

A Comprehensive Review of Machine Learning Algorithms Applied in Blended Learning Environments

Kuldeep Chauhan¹, Varun Bansal², Suryakant Pathak³, Anil Kumar⁴
^{1,2,3} Shobhit University, Gangoh

⁴Galgotias College of Engineering and Technology, Greater Noida

(E-mail: kuldeep.chauhan@shobhituniversity.ac.in, varun.bansal@shobhituniversity.ac.in,
skpathak.cs@gmail.com, vats.anilkumar@gmail.com)

Abstract—The usage of technology has an impact on all aspect of our life, including education. Blended learning has proven as a cutting-edge procedure to education that offers both professors and students new ways to study and teach. Blended-Learning (BL), which integrate in person traditional education with internet components, has become the new standard. This study examines 15 articles that were released between 2021 and 2024 and found in the Scopus, IEEE Access, Springer, and ResearchGate databases. According to the review, the majority of researchers concentrated on measuring student outcomes and learning behaviors utilizing any accessible educational datasets, machine learning algorithms, or appropriate technologies. However, only a small percentage of authors have employed machine learning approaches, and some researchers have not used any machine learning algorithms. We discovered that the authors of these studies are just interested for the end result and have not considered the factors that may influence the learning results. In order to work on the pre-find and selected novel parameters, this research presents the efficient use of machine-learning algorithms with blended learning data and novel parameters. We are the first to look for novel parameters that can affect the students' achievements in blended-learning environments.

Keywords—Blended Learning; Machine Learning; New age Technology; Novel Parameters; ICT Tools

I. INTRODUCTION

This is the era for advancement in every field where technology affecting everyone everywhere. In education, Blended Learning is that technology which affecting and changing the way of learning now a days. Blended-Learning merge the features of traditional face-to-face learning and online learning. Learning is the process of gaining knowledge, abilities, memories, and even wisdom via education and experiences. Modern education has been revolutionized by higher learning through internet technologies, and many universities now offer online courses in addition to entirely online programs [2]. For post-secondary education, especially for master's courses especially technical studies, the Blended-Learning (BL) with Flipped-Classroom (FC) approach may be a highly helpful teaching and learning strategy. Students and faculties have a variety of options and e-learning resources at

their disposal to plan these kinds of lectures, classes, and courses [10]. In order to help teachers identify underperforming students and raise their academic performance, it is critical to forecast and analyze student-performance in a blended-learning domain. In the meantime, choosing machine learning strategies that can yield the best score is necessary to achieve accurate forecasts [8]. The most recent machine-learning and data analyzing technologies can handle the kinds of data, like student and course information, and the connections among them, that call for big data management. Utilizing the streamlined computational-efficiency for various internet-learning model types is another overhead; machine-learning approach can also be used to tune the performance of e-learning models [14]. The application of machine learning in education has opened new avenues for automating and improving the structure of blended learning environments. A recent methodological approach focuses on distinguishing between theoretical instruction and hands-on laboratory practice through intelligent classification. This approach integrates a set of well-established machine-learning algorithms—namely Random-Forest (RF), SVM, Logistic-Regression (LR), and Decision-Tree—to build a classification framework tailored to hybrid learning contexts. By using these to identify theory based face to face sessions, SVM shows the highest accuracy of approximately 93%, and other side LR maintained a baseline performance close to 90%. When applied to detect practical lab sessions, models such as XGBoost, RF, and SVM achieved consistent accuracies near 80%. These outcomes shows the likely of machine learning methods to correctly recognize and categorize different learning styles, contributing notably to the automation and scalability of blended learning systems [1]. To understand the learning behaviors better in blended learning contexts, recent research has examined motivational factors like time-management and the frequency of student engagement levels. The evaluation process becomes recurring and adaptive with the addition of a dynamic window expansion mechanism, enabling more responsive analysis and prompt pedagogical involvement. This adaptive technique has proven to be applicable across different academic cohorts and learner groups, as evidenced by its congruence with current performance benchmarks. Particularly, this approach has a lot of dedication for early recognition of students who might be at risk of performing slower in blended learning courses, which would allow institutions to provide active support and increasing learning results [6].

II. LITERATURE REVIEW

Recently, the direction of educational systems distribute, monitor, and personalize training has significantly changes as a result of the association of machine-learning techniques into blended learning settings. Machine-learning has transform a key factor for assessing complicated student behaviors, forecasting academic achievement, and suggesting adaptive content as educational institutions increasingly embrace hybrid models that blend online and in-person components. Recent research has hold more dynamic and real-time method using supervised and unsupervised machine learning algorithms, whereas early applications mainly focus on static

data analysis, for example attendance or assessment scores. Now a days, activities like student categories, engagement tracking, dropout measures, and select the learning path creation are automated using methods like decision-trees, SVM, neural-networks, and many other models. These developments not only improve the effectiveness of instruction but also help to identify slow performers early on, allowing for prompt interventions. Consequently, machine learning is changing how intelligent blended learning systems are designed and implemented by providing data-driven insights that benefit teachers and students alike. This combination of technologies has been discussed as Table 1.

Table 1.

<u>Ref. No.</u>	<u>Methodology / Tools</u>	<u>Objective</u>	<u>Research Gap</u>	<u>Key Contribution</u>	<u>Publisher</u>
1	Shapley Additive Explanation (SHAP): An AI approach	TCLPI: Machine Learning-Driven Framework for Hybrid Learning Mode Identification	The research is based on student opinions, not conduct any practical on the students.	This Research find that blended-learning works more effectively for theoretical classes; it is less needed for practical classes for students.	IEEE Access
2	LMS with CART and regression-analysis models,	Analyzing Teaching Effects of Blended Learning with LMS: An Empirical Investigation	The research is conducted only on the non-technical course i.e. English and not use the resulted data analysis into instructional practices.	Check the learner's behavior and study habits	IEEE Access
3	Used Systematic-Reviews and Meta Analyses i.e. (PRISMA) method	Systematic Review of Self-Regulated Learning with Blended Learning in Digital Space	Focus on self-regulated-learning not used any ML techniques.	This study find the life-long impact of SRL and how various strategies strike the learning-performance.	IEEE Access
4	Five ML classifiers are used	The application of machine learning algorithms in a blended learning environment to predict student academic performance	Focus on the student performance and final grades of the external dataset of student.	Create a tailored model that accurately forecasts student performance in the blended-learning conditions and measure the performance of different ML algorithms.	National e-science, WITS University
5	Case Study	How to Plan and Manage a Blended Learning Course Module Using Generative Artificial Intelligence?	Only provides guideline to assist the Teachers not worked with data.	Provides practical guidelines to integrate the AI tools with Blended Learning	Springer Nature
6	Random Forest algorithm and SHAP tool	Predicting at-risk students in the early stage of a blended learning course via machine learning using limited data	Work with the data available online and not focused on the novel parameters which can affect the performance	Identify slow performing students in a blended-learning course domain at early stage.	Elsevier
7	ML Algorithms: CNNs	Blended Learning for Machine Learning-based Image Classification	Follow the educational approach	A specific emphasis on ML-based image classification	ICST Transaction e-Education and e-Learning
8	ML Algorithm and Surveys	Critical literature review on current state-of-the art in Predicting students' performance using machine learning Algorithm in blended learning environment	Work only on the final outcomes i.e. academic score	Predict student's performance in blended learning domain	African Journal of Emerging Issues (AJOEI)
9	Used a bDBL approach	Developing a Framework for Blended Design-Based Learning in a First-Year Multidisciplinary Design Course	Use the concept only on practical based learning	Proposed the concept of Blended Design-Based Learning and evaluate design competencies.	IEEE Transactions
10	Case Study	Interactivity – A Key Element of Blended Learning with Flipped Classroom Approach	Use only the Flipped classroom Method of teaching and measure performance.	A project-based and project-oriented learning approach is proposed.	IEEE Conference Paper (ICELIE)

11	Delphi study, FUZZY TOPSIS	A Toolkit to Support the Design of Blended Learning Courses	Focused on the course content delivery	Proposed a toolkit to assist in designing the blended learning courses	IEEE Access
12	Use theoretical and expert knowledge	A Blended Learning Model Based on Smart Learning Environment to Improve College Students' Information Literacy	Emphasize on the information literacy by content delivery, no algorithm is used.	This Paper proposed an information knowledge improvement model to the college students based CIAP	IEEE Access
13	Comprehensive Surveys and Analysis	Blended Learning Tools and Practices: A Comprehensive Analysis	Use traditional methods only not used ML models.	This paper regulate surveys based on various techniques, frameworks and models	IEEE Access
14	Different ML algorithms	Feature Evaluation of Emerging E-Learning Systems Using Machine Learning: An Extensive Survey	Use ML techniques but only with the E-Learning not with Blended Learning	This study surveillance the area of internet Learning with Machine-Learning.	IEEE Access
15	MOOCs platform, online learning tools	An Approach for Scaffolding Students Peer-Learning Self-Regulation Strategy in the Online Component of Blended Learning	Focused only on self-regulated-learning and peer learning	Proposed an approach for scaffolding students peer-learning self-regulation strategy in the online component of blended learning	IEEE Access

An overview of this trend has been given by a number of research. For example, Chaman Verma et al. [1], Zahra Azizah et al.[6] has used a SHAP tool of AI technology and predict the performance at early stage so that a student can take precautionary major to improve his/her grades and also suggest that this model is very best suitable for the theory courses as compare to the laboratory subjects.

More pure machine learning framework was suggested by Boitumelo Makgoba et al. [4], Shengpei Ye [7], Francis Ofori [8], Shabnam Mohamed Aslam et al. [14]. These researches suggest the effective utilization of ML algorithms to produce the effective result which can guide the students towards their success for a particular subject.

LMS (Learning Management System) and MOOCs online platforms plays a very significant role in this aspect and help to get more relevant data to process. Yiyun Wang et al. [2], Rasheed Abubakar Rasheed et al. [15] use the LMS and MOOCs online platform to make their research work to be effective enough for the better results.

Case studies are the key to understand the scenario of any research field and help researchers to understand the problems well. Mohammad Khalil et al. [5], Andrzej Ożadowicz [10] use conduct a detailed case study about the use to ML with blended learning concept and suggest the key elements for the study in this area.

The expert knowledge and traditional content delivery towards achieving a higher side of information literacy are the key factors which can affect the entire process of Blended Learning. Yong Shi et al. [12] and Adarsh Kumar et al. [13] suggest the overall structure to follow to get a high quality dataset and also reveal the front and initial side of the studies in this field.

Self-regulated learning and design based learning play a vital role in blended learning setup for initial starting with this technique. M. S. Faathima Fayaza et al. [3], Jac Ka Lok Leung

et al. [9] and Ali Saleh Alammery [11] focused on these area of blended learning concept and reveals the impact of SRL, DBL like phenomenon on their studies.

III. CHALLENGES AND SCOPE

By this comprehensive review, we find some significant research gaps. Those are as:

1. Many existing models focused the traditional approaches to find the performance of students in the blended learning setup. They haven't use any Machine Learning approach to find the results.
2. The researchers who used the ML approach, not collect the data by self for their study. They collect the available dataset and perform their study.
3. Only few studies use ML techniques with Blended Learning environment but not focused on the novel parameters which can affect the students' performance.

The review shows that no researcher have touch the area of novel parameters (finding and utilizing) for their studies. So there is an urgent need to find those parameters which can affect the student's performance in the blended learning environment. Our research would be able to handle both the aspects of blended learning to find performance and the parameters that affect this performance. We'll suggest a powerful model that combine the aspects of Blended Learning and Machine Learning. Based on this review, the following research goals are set:

- a. **To identify novel parameters influencing blended learning outcomes**, which has the capacity to alter the level of learning too.
- b. **To investigate the relationships between these parameters and learning outcomes using ML**, how these parameters are interlinked with each other to affect the model accuracy and results.

- c. To check the impact of novel parameter variations on the accuracy and consistency of machine learning models predicting blended learning outcomes, how these parameters can affect the machine learning accuracy?
- d. To collate the achievement of different ML algorithms to adjusting blended learning outcomes, by using blended learning environment it is necessary to check the achievement of different Machine-Learning- algorithms and find what will show greater performance.

IV. PROPOSED METHODOLOGY

This research proposes a hybrid model that combine the two major areas with each other (Blended Learning and Machine Learning). Finding and measuring novel parameters effects will be the major findings of this research.

A. Architecture of the model

This is the narrow view of the structure by which the research will carried out. There are always be two major parts of any research namely

1. Data Collection
2. Data processing

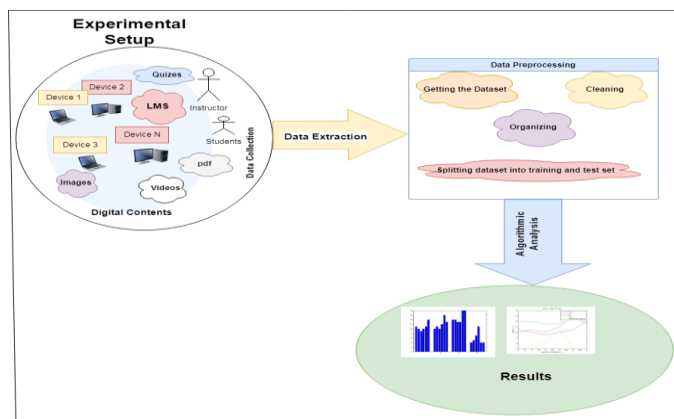
Same as this research will be done in two steps:

i. Step One: Creation of a controlled environment for Blended Learning Activities like:

- Selection of Course and Subject
- Preparation and Delivery of subject contents

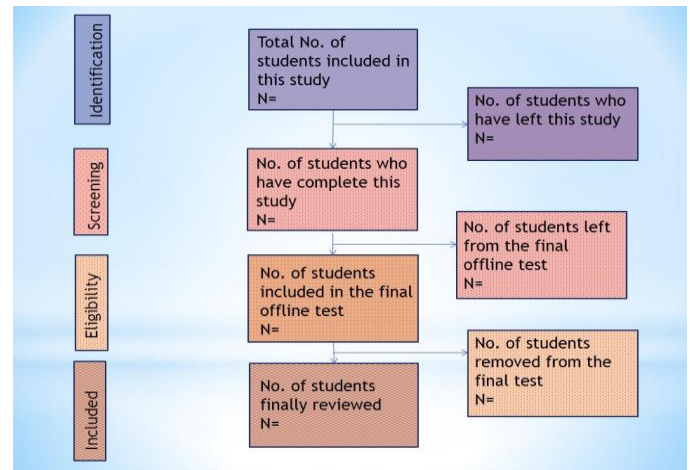
ii. Step Two: After step one collect the relevant data and prepare the Data Set and perform the necessary action according to the need

- All the collected data are taken into consideration.
- Select the fittest data set for the further processing.
- Train the selected data as per the need of the model.



B. Data Collection

The conceptual framework to collect the relevant data is truly a major part to discuss. We have select the following framework to collect the data:



C. Data Preprocessing

In machine learning workflows, data cleaning is an essential preprocessing step, especially for educational datasets where noisy, inconsistent, or missing records are prevalent. To maintain data integrity, it entails handling and identifying null values, fixing formatting mistakes, standardizing categorical variables, and getting rid of duplicates.

By insuring that learner behavior, engagement logs, and assessment records are trustworthy and consistent, appropriate data cleansing improves model accuracy in the context of blended learning systems. Effective feature extraction and model training are predicated on clean, well-structured data.

D. Proposed algorithm for this model

1. $D_{raw} \leftarrow \text{collect_dataset}()$
2. $D_{clean} \leftarrow \text{clean_data}(D_{raw})$
3. $F \leftarrow \text{extract_features}(D_{clean})$
4. $F_{norm} \leftarrow \text{scale}(F)$
5. $\text{labels} \leftarrow \text{define_target}(D_{clean})$
6. $\text{Split} \leftarrow \text{train_test_split}(F_{norm}, \text{labels}, \text{test_size}=0.2, \text{stratify}=\text{labels})$
7. $\text{models} \leftarrow \{\text{Logistic Regression, Random Forest, SVM, KNN, XGBoost}\}$
8. For each model in models:
 - a. Train the model on the training set.
 - b. Test it using Accuracy, F1, the Confusion Matrix, and ROC-AUC.
 - c. Perform cross-validation
9. $\text{best_model} \leftarrow \text{select model with highest validation performance}$
10. $\text{important_features} \leftarrow \text{feature_importance}(\text{best_model})$
11. Return $\text{best_model}, \text{performance_metrics}, \text{important_features}$

V. CONCLUSION AND DISCUSSION

This research will be based on some new aspect in the field of Blended Learning and Machine Learning. It combines the two different areas i.e. Blended Learning and Machine Learning. Both the fields are rapidly growing and attracting the researchers to find the new ways for these processes. The

inclusion of machine learning (ML) into blended learning (BL) conditions has demonstrated pushing potential in transforming the analysis, interpretation, and application of educational data to enhance academic performance. It is clear from the analysis of fifteen varied and highly significant studies that machine learning (ML) techniques—from conventional classifiers like Support Vector Machines and Logistic Regression to ensemble approaches like Random Forest and XGBoost—have been effectively used to forecast student performance, identify at-risk students, and tailor teaching methods. These methods not only offer automated insights into student engagement and learning paths, but they also help teachers make timely interventions, particularly in hybrid learning environments where it's necessary to combine data from offline and online sources. The literature analysis also shows that there is increasing interest in figuring out novel variables that affect students' performance in mixed learning environments. Emerging characteristics like digital engagement patterns, submission timing, forum involvement, and behavioral consistency have been investigated as potential predictors in addition to traditional indicators like attendance and test results. The generalizability of models across institutions, data privacy difficulties, and the scarcity of datasets are still major problems, nevertheless. All things considered, the intersection of ML and BL is a promising area of study that is still developing; future research is anticipated to concentrate on interpretable models, cross-platform data integration, and scalable solutions that are suited to various learning settings.

ACKNOWLEDGMENT

Kuldeep Chauhan like to thank the School of Engineering and Technology, Shobhit University, Gangoh, India for supporting and encouraging him to work on this research.

REFERENCES

- [1] Verma, C., Illés, Z., & Kumar, D. (2024). TCLPI: Machine Learning-Driven Framework for Hybrid Learning Mode Identification. *IEEE Access*.
- [2] Wang, Y., Wu, J., Chen, F., & Li, J. (2024). Analyzing teaching effects of blended learning with LMS: an empirical investigation. *IEEE Access*, 12, 42343-42356.
- [3] Fayaza, M. F., & Ahangama, S. (2024). Systematic review of self-regulated learning with blended learning in digital space. *IEEE Access*.
- [4] Makgoba, Boitumelomakgoba & Martins, Dr & Arasomwan, Martins. (2024). The application of machine learning algorithms in a blended learning environment to predict student academic performance.. 10.13140/RG.2.2.18367.28326.
- [5] Khalil, Mohammad & Shakya, Ronas & Liu, Qinyi & Ebner, Martin. (2024). How to Plan and Manage a Blended Learning Course Module Using Generative Artificial Intelligence?. 10.1007/978-981-97-9388-4_4.
- [6] Azizah, Zahra & Ohyama, Tomoya & Zhao, Xiumin & Ohkawa, Yuichi & Mitsuishi, Takashi. (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. 10.1016/j.caeai.2024.100261.

- [7] Ye, Shengpei. (2023). Blended Learning for Machine Learning-based Image Classification. *EAI Endorsed Transactions on e-Learning*. 9. 10.4108/eetel.4509.
- [8] Ofori, Francis & Matheka, Abraham & Maina, Elizaphan. (2023). Critical literature review on current state-of-the art in predicting students' performance using machine learning algorithm in blended learning environment. Vol (5). 23-38.
- [9] Leung, J. K. L., Chu, S. K. W., Pong, T. C., Ng, D. T. K., & Qiao, S. (2021). Developing a framework for blended design-based learning in a first-year multidisciplinary design course. *IEEE Transactions on Education*, 65(2), 210-219.
- [10] Ożadowicz, Andrzej. (2022). Interactivity —A Key Element of Blended Learning with Flipped Classroom Approach. 1-6. 10.1109/ICELIE55228.2022.9969439.
- [11] Alammery, A. S. (2022). A toolkit to support the design of blended learning courses. *Ieee Access*, 10, 85530-85548.
- [12] Shi, Y., Peng, F., & Sun, F. (2022). A blended learning model based on smart learning environment to improve college students' information literacy. *Ieee Access*, 10, 89485-89498.
- [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] Aslam, S. M., Jilani, A. K., Sultana, J., & Almutairi, L. (2021). Feature evaluation of emerging e-learning systems using machine learning: An extensive survey. *IEEE Access*, 9, 69573-69587.
- [15] Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2021). An approach for scaffolding students peer-learning self-regulation strategy in the online component of blended learning. *Ieee Access*, 9, 30721-30738.



Kuldeep Chauhan received the B.Tech. and MBA degree from Uttar Pradesh Technical University, Lucknow and the M.Tech. degree from Shobhit University, Gangoh. He is currently pursuing the Ph.D.degree in computer science and engineering. He is also an Assistant Professor in the Department of CSE, School of Engineering and Technology, Shobhit University, Gangoh, India. He is an experienced educator with over ten years of teaching and research experience in the area of computer science. He has published several research articles in international journals and has presented at reputed international conferences. His research interests include Blended Learning, Machine Learning technology and Business Intelligence.



Dr. Varun Bansal serves as a Professor in the Department of computer science and engineering at Shobhit University, Gangoh, bringing over 20 years of academic and research expertise. His core research areas include Honeypot Technology, machine learning and Blended Learning. Dr. Bansal has authored several publications in prominent international journals and conferences and has mentored many postgraduate and doctoral researchers. He remains actively engaged in academic partnerships and has played a vital role in advancing curriculum innovation and promoting research-driven learning within the department.



Dr. Suryakant Pathak received his MCA, M.Tech.in Computer Science & Engineering & Ph.D. in Wireless Communication from BBD University Lucknow, India. He had worked in Indian Air Force, 1983-2003, Faculty at NIELIT Deemed to be University Gorakhpur (UP), India, 2004-2006. Lecturer at

Amity University Lucknow (UP), India, 2007-2008. Assistant Professor at Sagar Institute of Technology & Management Barabaki (UP), India 2009 -2013. Associate Professor and Head of Department Computer Science & Engineering at Dr. KN Modi University Newai, Rajasthan, India, 2014-2018. Director at KIPM College of Engineering & Technology GIDA Gorakhpur (UP), India, 2019-2023. At present working as Professor in the department of Computer Science & Engineering at Shobhit University Gangoh Saharanpur (UP), India.



Dr Anil Kumar was awarded a Ph.D. in Computer Science and Engineering from Guru Gobind Singh Indraprastha University, New Delhi. His research expertise spans privacy preservation, information security, the Medical Internet of Things (MIoT), and machine learning. His doctoral work focused

on developing a privacy-preserving framework for electronic health records using soft computing techniques.

Dr Kumar has a strong research background in identifying critical gaps, formulating mathematical models, and conducting in-depth experimental analyses. He has published over 20 research papers in reputed SCI/SCIE and Scopus-indexed journals, edited three books in the field of computer science, and holds six national and international patents.

With over 15 years of teaching experience following his M.Tech. in Software Engineering and 3.5 years of industry experience after completing an M.Sc. in Computer Science, he currently serves as an Associate Professor in the Department of Artificial Intelligence and Machine Learning at Galgotias College of Engineering and Technology, Greater Noida, India.