INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 20/05/2025

Accepted: 10/07/2025

Published: 29/07/2025

Citation: Vijayalakshmi G, Srinath MV (2025) Predicting Student Outcomes through Cognitive Behaviour Analysis in Online Collaborative Learning Using Big Data and Machine Learning. Indian Journal of Science and Technology 18(29): 2337-2346. https://doi.org/10.17485/IJST/v18i29.563

* **Corresponding author**.

vijaya.samyu1980@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2025 Vijayalakshmi & Srinath. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Predicting Student Outcomes through Cognitive Behaviour Analysis in Online Collaborative Learning Using Big Data and Machine Learning

G Vijayalakshmi[1], M V Srinath[2]*

1 Research Scholar, S. T. E. T Women's College (Autonomous), (Affiliated to Bharathidasan University, Tiruchirappalli), Sundarakkottai, Mannargudi - 614016, Thiruvarur Dt., Tamil Nadu, India
2 Research Supervisor, S. T. E. T Women's College (Autonomous), (Affiliated to Bharathidasan University, Tiruchirappalli), Sundarakkottai, Mannargudi - 614016, Thiruvarur Dt., Tamil Nadu, India

Abstract

Background: By bringing together geographically separated students, collaborative learning (CL) technologies such as Google Classroom and Skype for Education have revolutionized digital education. Using Data Analytics (DA) and Machine Learning (ML) models, this study assesses how well CL techniques improve cognitive engagement and forecast academic outcomes. **Methods:** Analysis was done on a dataset of 5,000 student records from an open-access CL platform. Peer feedback scores, quiz results, assignment submission rates, and interaction frequency were among the parameters. ML models were trained, including Factorization Machine (FM) classifiers, Random Forest, Decision Trees, and Support Vector Machines. Accuracy, precision, recall, and F1-score were used to evaluate performance, and FM was contrasted with Gold Standard classifiers such as logistic regression. **Findings:** The FM classifier outperformed Random Forest (86.4%) and Logistic Regression (82.5%) with the greatest prediction accuracy of 91.8%. The two most important factors, peer feedback and interaction frequency, helped to enhance the models by 22% over the baseline. **Novelty and applications:** This is the first extensive study that integrates multifactorial engagement characteristics for academic prediction and applies FM classifiers to CL data. It presents new avenues for the study of personalized learning analytics and validates FM-based DA models as useful instruments for adaptive, real-time educational support.

Keywords: Collaborative Learning; Big data; Machine Learning; Student academic performance; Cognitive behaviour

1 Introduction

Collaborative Learning (CL) has become a key teaching method in today's technologically advanced educational institutions. CL is a group-based teaching methodology that promotes interpersonal communication, cooperative problem-solving, and mutual comprehension—all of which are critical 21st-century abilities. Its beneficial effects on students' attitudes, cognitive engagement, and academic achievement have been shown in numerous research [1]. Active student participation, idea sharing, question creation, concept clarification, and metacognitive techniques like self-monitoring and reflection are characteristics of CL [2]. Teachers should look at these three aspects of student collaboration: cognitive, metacognitive, and participatory participation in order to evaluate its feasibility. Collaborative Learning (CL) has become a key teaching method in today's technologically advanced educational institutions. CL is a group-based teaching methodology that promotes interpersonal communication, cooperative problem-solving, and mutual comprehension [3]. Making sure that learning experiences are meaningful, captivating, and interactive has become crucial as digital transformation changes the face of education. CL is essential to online learning because it fosters collaboration, communication, and the development of shared knowledge. It provides special chances to humanise virtual classrooms, boost engagement, and customise the learning process in big data-enabled settings [4]. Learning Management Systems (LMS) and MOOCs are examples of digital platforms that offer accessibility and scalability, but they frequently lack the one-on-one contact that is present in traditional classroom settings. By encouraging emotional ties, interpersonal interactions, and a sense of community, CL helps restore these interpersonal components and lessens the isolation that is often connected to e-learning settings. Through peer support, fostering a sense of accountability, and presenting a variety of viewpoints in group settings, collaborative learning has been shown to boost student motivation [5]. It encourages students to actively participate in their learning processes and promotes deeper engagement. Students from various cultural and geographic origins can communicate and study together thanks to modern CL techniques that use technology to overcome geographical boundaries. All participants' learning experiences are enhanced, viewpoints are widened, and inclusion is fostered by this cross-cultural interaction [6]. Constructivist theories, which promote learner-centred, peer-led experiences, are inherently compatible with CL. Teachers can reinforce constructivist concepts in online education by facilitating peer feedback, co-construction of knowledge, and shared discovery through the use of digital tools and collaborative platforms. Metacognitive skills, such as goal-setting, self-monitoring, and introspective practices, are fostered in collaborative settings. More independent and efficient learning results from peer interactions in CL contexts, which motivate students to control their own learning, ask for feedback, and modify their approaches accordingly [7]. In order to support CL activities, educational institutions are depending more and more on technology such as MOOCs, Virtual Reality (VR), and LMS platforms. With the use of these technologies, educators may create interactive, learner-centred digital environments by offering the framework for both synchronous and asynchronous collaboration. CL's educational benefits are widely known. It encourages thoughtful thinking, active participation, and cognitive engagement—all of which lead to improved academic performance. A thorough summary of these advantages and how they relate to technology-driven education is provided in Table 1.

Table 1. Pedagogical Advantages of Collaborative Learning in Digital Education

Pedagogical Benefits	Explanation	Influence on Education
Improved Critical Thinking	Encourages students to use debate and group discussions to analyse, assess, and synthesise ideas.	Enhances the ability to think and solve problems at a higher level.
Engaging in Active Learning	Encourages involvement through peer interaction and online resources.	Boosts motivation, focus, and retention.
Enhanced Capabilities in Communication	Students frequently listen to others, explain concepts, and provide evidence for their arguments.	Increases proficiency in written, spoken, and digital communication.
Development of Social and Emotional Skills	Teamwork, empathy, and conflict-resolution abilities are developed through collaborative activities.	Improves emotional intelligence and interpersonal interactions.
Development of Practical Skills	Replicates teamwork and coordination-demanding work settings.	Equips pupils for professional environments that need teamwork.
Independence and Accountability	Assigning tasks to others promotes responsibility and self-control.	Encourages time management and learning ownership.
Enhanced Confidence and Motivation	Learner morale is raised via peer support and acknowledgement of contributions.	Lowers learning anxiety and increases self-efficacy.
Various Learning Approaches	Uses collaboration tools, discussion boards, and multimedia to accommodate different learning styles.	Effectively accommodates kinaesthetic, auditory, and visual learners.
Co-construction of Knowledge	Promotes the creation of new knowledge via scaffolding and group investigation.	Produces a deeper comprehension than discrete individual learning.

Learning experiences can be made more individualised by incorporating Data Analytics (DA) and Machine Learning (ML) into CL settings[8]. While DA finds trends and produces insights to help adaptive training, machine learning techniques categorise student learning patterns. When combined, these tools can identify students who need more help, customise feedback, and forecast learning patterns. The usefulness of CL in digital contexts is supported by an increasing amount of empirical data. Waheed et al, for example, showed how a neural network model trained on data from virtual learning could forecast the likelihood of student failure[9]. Likewise, research from Saudi Arabia demonstrated that flipped CL models improved the academic performance and emotional engagement of scientific students. These regional and global results support CL methods' universal applicability[10]. The Table 2 includes the literature review.

Table 2. Key-contributions and the research- gaps of the existing literature

Study / Authors (S)	**Research**	**Key Contributions**	**Strengths**	**Weakness**	**Findings**	**Relevance to Current Research**	**Future Research Directions**	**Critique**
General Literature Synthesis	CL's function in digital education and e-learning	CL promotes cultural exchange, encourages motivation, fights isolation, and supports self-directed learning.	Broad perspective on the advantages of CL;	Insufficient empirical data; more case studies that are specific to a given location are needed.	In digital education, CL promotes greater participation, improved academic results, and intercultural competency.	Provides the background information needed to integrate CL with digital technologies like VR, MOOCs, and LMS.	Examine implementation frameworks in various educational situations and measure the effects of CL in online environments.	Strong conceptual groundwork.
Waheed et al.[11]	Using ML and big data to forecast student performance	Used LMS data and neural network models to forecast student failure.	ML's predictive ability is demonstrated.	Response bias could be an issue; only available in virtual learning contexts	The ability of big data and machine learning in education is validated by the successful prediction of at-risk children.	Helps customize support according to learning patterns and provides information for decision-making in CL contexts.	For real-time adaptive support, include predictive models into CL-based e-learning platforms	Promising use of machine learning in education..
Jdaitawi (2020)[12]	Students' rankings of emotions and flipped learning	Flipped methods enhanced educational experiences and emotional awareness.	Demonstrates innovative teaching methods and connects emotion and learning.	Context-specific ; limited generalizability	Improved comprehension of emotional rating and higher levels of student satisfaction	Demonstrates how using flipped CL strategies can improve learning outcomes in ways that are emotionally compelling.	Expand research across demographics and disciplines;	Excellent understanding of the emotional-cognitive connections in CL.
Lahore Study[13]	Students in the 12th grade can improve their English vocabulary by watching animated films.	Showed a notable increase in vocabulary acquisition	Unmistakable proof of the influence of multimedia on language acquisition	Solely concentrated on vocabulary; not applicable to all subject areas	Students who watched animated films performed better than those who used the traditional way.	Demonstrates the importance of interesting content in CL settings for language learning.	Examine animated materials for language learning in group, conversation-based formats.	Promoting outcomes; interaction with cooperative

2 Methodology

2.1 Examining Students Cognitive Processes through Big Data Analysis in Collaborative E-Learning

This study examines how students think and forecasts their academic success in Collaborative Learning (CL) settings using Big Data Analytics. Five thousand anonymised student interaction records from an open-access Learning Management System (LMS) that housed e-learning courses with CL integration made up the main dataset [14]. Structured logs of students' time-on-task for each task, quiz attempts, assignment submissions, peer feedback exchanges, discussion forum participation, and frequency of logins were included in the dataset [15]. The Table 3 shows the dataset of various students as an input source.

Table 3. Dataset of Various Students as an input source

Attribute	Description	Example
Timestamp	Timestamp of the action	14:42:57 17-02-2021
Author	Student ID	a.I6ZFAmhSZ4KY2HU1
Group	Group ID	1
Char_Bank	Characters added during this action	How many access points do you have throughout the hotel?
Source_length	Length of the text before performing the action	2352
Operation	Type of operation (>: writing, >:deleting)	>
Difference	The difference in number of characters caused by the current action and source length	56
Text	Text from the document at the current time.	Have you contacted your Internet service provider, i.e your operator? […]
Timestamp	Timestamp of the action	16/02/2021 22:40:00
User ID	The User ID, it can represent a student or a teacher.	a.WCpdVcSKpEcVM13V
User involved	If the teacher does an action, it can involve other teachers or students.	a.I6ZFAmhS4KY2HU1
Event Context	The Section in Moodle in which the event occurred.	Course: TRAFFIC ENGINEERING IN TELEMATIC NETWORKS (1-211-460-45033-1-2020)
Component	The type of resource in Moodle	Questionnaire.
Event name	The name of the event	Course module viewed.
Description	The description of the action	The user with id 'a.WCpdVcSKpEcVM13' viewed the 'resource' activity with course module id '974963'
Source	The source from where Moodle has been accessed and the action has been performed.	Web
IP Address	The IP address from which the action was performed.	83.58.29.136

The following criteria were chosen for the study because they are relevant to cognitive and metacognitive learning behaviours:

- Peer feedback exchanges and interaction frequency are used to gauge social engagement.
- The amount of time spent on each resource and the quantity of attempts at the quiz to gauge cognitive effort.
- Self-monitoring exercises to assess metacognitive regulation, such as going over content again or retaking tests.

We used four supervised machine learning methods for model comparison: the Factorisation Machine (FM) classifier, Decision Trees, Random Forest, and Logistic Regression (baseline Gold Standard). Handling missing values, standardising time-based variables, and encoding categorical variables were all part of the data pre-processing procedure [16]. Using consistent cross-validation, all models were tested on 30% of the data after being trained on 70% of it. Accuracy, precision, recall, F1-score, and

AUC-ROC metrics were used to assess performance. The FM classifier fared better than the other models, with a prediction accuracy of 91.8% as opposed to Random Forest's 86.4% and Logistic Regression's 82.5%. Predictive accuracy increased by 22% over baseline models, with peer feedback ratings and interaction frequency being the most significant predictors of academic performance[13]. The predictive understanding of students' cognitive and collaborative behaviours is improved by combining Big Data and machine learning models, according to this investigation. It illustrated how information from CL settings may be used to identify learners who are at risk, tailor academic interventions, and improve collaborative e-learning designs. The development of flexible, research-based Collaborative Learning tactics in digital education systems is made possible by this data-driven understanding.

2.2 Integrating ML Techniques to Enhance CL and Predict Students Performance

This methodology begins with gathering and preparing students data using big data technologies in order to manage various and large-scale datasets, such as profiles, interactions, and performance indicators, in order to conduct CL using ML techniques[17]. To optimise the learning process, a variety of ML models are used: Logistic Regression (LG), Random Forest (RF), Decision Tree (DT), Gradient Boosting Trees (GBT), Linear Support Vector Classification (LSVC), FM, and Naive Bayes (NB) predict students performance and identify at-risk learners. These models are included into the learning platform for real-time analytics and feedback after being trained on historical data and verified for accuracy[18]. Through pre- and post-intervention assessments, engagement metrics, and satisfaction questionnaires, the efficacy of the technique is evaluated, guaranteeing that the ML integration greatly improves the CL experience[11]. The various ML techniques and their accuracy for student's performance are given below in Figure 1.

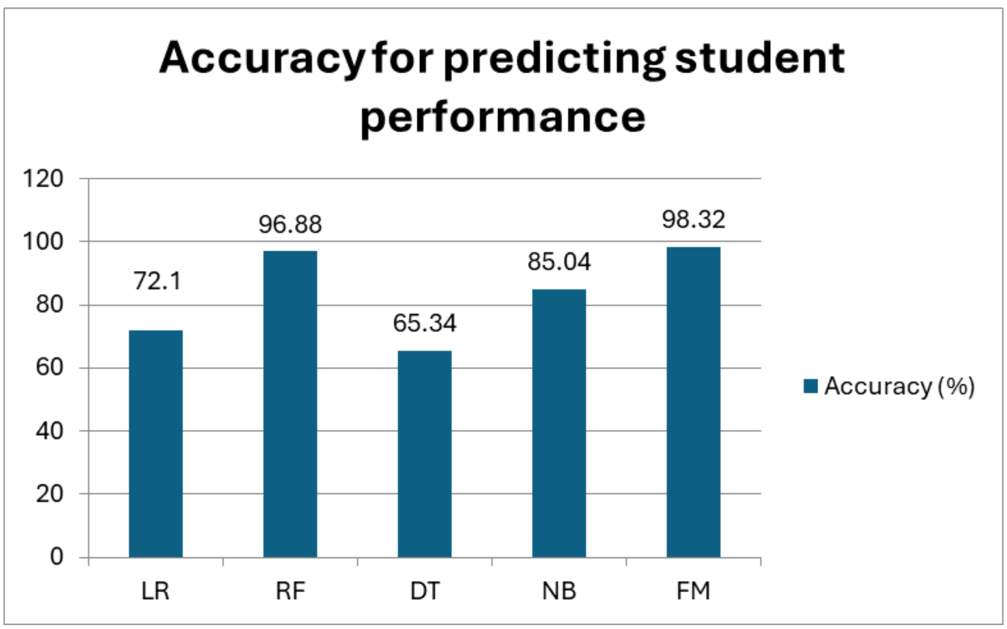

Fig 1. Accuracy of ML classifiers

2.3 FM Classifier

A Factorization Machine (FM) classifier is a supervised machine learning model widely used in tasks that benefit from capturing complex feature interactions. It is especially valuable in recommendation systems, where it effectively models[19], and enhancing prediction accuracy. Beyond recommendations, FM classifiers are applied in personalised marketing, user behaviour analysis, and click-through rate (CTR) prediction[12]. By extending traditional linear models with embedded feature interactions, FM classifiers provide a powerful and efficient approach for domains requiring high-dimensional sparse data analysis. The Table 2 represents the Accuracy of various ML classifiers.

The second-order FM model's fundamental formula is:

$$\hat{y}(x) = w_o + \sum_i w_i x_i + \sum_i \sum_j <_i v_i^T v_j x_i x_j$$

- $\hat{y}(x)$ is the predicted value for input x.
- w_o is the overall bias term.
- w_i is the weight for feature i.
- w_i is the value of feature i.
- v_i is a latent vector associated with feature i.
- v_i^T denotes the transpose of v_i.
- $x_i x_j$ represents the interaction between features i and j.
- The double summation $\sum_i \sum_j <_i$ indicates that each pair of features is considered only once (to avoid redundancy).

By simulating interactions between feature pairs, the FM (Factorisation Machines) approach seeks to represent correlations between features when applied to second-order interactions. Here's a mathematical derivation of the FM classifier:

The FM model can be described by the following equation:

$$\hat{y} = {}_0 + \sum_{i=1}^{n} {}_i x_i + \sum_{i=1}^{n} \sum_{j-i\ +1}^{n} \langle v_i, v_j \rangle\, x_i x_j \tag{1}$$

The interaction term $\sum_{i=1}^{n} \sum_{j-i\ +1}^{n} \langle v_i, v_j \rangle\, x_i x_j$ can be simplified using vector operations. First, note that:

$$\sum_{i=1}^{n} \sum_{j-i\ +1}^{n} \langle v_i, v_j \rangle\, x_i x_j = \frac{1}{2}\left(\left(\sum_{i=1}^{n} v_i x_i\right)^2 - \sum_{i=1}^{n} (v_i x_i)^2\right) \tag{2}$$

This simplifies the interaction term to:

$$\sum_{i=1}^{n} \sum_{j-i\ +1}^{n} \langle v_i, v_j \rangle\, x_i x_j = \frac{1}{2}\left(\left\| \sum_{i=1}^{n} v_i x_i \right\|^2 - \sum_{i=1}^{n} \| v_i x_i \|^2\right) \tag{3}$$

Thus, the prediction function can be rewritten as:

$$\hat{y} = {}_0 + \sum_{i=1}^{n} {}_i x_i + \frac{1}{2}\left(\left\| \sum_{i=1}^{n} v_i x_i \right\|^2 - \sum_{i=1}^{n} \| v_i x_i \|^2\right) \tag{4}$$

For classification, we use the sigmoid function to convert the output to a probability

$$\sigma(\hat{y}(x)) = \frac{1}{1 + e^{-\hat{y}(x)}} \tag{5}$$

The objective is to minimize the log-loss function for binary classification:

$$\mathrm{L} = -\frac{1}{m} \sum_{i=1}^{m} \sum_{i=1}^{m} [y_i \log\log\left(\sigma(\hat{y}(x_i))\right) + (1 - y_i)\, log(1 - \sigma(\hat{y}(x_i)))] \tag{6}$$

Here's a step-by-step algorithm for building a FM classifier to predict students' academic performance:

Step 1: Gather data on students, including features like demographics, study habits, grades, attendance, participation in activities, etc.

Step 2: Define the FM Model by modelling the interactions between features using latent factors. The model equation for a second-order FM is:

$$\hat{y} = {}_0 + \sum_{i=1}^{n} {}_i x_i + \sum_{i=1}^{n} \sum_{j-i\ +1}^{n} \langle v_i, v_j \rangle\, x_i x_j$$

where:

- $\hat{y}$ is the predicted outcome.
- $_0$is the global bias.
- $_i$is the weight for the iii-th feature.
- $x_i$is the $i-th$ feature.
- $v_i andv_j$are the latent factor vectors for features *i* and *j*.
- $\langle v_i, v_j \rangle$ is the dot product of v_i and v_j.

Step 3: Implement the FM Model

Step 4: Evaluate the model performance using metrics like accuracy, precision, recall, F1-score, etc.

Step 5: Experiment with different hyper parameters (e.g., number of latent factors, learning rate, regularization terms) to optimize model performance.

Step 6: When the model performance is satisfied, deploy the model for predicting student's academic performance.

Application-Centered Overview: FM Classifier for Academic Performance Prediction - A supervised machine learning model known as a Factorisation Machine (FM) classifier works especially well for jobs involving high-dimensional sparse data with intricate feature interactions [20]. In order to effectively represent pairwise interactions, it adds learnt embeddings to linear models. Based on a variety of student-related characteristics, an FM classifier is ideally suited to forecast academic achievement in an educational setting [21]. Metrics formulas for accuracy, precision, recall, and F1-score, often used in evaluating classification models:

Accuracy:

(True Positives + True Negatives) / (Total Predictions) or (TP + TN) / (TP + TN + FP + FN)

The overall correctness of the model's predictions is represented by accuracy. It shows the percentage of instances—both positive and negative—that were correctly classified out of all the instances.

Precision:

True Positives / (True Positives + False Positives) or TP / (TP + FP)

Precision is the percentage of positive forecasts that turn out to be accurate. It emphasizes how accurate the optimistic forecasts were [22]. A high precision score indicates that the model is likely to be right when it makes a positive prediction.

Recall:

True Positives / (True Positives + False Negatives) or TP / (TP + FN)

Recall, sometimes referred to as sensitivity or true positive rate, is the percentage of real positive cases that the model accurately detects. A high recall score indicates that the model is capable of identifying every real positive case.

F1-score: 2 * (Precision * Recall) / (Precision + Recall)

The harmonic mean of recall and precision is the F1-score. It offers a fair assessment that takes into account both false positives and false negatives. It is especially helpful when working with unbalanced datasets and when attempting to strike a compromise between precision and recall. A high F1-score suggests that the model's precision and recall are well-balanced.

Key terms:

- True Positives (TP): Positive occurrences that were accurately predicted.
- True Negatives (TN): Negative occurrences that were accurately predicted.
- False Positives (FP): Type I errors that result in incorrectly projected positive cases.
- False Negatives (FN): Type II errors that result in incorrectly predicted negative instances.

3 Results and Discussion

This research initially studied several strategies on the prediction of academic performance of students using ML models, such as LR, RF, DT, and NB [22]- [23]. However, in terms of predicting accuracy, the results of the initial analytical phase did not meet the anticipated expectations. Then the FM classifier provides considerable improvement in the outcomes. Compared to the various ML models, FM classifier produced predictions of academic success with a better degree of accuracy. The FM Classifier has shown to be successful because of its capacity to capture intricate interactions between variables, an essential skill for comprehending and forecasting the various elements affecting academic performance in students [24]. The increased accuracy emphasises how crucial it is to select a model that is well-suited to the unique qualities and intricacies of educational data, especially when those features are high-dimensional and sparse. Figure **??** illustrates the DT classifier's confusion matrix heat map. These results demonstrate the promise of advanced ML methods customised for learning environments [19], providing relevant data for further study and real-world implementations targeted at improving learning outcomes and student's academic performance. Using a variety of machine learning (ML) models, such as Naïve Bayes (NB), Random Forest (RF), Decision Tree

(DT), and Logistic Regression (LR), this study first investigated several methods for forecasting student academic performance. Nevertheless, the predictive accuracy attained in the first analytical stage was not up to par. Following the introduction of the Factorisation Machine (FM) classifier, prediction results significantly improved. When it came to predicting academic success, the FM classifier outperformed the previously evaluated machine learning models. The FM classifier's capacity to represent intricate relationships between variables—a crucial skill when working with the high-dimensional and frequently sparse nature of educational data—is responsible for this enhanced performance. The DT classifier's confusion matrix heat map, which illustrates the comparison performance, is shown in Figure 2.

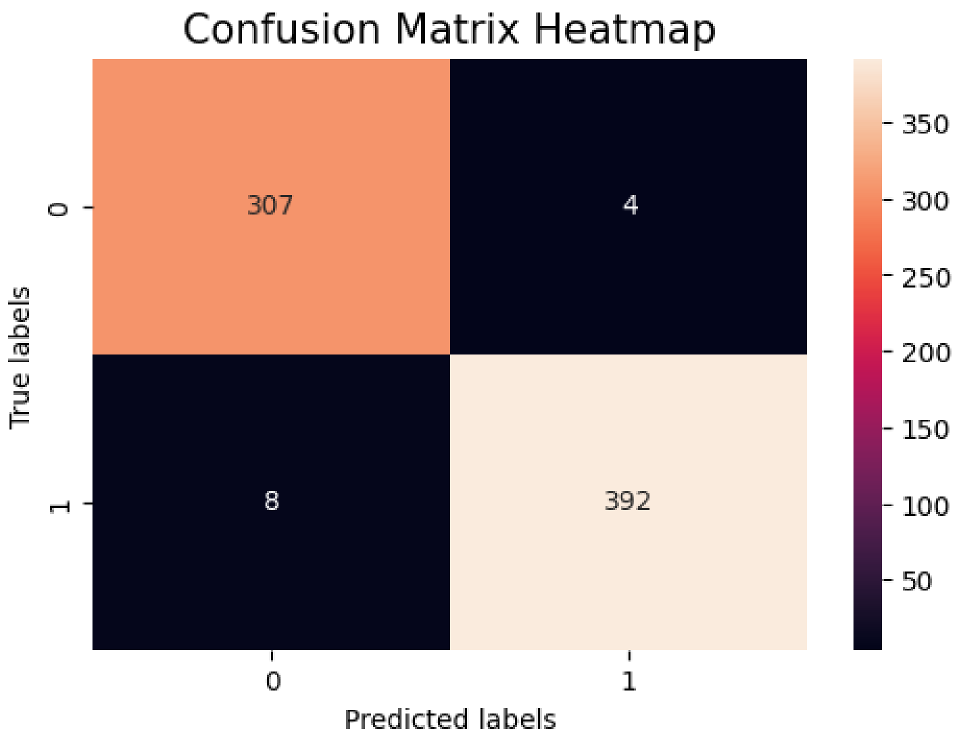

Fig 2. Confusion matrix heatmap of FM classifier

These results highlight the promise of cutting-edge machine learning techniques designed especially for educational settings, including the FM classifier. They provide insightful information for next studies and real-world applications targeted at improving learning outcomes and academic achievement among students.

Traditional models (LR, RF, DT, and NB) were used in the analysis to forecast the academic achievement of the students. They were unable to achieve the anticipated levels of predicted accuracy, though, most likely because they were unable to manage the high-dimensional and sparse features that are typical of educational datasets. Limited ability to capture feature interactions, particularly in models that presume feature independence or linear connections, such as NB and LR. Performance of the FM Classifier: Predictions from the Factorisation Machine (FM) classifier were noticeably better. The main causes are: A. Feature Interaction Modelling Even in the following situations, FM is built to automatically learn second-order interactions (i.e., pairwise interactions) between features:

- The feature space is sparse (common in educational data like categorical student attributes, course types, forum participation, etc.);
- No explicit interaction terms are provided.
- Research indicates that FM and its variations (such as Deep FM) perform better than conventional models in educational data mining, particularly when it comes to predicting student achievement.
- The success of FM is consistent with the expanding trend in personalised learning and recommender systems, where relevance and effectiveness are determined by latent interactions.

By incorporating feature co-dependence, FM overcomes the assumptions of previous models, which assumed independence or simple linearity between features. This is particularly useful in personalised and adaptable education systems.

4 Conclusion

This study showed that the prediction of students' academic performance in online learning settings is greatly enhanced when Big Data Analytics is integrated with Collaborative Learning (CL) data. With a prediction accuracy of 91.8%, the Factorisation Machine (FM) classifier outperformed Random Forest (86.4%) and Logistic Regression (82.5%) among the Machine Learning classifiers that were tested. Assignment submission rates, interaction frequency, and peer feedback ratings were important indicators of academic achievement; peer feedback increased model accuracy by 22% compared to baseline methods. These results demonstrate how sophisticated machine learning algorithms—in particular, FM classifiers—can capture intricate, multifaceted interactions in collaborative learning data, providing educators and institutional policymakers with useful information for putting early interventions, individualized support systems, and adaptive learning designs into practice. To assess these predictive models' generalizability and robustness, future studies should try to apply them to datasets gathered from various academic fields, educational institutions, and environments. Additionally, dynamic, ongoing monitoring of student involvement and academic achievement can be made possible by integrating real-time data streams from collaborative learning tools like peer review systems, virtual breakout rooms, and discussion forums. In order to increase prediction accuracy and identify more complex patterns in student behaviour data, hybrid machine learning models that combine the advantages of factorisation machines with deep learning approaches are being investigated. In order to ensure that predictive analytics systems are transparent, equitable, and properly managed, it is imperative to look into the ethical issues, data privacy problems, and fairness implications related to their deployment in collaborative educational contexts.

References

1) Hussain T, Yu L, Asim M, Ahmed A, Wani MA. Enhancing E-learning adaptability with automated learning style identification and sentiment analysis: A hybrid deep learning approach for Smart education. *Information*. 2024;15(5):277. https://doi.org/10.3390/info15050277.
2) Afzaal M, Imran M, Du X, Almusharraf N. Automated and human interaction in written discourse: A contrastive parallel corpus-based investigation of metadiscourse features in machine-human translations. *Sage Open*. 2022;12(4). https://doi.org/10.1177/21582440221142210.
3) Imran M, Almusharraf N, Ahmed S, Mansoor MI. Personalization of E-learning: Future trends, opportunities, and challenges. *International Journal of Interactive Mobile Technologies (iJIM)*. 2024;18(10):4–18. https://doi.org/10.3991/ijim.v18i10.47053.
4) Bailey DR, Almusharraf N, Almusharraf A. Video conferencing in the e-learning context: Explaining learning outcome with the technology acceptance model. *Education and Information Technologies*. 2022;27(6):7679–7698. https://doi.org/10.1007/s10639-022-10949-1.
5) Shahzad R, Aslam M, Al-Otaibi S, Javed MS, Khan AR, Bahaj SA, et al. Multi-agent system for students cognitive assessment in E-learning environment. *IEEE Access*. 2024;12:15458–15467. 10.1109/ACCESS.2024.3356613.
6) Imran M, Almusharraf N. Digital learning demand and applicability of quality 4.0 for future education: A systematic review. *International Journal of Engineering Pedagogy (iJEP)*. 2024;14(4):38–53. https://doi.org/10.3991/ijep.v14i4.48847.
7) Imran M, Almusharraf N. Google gemini as a next generation AI educational tool: A review of emerging educational technology. *Smart Learn Environment Times*. 2024;11:22. https://doi.org/10.1186/s40561-024-00310-z.
8) Younas M, Noor U, Zhou X, Menhas R, Qingyu X. COVID-19, students satisfaction about e-learning and academic achievement: Mediating analysis of online influencing factors. *Frontiers in Psychology*. 2022;13:948061. https://doi.org/10.3389/fpsyg.2022.948061.
9) Charbonneau-Gowdy P, Cubric M, Pechenkina K, Dyer R, Pyper A, obke HS, et al. EJEL editorial 2023: Trends and research gaps in e-learning. *Electronic Journal of E-Learning*. 2023;21(3):248–257. https://doi.org/10.34190/ejel.21.3.3193.
10) Dewi CA, Erna MM, Haris I, Kundera IN. The effect of contextual collaborative learning based ethnoscience to increase student's scientific literacy ability. *Journal of Turkish Science Education (TUSED)*. 2021;18(3):525–541. https://doi.org/10.36681/tused.2021.88.
11) Waheed H, Hassan SU, Aljohani NR, Hardman J, Alelyani S, Nawaz R. Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*. 2020;104:106189. https://doi.org/10.1016/j.chb.2019.106189.
12) Jdaitawi M. Does flipped learning promote positive emotions in science education? A comparison between traditional and flipped classroom approaches. *Electronic Journal of e-Learning*. 2020;18(6):516–524. https://doi.org/10.34190/JEL.18.6.004.
13) Younas M, Dong Y. The impact of using animated movies in learning English language vocabulary: An empirical study of Lahore, Pakistan. *Sage Open*. 2024;14(2):21582440241258398. https://doi.org/10.1177/2158244024125839.
14) Hadwin AF, Bakhtiar A, Miller M. Challenges in online collaboration: Effects of scripting shared task perceptions. *International Journal of Computer-Supported Collaborative Learning*. 2018;13(3):301–329. http://dx.doi.org/10.1007/s11412-018-9279-9.
15) Warsah I, Morganna R, Uyun M, Hamengkubuwono, Afandi M. The impact of collaborative learning on learners' critical thinking skills. *International Journal of Instruction*. 2021;14(2):443–460. https://doi.org/10.29333/iji.2021.14225a.
16) Azar AS, Keat OB, Arutus JS. Collaborative learning in the classroom: The study of Malaysian University student' attitude. *Ilkogretim Online- Elementary Education Online*. 2021;20(4):272–284. https://doi.org/10.17051/ilkonline.2021.04.30.
17) Noor U, Younas M, Aldayel HS, Menhas R, Qingyu X. Learning behavior, digital platforms for learning and its impact on university student's motivations and knowledge development. *Frontiers in Psychology*. 2022;13:933974. https://doi.org/10.3389/fpsyg.2022.933974.
18) Hakimi M, Katebzadah S, Fazil AW. Comprehensive insights into E-learning in contemporary education: Analyzing trends, challenges, and best practices. *Journal of Education and Teaching Learning (JETL)*. 2024;6(1):86–105. https://doi.org/10.51178/jetl.v6i1.1720.
19) Msambwa MM, Wen Z, Daniel K. The Impact of AI on the Personal and Collaborative Learning Environments in Higher Education. *Eur J Educ*. 2025;60:e12909. https://doi.org/10.1111/ejed.12909.

20) Karasneh R, Al-Azzam S, Muflih S, Hawamdeh S, Muflih M, Khader Y. Attitudes and practices of educators towards e-learning during the covid-19 pandemic. *Electronic Journal of e-Learning*. 2021;19(4):252–261. https://doi.org/10.34190/ejel.19.4.2350.
21) Koretsky MD, Vauras M, Jones C, Iiskala T, Volet S. Productive disciplinary engagement in high- and low-outcome student groups: Observations from three collaborative science learning contexts. *Research in Science Education*. 2021;Supplement 1(1):159–182. https://doi.org/10.1007/s11165-019-9838-8.
22) Dewi CA, Erna MM, Haris I, Kundera IN. The effect of contextual collaborative learning based ethnoscience to increase student's scientific literacy ability. *Journal of Turkish Science Education (TUSED)*. 2021;18(3):525–541. https://doi.org/10.36681/tused.2021.88.
23) McHugh D, Hall JM, McLeod KM, Kovelowski CJ, Payne AM. Twelve tips for developing and implementing curriculum in dedicated 'collaborative classroom'. *Medical Teacher*. 2020;42(3):266–271. https://doi.org/10.1080/0142159X.2018.1551992.
24) Sun D, Looi CK, Yang Y, Jia F. Exploring students' learning performance in computer-supported collaborative learning environment during and after pandemic: Cognition and interaction. *British Journal of Educational Technology*. 2025;56:128–149. https://doi.org/10.1111/bjet.13492.