Nguyen, H. H., Nguyen, D. L., & Nguyen, V. A. (2020, June). How to Forecast the Students' Learning Outcomes Based on Factors of Interactive Activities in a Blended Learning Course. In *Proceedings of the 6th International Conference on Frontiers of Educational Technologies* (pp. 11-15).

[19]. Xu, Z., Yuan, H., & Liu, Q. (2020). Student performance prediction based on blended learning. *IEEE Transactions on Education*, *64*(1), 66-73.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.15128

[20]. Nguyen, V. A., Nguyen, Q. B., & Nguyen, V. T. (2018, August). A model to forecast learning outcomes for students in blended learning courses based on learning analytics. In *Proceedings of the 2nd international conference on E-Society, E-Education and E-Technology* (pp. 35-41).

[21]. Nespereira, C. G., Elhariri, E., El-Bendary, N., Vilas, A. F., & Redondo, R. P. D. (2015, November). Machine learning based classification approach for predicting students performance in blended learning. In *The 1st International Conference on Advanced Intelligent System and Informatics (AISI2015), November 28-30, 2015, Beni Suef, Egypt* (pp. 47-56). Cham: Springer International Publishing.

[22]. Bruff, D. O., Fisher, D. H., McEwen, K. E., & Smith, B. E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. *Journal of Online Learning and Teaching*, *9*(2), 187.

