

Research Article

Open Access



Decoding ChatGPT's impact on student satisfaction and performance: a multimodal machine learning and explainable AI approach

Nisar Ahmed Dahri^{1,2}, Fida Hussain Dahri³, Asif Ali Laghari⁴, Muhammad Javed⁵

¹Faculty of Educational Sciences and Technologies (FEST), University Technology Malaysia, Johor Bahru, Johor 81310 Malaysia.

²Faculty of Language Studies, Sohar University, Sohar 311, Oman.

³School of Computer Science and Engineering, Southeast University, Nanjing 211189, Jiangsu, China.

⁴Software College, Shenyang Normal University, Shenyang 110136, Liaoning, China.

⁵Department of Computer Science and Technology, College of Computer Science, Donghua University, Shanghai 54100, China.

Correspondence to: Dr. Asif Ali Laghari, Software College, Shenyang Normal University, Shenyang 110136, Liaoning, China.
E-mail: Asiflaghari@synu.edu.cn

How to cite this article: Dahri, N. A.; Dahri, F. H.; Laghari, A. A.; Javed, M. Decoding ChatGPT's impact on student satisfaction and performance: a multimodal machine learning and explainable AI approach. *Complex Eng. Syst.* 2025, 5, 7. <https://dx.doi.org/10.20517/ces.2025.07>

Received: 26 Jan 2025 **First Decision:** 13 May 2025 **Revised:** 17 May 2025 **Accepted:** 26 May 2025 **Published:** 29 May 2025

Academic Editors: Vijayakumar Varadarajan, Ding Wang **Copy Editor:** Fangling Lan **Production Editor:** Fangling Lan

Abstract

The extensive use of artificial intelligence (AI) tools in education and their effectiveness and explicability have become an essential area of investigation because of their scope for ethical and fair use. Therefore, this innovative approach proposed a multimodal machine learning and explainable AI (XAI) approach to predict how ChatGPT's usage impacts students' academic outcomes, including satisfaction and performance. Three machine learning models, XGBoost, Random Forest, and Support Vector Machine, were used to predict students' performance and level of satisfaction. The XAI techniques, SHapley Additive exPlanations and Local Interpretable Model-Agnostic Explanations, were used to investigate how models make decisions with ethical standards and ensure models are reliable and fair in their usage. A custom-created dataset, ChatGPT Survey Data, was utilized for the experiment. All three models gave the promised classification, with Support Vector Machine having the highest accuracy at 89% and the receiver operating characteristic curve and the area under this curve (AUC-ROC) at 92%. Random Forest has 87% accuracy and an AUC-ROC of 90%. XGBoost showed the best accuracy with 92% (R^2). XAI analysis revealed ChatGPT usage and satisfaction as key predictors of academic performance, advancing AI's role in education.

Keywords: ChatGPT, academic performance, student satisfaction, machine learning, explainable AI (XAI)



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



1. INTRODUCTION

Artificial intelligence (AI) has changed education around the world at a fast pace. Generative AI tools such as ChatGPT are significant in teaching and learning^[1,2]. The sophisticated language models of ChatGPT enable it to offer personalized tutoring, automatic grading, research assistance, and support with academic writing^[3,4]. With the increased usage of AI tools in education, there is increased interest in how they influence student learning and academic performance^[5,6]. Recent studies have shown that AI tools make students more engaged, help them explain complex topics better, and provide rapid feedback, which makes learning easier and more individualized for each student^[7,8]. The ability to create smart, human-like conversations with chatbots such as ChatGPT can make the learning environment engaging and dynamic. It is an invaluable tool for most learning scenarios, including self-paced education and teamwork in learning^[9]. Increased practice using generative AI in education remains largely unsupported empirically, with the full impact of this technology on student academic outcomes and factors that explain variations in student satisfaction and engagement in traditional tasks with these tools.

Most research on generative AI tools in education has focused on their use and specific cases^[10]. It has not paid much attention to how these tools affect student satisfaction and performance. Past studies have examined how ChatGPT can automate regular academic tasks or provide knowledge when studying for an exam^[11,12]. These studies do not look into crucial areas, such as how students using these tools can impact their motivation, critical thinking, and satisfaction with learning^[13]. Moreover, unclear explanations of AI models applied in education make it hard to determine why and how something impacts performance. Without a clear answer, education policymakers and practitioners cannot draw helpful insights into making AI tools for various types of students. Wu *et al.*^[14] and Rehman^[15] stated that only simple statistics or ordinary ML models were used in studies to determine the impact of AI on academic performance of students. This limits how deeply such studies can probe for insights from predictions and to what extent the findings can be transported into different contexts. Studies on AI applications in education rarely use an explainable AI (XAI) approach to present explanations and what is discovered. This is required because it connects the technical sophistication of AI with how AI is implemented in practice within education^[16]. These limitations demand a more integrated approach that merges the best ML techniques with the power of XAI, thereby yielding rich insights into the dynamics between the applied AI tools, academic performance, and resultant student satisfaction^[17,18]. Our study fills these gaps by using advanced ML models coupled with XAI to investigate the effects of the ChatGPT on students' academic outcomes.

This research introduces a new way to use ChatGPT and its impact on students' academic performance with the help of machine learning (ML) and XAI. It does not just look at a basic test to see whether ChatGPT is effective; it also studies how this relates to learning satisfaction, academic results, and student engagement. This research uses ML and XAI techniques to explore the link between ChatGPT and academic performance. However, the ML models mainly act as "black boxes", which means they provide less information about why they made their predictions. We use XAI methods to understand the models' predictions and decision-making and seek essential factors. The combination of ML and XAI ensures our results are accurate and useful.

This study aims to determine how ChatGPT relates to student learning and performance in school. The research will: (1) Forecast changes in satisfaction and grades among students who use ChatGPT using ML models; (2) Deploy transparent AI techniques to identify the primary motivations among the most common reasons for students using ChatGPT; and (3) Provide valuable recommendations to teachers, policymakers, and platform developers on using generative AI effectively. This research contributes to the growing knowledge of AI in education. This study offers worthwhile concepts that would help develop better

planning for improving the effectiveness of generative AI tools in schools to achieve these objectives.

2. LITERATURE REVIEW

AI has dramatically helped improve education. It aids the teacher in teaching and in determining how the students learn. AI tools help personalize learning experiences, simplify administrative work, and provide data-driven insights that improve teaching methods^[19,20]. Such new tools have helped meet the different needs of students, thereby allowing teachers to spend more time on teaching well and less on repetitive work. AI is used in learning platforms such as DreamBox and Knewton, which change educational content based on each student's strengths and weaknesses^[21]. These platforms make learning more engaging and improve the experience, especially for students who have trouble with traditional methods. The most significant application of AI in education is to make learning easier for everyone. It can translate speech into text, translate languages in real time, and provide content according to the needs of each learner^[22,23]. AI tools also help teachers know how students are doing in school, which allows them to step in when problems arise. Even with all these changes, issues still abound.

Generative AI models are becoming increasingly popular among education tools, such as the model used for ChatGPT. It is a conversational AI based on OpenAI's GPT-4 design^[24]. This model enables the generation of text similar to human writing and gives detailed explanations to help with various educational tasks, including automated grading, personalized tutoring, content generation, and assistance with academic writing^[25,26]. For example, students can use it to explain complicated topics, request feedback on their essays, or brainstorm project ideas^[26,27]. Similarly, instructors can use it to automatically make lesson plans, answer students' queries, and create exciting learning experiences^[28,29]. ChatGPT has greatly supported self-paced learning, wherein students can freely ask questions related to various subjects, developing confidence and facilitating a better understanding of them^[30]. Still, there are concerns about over-reliance on AI, loss of critical thinking skills, and the ethics of using AI-made content in assessments. According to ElSayary *et al.* (2024), research has proven that despite all these benefits, these tools should be used carefully to support classroom teaching instead of taking their place^[31].

ML is a part of AI and has helped many areas, including education. It started from fundamental computer theories in the middle of the 20th century and grew rapidly in the 1990s because of new algorithms that could sort and study large amounts of data^[32,33]. ML models are highly utilized in education for various purposes, including predicting student performance, learning habit analysis, and curriculum design enhancement^[32,34,35]. Among the best ML methods, Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) can help find patterns in student data for education improvement^[36,37]. For example, ML models can predict student risk based on attendance data, grades, and behavior, so help may be pursued whenever required^[34,37]. Similarly, k-means techniques group students based on their learning preferences to create a personalized learning plan^[36]. Even though ML models are compelling, their "black-box" nature often deters people from using them in crucial sectors such as education, where clarity and responsibility are essential^[38-40]. XAI solves the problem by making the ML model easy to understand so that teachers and decision-makers know the reasons behind the predictions made by the model^[36,40]. Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) make complex machine-learning models easily comprehensible. Methods are used to show how each feature affects predictions^[41,42]. XAI has been applied in education to reveal the primary reasons students drop out: their economic background, attendance, and grades^[41]. XAI helps teachers make decisions based on data, improving educational planning and student support^[43,44].

Introducing ML and XAI in education has opened new avenues for data-driven teaching and learning strategies. ML and XAI also play a significant role in developing intelligent tutoring systems (ITS). ML models have been used to predict students' exam performance, analyze student feedback, and automate administrative staff's work. With XAI, these predictions are still interpretable and actionable^[39,43,44]. Huang *et al.* (2024) examined factors influencing students' engagement with XGBoost and SHAP. According to them, time management and participation in group activities are some of the most important factors. This allows the teachers to devise specific plans for enhancing engagement^[45].

A clear example is using XAI to check the effectiveness of generative AI tools such as ChatGPT. The researchers use XAI techniques to identify the aspects, such as how often people engage with ChatGPT and the difficulty of questions that affect learning performance. This can lead to better tools for more tremendous educational success^[46]. Research studies indicate that ML and XAI can positively enhance academic results. Villegas-Ch *et al.* (2024) also highlighted how applying ML models for predicting students' performance in online courses may be related to engagement metrics as critical predictors^[47]. In addition, the SHAP explanation provided actionable insights for course designers^[48]. demonstrated the use of XGBoost and LIME in data analysis concerning social media addiction and how it impacts academic performance. It was found that too much screen time and a lack of self-control were the primary reasons. Lo and Kwan^[8] further discussed how ChatGPT can be used in teaching students. Using the empowerment and elements of XAI, the researchers established that ChatGPT's instant responses and compression of difficult topics further helped the students improve their understanding and confidence. However, it also calls for caution against over-reliance on this tool and demands a harmonious blend with the traditional teaching system^[49].

Combining AI, ML, and XAI in education will create significant opportunities for improving teaching and learning. For example, tools such as ChatGPT show what generative AI can do to make learning more personal and engaging. However, explainability, ethical issues, and dependency problems highlight the need for careful integration. The current research fills this gap by bridging technological innovation to practical applications in utilizing state-of-the-art ML models and XAI techniques that give much-needed insight into further enhancing AI tools in pursuit of educational achievement.

3. METHODOLOGY

Our study's methodology starts with the first step, data preparation to deal with missing values due to imputation techniques and encoding categorical variables such as gender and institution type using one-hot encoding^[50]. This puts the data in a standard format and prepares it for analysis. Four dimensions were used to classify variables executed in feature engineering^[51]: Academic performance, satisfaction and perception, usage and interaction, and stress and motivation. These dimensions allowed for better interpretability of the models and their effectiveness. An Exploratory Data Analysis (EDA) was conducted through visualizations and correlation matrices to understand the patterns and relationships in the data. The dataset was split into a ratio of 80%-20% for training and testing to test the models sufficiently. Further, our novel study explored the utilization of ML and the XAI approach to predict the effects of ChatGPT on several academic outcomes: performance, satisfaction, motivation, and related variables. The experiment used three ML models: RF, XGBoost, and SVM on a custom-created dataset. Hyperparameter tuning was also performed to maximize predictive accuracy. The whole process of the training methodology ensured the most reliable and accurate results in predicting the outcome of the education process. XAI tools such as SHAP and LIME were also applied to investigate models, their usage, how a model arrives at its predictions or decisions, and to explain the working process. As demonstrated, the applied data analysis methods supported a more in-depth understanding of global and local features explaining predictions, with Q2_ChatGPT_Satisfaction,

Q3_Performance_Impact, and Q4_Usage_Frequency variables being key predictors. The model's performance is evaluated using classical metrics, such as accuracy, precision, recall, F1-score, and the receiver operating characteristic curve and the area under this curve (AUC-ROC). [Figure 1](#) depicts the overall workflow of the proposed approach.

3.1. Data collection and survey design

This section describes the participants and demographics, survey design, data collection process, and analysis methods used to examine ChatGPT's impact on academic performance, satisfaction, and support in higher education.

3.1.1. Participants and demographics

Data was collected from 1,000 higher education students from various institutions to create a custom ChatGPT Survey Data dataset. The demographic profile of the participants was categorized by gender, academic year, and institution type. A summary of these demographics is presented in [Table 1](#).

The demographic distribution observed in [Table 1](#), including the male majority (68%) and variation across academic years, reflects real-world usage patterns of ChatGPT among higher education students. Rather than applying artificial demographic balancing, we retained this natural representation to preserve ecological validity. To address potential bias, demographic variables such as gender and academic year were incorporated as input features in our ML models. Their contributions were further analyzed through feature importance and SHAP interpretation techniques. This strategy allowed the models to account for demographic variance rather than ignore it. Future research may explore stratified sampling or weighting schemes to enhance demographic representativeness and cross-context generalizability.

3.1.2. Survey design and questions

The study employed a structured survey to gather responses from participants. The survey included 12 Likert scale questions addressing ChatGPT usage frequency, perceived effectiveness, satisfaction, and its impact on academic performance. Items were adopted from^[20,49]. Respondents rated each question on a 5-point Likert scale, where 1 represented "Strongly Disagree" and 5 represented "Strongly Agree". The survey questions are listed in [Table 2](#).

3.1.3. Data collection

The survey was distributed through online platforms, ensuring anonymity and voluntary participation. A total of 1,000 valid responses were collected and used for analysis. The description of dataset variables is included in [Table 3](#), and a summary of the dataset is included in [Table 4](#).

3.2. Data preparation

Several data preparation steps were used to ensure the dataset was adequate for input into ML classifiers. These steps aided in achieving smooth analysis, such as eliminating missing values where less than 5% accounted for missing data^[52,53]; the maximum variable "Q7_Accuracy" is shown to possess missing values by approximately 4.2%. In this regard, continuous variables such as "Q4_Usage_Frequency" and "Q5_Time_Reduction" were imputed using mean values. Categorical variables such as "Gender" were imputed using the mode to ensure completeness of the dataset without introducing bias. Min-max scaling was used to scale continuous variables, ensuring all continuous features contributed equally to the analysis. The Min-Max scaling is determined by

$$x' = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where:

- X is the original value of the feature.
- X_{\min} is the minimum value of the feature.
- X_{\max} is the maximum value of the feature.
- X_{scaled} is the scaled value, now in the range $[0, 1]$.

The categorical variable "Gender" was represented as "Male, Female, Other". One-hot encoding was done, transforming the above categorical variables to a numerical format to convert each binary column representing the categories. Using these variables makes the ML algorithms work effectively on these variables without any ordinal relation implied among the levels of categorical values. This sets a strong footing for further exploration and model development based on data.

3.3. Exploratory data analysis

The exploratory data analysis (EDA) supports understanding the nature of our custom-made dataset's characteristics and relationships or patterns before applying advanced modeling techniques. The task of this study was to apply EDA to gain further insight into and analyze the responses made on the survey, detect problems such as outliers or skewed distributions, and finally better decide on data preprocessing and feature selection. [Figure 2](#) plots the histograms showing distributions for every item of the survey. These histograms provide a clear view of the frequency of various responses across numeric variables, highlighting skewness or imbalance in the data. This visualization helped identify if any specific items were over- or underrepresented in the survey, which might affect the accuracy and generalizability of the models. It also showed whether the responses followed a normal distribution or had noticeable deviations.

Moreover, [Figure 3](#) shows feature box plots for the individual variables in the dataset, representing the dataset's central tendency, spread, and possible outliers in a graphical way. These box plots would prove helpful in identifying extreme values or variables with extensive interquartile ranges. We could see a need for features that may require further attention, either for normalization or to treat outliers, so as not to distort the model while in training.

Finally, [Figure 4](#) is the correlation matrix heatmap, presented to display relationships between numeric features of our self-crafted dataset, namely ChatGPT Survey Data. The color intensities of the heatmap point toward the strength and direction of correlation to identify highly correlated variables. In this regard, finding multicollinearity degrading ML models' predictive ability was essential. This helped us to refine the feature selection process so that only the most relevant and independent variables were included in the model.

3.4. Multicollinearity assessment

As shown in the correlation heatmap [[Figure 4](#)], inter-feature correlations are relatively low, with most values under 0.2, and only a weak correlation is observed between Q6_Problem_Solving and Q12_Learning_Motivation ($r = 0.13$). This indicates a low multicollinearity risk in the dataset. Since tree-

Table 1. Demographics of participants

Category	Subcategory	Frequency	Percentage
Gender	Male	680	68%
	Female	320	32%
Academic year	First year	150	15%
	Second year	200	20%
	Third year	150	15%
	Fourth year	200	20%
	Postgraduate	300	30%
Institution type	Public	700	70%
	Private	300	30%

Table 2. Likert scale survey questions

No.	Survey question	Likert scale options
1	ChatGPT has enhanced my understanding of academic concepts	1 2 3 4 5
2	I am satisfied with the academic support provided by ChatGPT	1 2 3 4 5
3	ChatGPT has positively impacted my academic performance	1 2 3 4 5
4	I use ChatGPT frequently to complete academic tasks	1 2 3 4 5
5	ChatGPT has reduced the time I spend using traditional learning resources	1 2 3 4 5
6	I rely on ChatGPT to solve complex academic problems	1 2 3 4 5
7	ChatGPT provides accurate and reliable academic information	1 2 3 4 5
8	ChatGPT makes me feel more confident about my academic abilities	1 2 3 4 5
9	I believe ChatGPT is a valuable tool for improving my academic performance	1 2 3 4 5
10	I would recommend ChatGPT to other students for academic purposes	1 2 3 4 5
11	ChatGPT helps me manage academic stress by offering quick solutions	1 2 3 4 5
12	ChatGPT has motivated me to explore new areas of learning	1 2 3 4 5

Table 3. Dataset structure and variables

Variable name	Description
Gender	Gender of the participant (Male/Female/Other)
Academic_Year	Year of study (First/Second/Third/Fourth/Postgrad)
Institution_Type	Institution type (Public/Private)
Q1_ChatGPT_Understanding	ChatGPT has enhanced my understanding of academic concepts
Q2_ChatGPT_Satisfaction	I am satisfied with the academic support provided by ChatGPT
Q3_Performance_Impact	ChatGPT has positively impacted my academic performance
Q4_Usage_Frequency	I use ChatGPT frequently to complete academic tasks
Q5_Time_Reduction	ChatGPT has reduced the time I spend using traditional learning resources
Q6_Problem_Solving	I rely on ChatGPT to solve complex academic problems
Q7_Accuracy	ChatGPT provides accurate and reliable academic information
Q8_Confidence	ChatGPT makes me feel more confident about my academic abilities
Q9_Valuable_Tool	I believe ChatGPT is a valuable tool for improving my academic performance
Q10_Recommendation	I would recommend ChatGPT to other students for academic purposes
Q11_Stress_Management	ChatGPT helps me manage academic stress by offering quick solutions
Q12_Learning_Motivation	ChatGPT has motivated me to explore new areas of learning

based models such as XGBoost and RF are less sensitive to multicollinearity and do not rely on coefficient estimation, traditional VIF metrics were not applied. Nonetheless, interaction features were explicitly engineered to isolate combined effects, reducing redundant feature influence.

Table 4. Dataset summary

Response	Item_1	Item_2	Item_3	Item_4	Item_5	Item_6	Item_7	Item_8	Item_9	Item_10	Item_11	Item_12
1 (SD)	108	108	104	106	99	101	106	107	100	92	96	97
2 (D)	211	188	194	209	204	200	198	208	208	202	196	200
3 (N)	294	274	296	294	310	304	322	314	323	324	304	288
4 (A)	234	278	252	261	246	242	220	231	234	246	243	243
5 (SA)	153	152	154	130	141	153	154	140	135	136	161	172

EDA provided in-depth visualizations and a comprehensive understanding of the dataset, identifying key trends and potential issues that could affect the accuracy of the models. The insights gained from this analysis informed the preprocessing steps, including imputation, scaling, and feature engineering, ensuring that the dataset was well-prepared for the subsequent machine-learning techniques.

3.5. Feature engineering

Using feature engineering techniques helped improve the predictive capability of the dataset in our study by developing new variables that better captured the relationships between student engagement with ChatGPT and associated academic satisfaction and performance^[52,54]. This paper discusses two explicit feature engineering techniques applied during this study to refine a dataset toward a more detailed analysis of student results' determinants^[54].

First, the features were combined into one that would create an overall measure of the impact of ChatGPT on student learning. The new feature, Overall ChatGPT Impact, was derived by aggregating the results from three key survey questions: Q1_ChatGPT_Understanding, which measures the clarity of explanations provided by ChatGPT; Q2_ChatGPT_Satisfaction, which assesses overall satisfaction with ChatGPT as a learning tool; and Q3_Performance_Impact, which captures perceived impact on academic performance. These three variables were averaged to create the Overall ChatGPT Impact score, thus providing a consolidated measure of the student experience with ChatGPT. This feature is aggregated and gives an overall view of how students perceive the tool's effectiveness; it helps quantify its influence on the students' academic journey.

The second technique focused on interaction features, investigating the relationship between students' problem-solving abilities and their intrinsic motivation for learning. An interaction term was created by multiplying Q6_Problem_Solving, which evaluates students' ability to solve complex problems with ChatGPT, and Q12_Learning_Motivation, which measures their intrinsic motivation towards learning. This new variable, the Interaction Feature, captures how motivation influences the students' ability to use ChatGPT effectively in solving problems. This would test whether the more substantial motivational level would be associated with achievement in solving problems using ChatGPT.

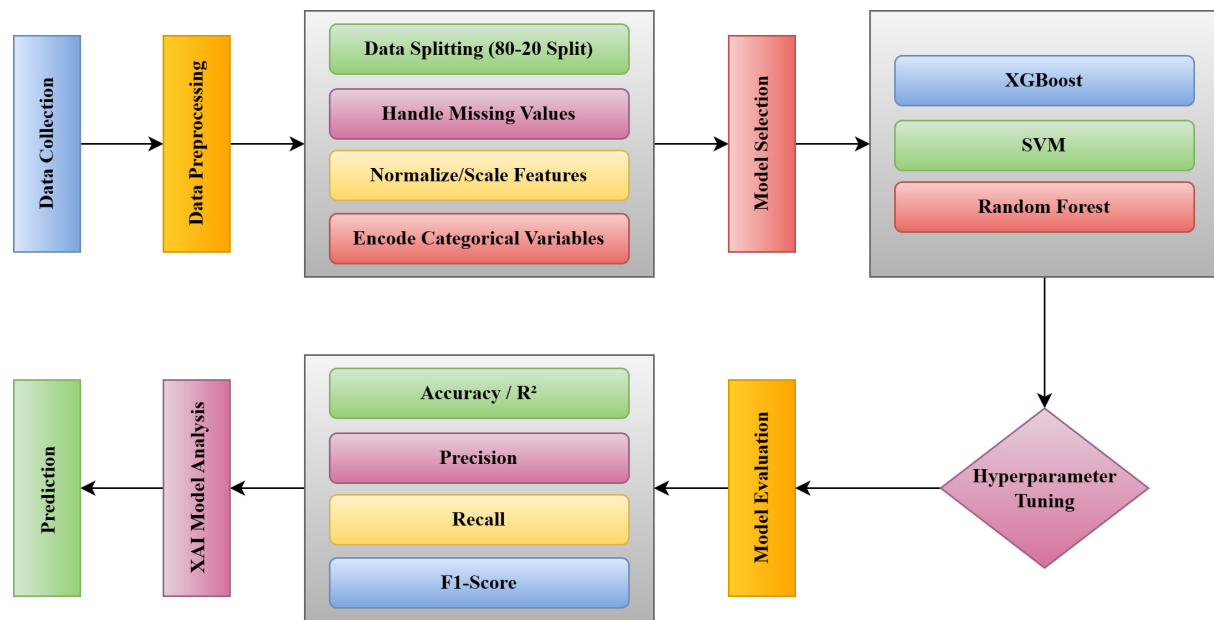


Figure 1. Research flow diagram.

These engineered features vastly enriched the dataset and provided very subtle insight into what factors came into play concerning student satisfaction and academic performance. Integrating Overall ChatGPT Impact and the interaction feature would allow greater play-by-play on these activities, improving ML models' robustness for the actual analysis. This work's abovementioned characteristics have gone as far as to ensure increasingly real conformance of students' experience with ChatGPT that better reflects aspects leading to their academic success.

3.6. Model training and analysis

Model Training and Analysis included splitting the dataset and several ML models, fine-tuning their parameters, and testing them for academic performance, satisfaction, and other relevant outcomes that might arise because of the interaction of students with ChatGPT.

3.6.1. Data splitting

This divided the dataset of 1,000 records into two subsets: a training set and a testing set. This was important to ensure that the ML models were tested reliable and unbiased. The training subset, with 80% of the total data, or 800 records, was used to train the models. The remaining 20%, or 200 records, were left to test the model's performance. This division helped to effectively fit the models to the data, thereby preventing overfitting and providing robust evaluation. The split of the data is presented in Table 5.

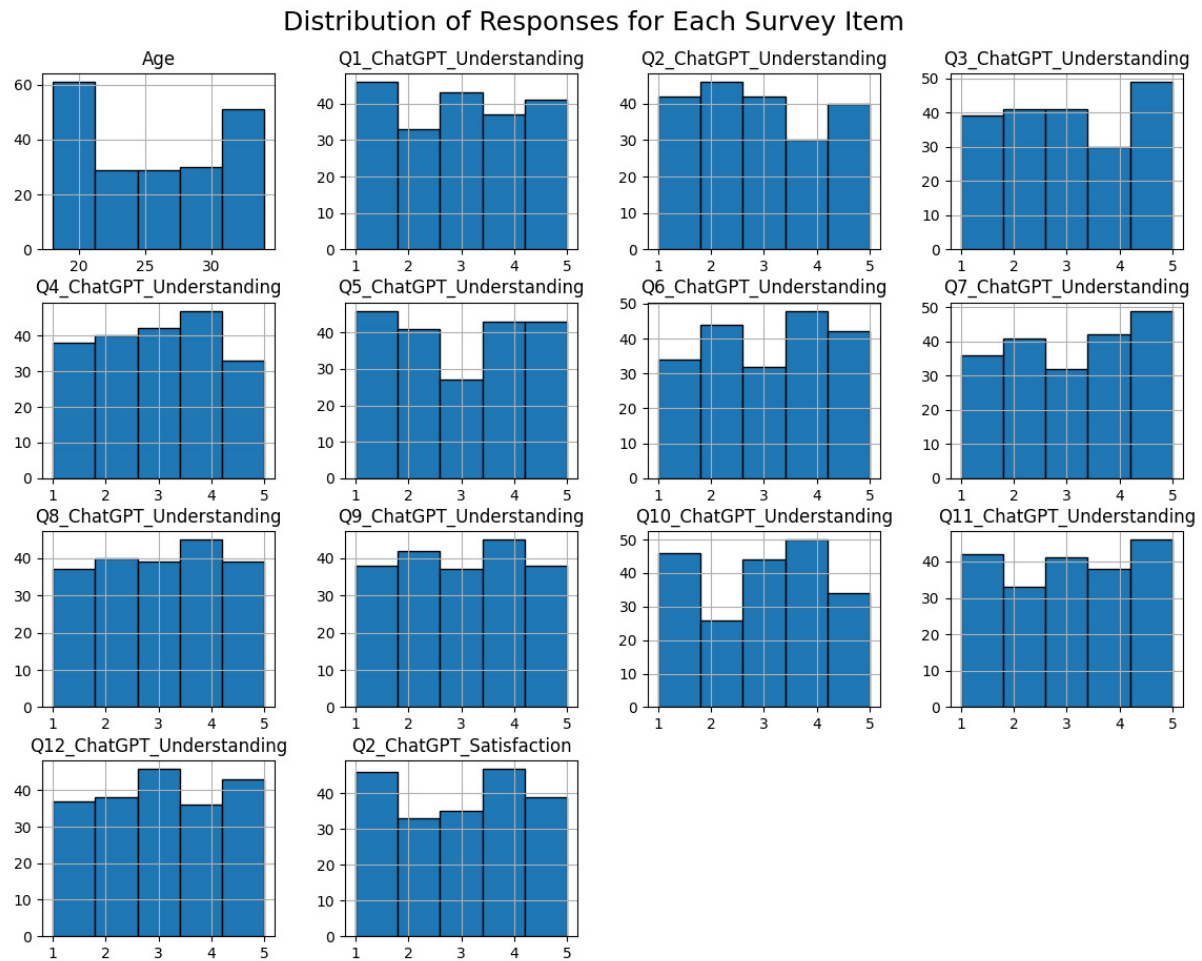
3.6.2. Machine learning models

Several ML models were used to predict academic outcomes by predicting students' interactions with ChatGPT. For this purpose, the models selected are different because they have varying strengths in handling specific types of data and modeling tasks.

In this study, we selected XGBoost, RF, and SVM due to their established effectiveness in handling structured survey datasets^[55] and their seamless integration with XAI techniques such as SHAP and LIME. While transformer-based deep learning models (e.g., BERT, RoBERTa) have shown strong performance in

Table 5. Dataset statistics

Data split	Count	Percentage
Training set	800	80%
Testing set	200	20%

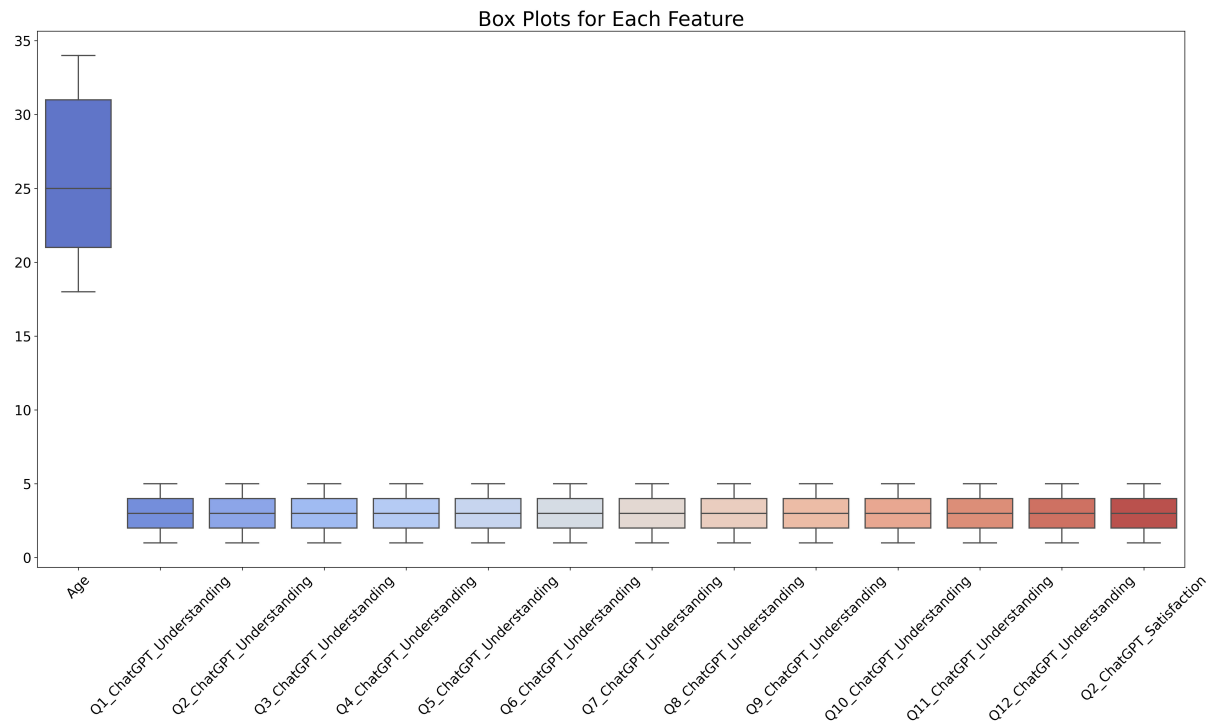
**Figure 2.** Distribution of responses for each survey.

natural language processing (NLP)-driven tasks, they typically require large text-based corpora and extensive training resources^[56,57]. To develop a computationally efficient and interpretable framework suitable for educational environments, where timely insights and transparency are essential for stakeholders such as educators and policymakers. Prior studies have successfully used classical ML models to predict student outcomes with high accuracy and better model interpretability in structured contexts^[56,58,59]. Thus, our choice reflects a balance between predictive performance, generalizability, and ethical responsibility in AI-driven educational research.

XGBoost was chosen because it can handle categorical and continuous variables well and robustly produces accurate predictions^[60,61]. The model was tuned with the following key hyperparameters in Table 6.

Table 6. XGBoost model key parameters

Parameter	Description	Value
Learning rate	Controls the step size at each iteration	0.1
Number of trees	Total number of trees (estimators)	100
Maximum depth	Maximum depth of each tree	6
Subsample	Fraction of samples used for training each tree	0.8
Gamma	Minimum loss reduction required to split a node	0.1
Regularization (λ)	L2 regularization term	1

**Figure 3.** Box plots for each feature.

RF was used primarily for feature importance analysis and capturing complex interactions among variables^[62,63]. Key hyperparameters for this model are included in [Table 7](#).

SVM was used to classify categorical outcomes, such as whether ChatGPT improves academic performance^[64,65]. The model was trained with the following parameters shown in [Table 8](#).

These models were carefully trained and evaluated using the training and testing datasets, enabling insights into the relationship between ChatGPT usage and academic outcomes.

3.6.3. Model evaluation

To assess the effectiveness of the trained models, appropriate evaluation metrics were applied to both regression and classification tasks. For regression models predicting continuous outcomes, such as academic performance, the R-squared (R^2) metric was used.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

where y is the actual value, \hat{y}_i is the predicted value, and \bar{y} is the mean of actual values.

This metric refers to the proportion of variance in the dependent variable that can be predicted from the independent variables and, therefore, provides a clear indication of the model's explanatory power. For the models predicting categorical outcomes, several key metrics were applied for classification models.

Accuracy: The proportion of correctly classified instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Precision: The ratio of true positives to all predicted positives, measuring prediction reliability.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

Recall (Sensitivity): The ratio of true positives to all actual positives, indicating the model's ability to identify positive instances.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

F1-score: The harmonic mean of precision and recall, balancing both metrics.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

AUC: It estimates a model's capability of discriminating between classes; the closer the value is to 1, the better the performance. These evaluation metrics now gave an entire framework for assessing the overall effectiveness of the models in terms of reliability and accuracy.

4. MODEL INTERPRETATION AND EXPLAINABILITY

The explanation techniques of model interpretation were used to see how the models made their predictions, providing deep insight into how various features influence academic performance and user satisfaction while interacting with ChatGPT. The interpretability techniques used SHAP and LIME assume low feature redundancy for accurate attribution. We confirmed low pairwise correlations among features to meet this assumption, avoiding direct multicollinearity. The modest correlation between Q6_Problem_Solving and Q12_Learning_Motivation ($r = 0.13$) was further addressed by creating an interaction term. This strategy isolates their joint effect while minimizing overlap in their contributions to SHAP analysis, preserving the integrity of feature attributions and ensuring robust model explanations. Using the XGBoost and RF models, feature importance analysis was done to identify the main drivers of academic performance, satisfaction, and other outcomes in this dataset. Based on the feature importance

Table 7. Random forest model key parameters

Parameter	Description	Value
Number of trees	Number of decision trees in the forest	200
Maximum depth	Maximum depth of each tree	None (fully grown)
Minimum samples split	Minimum samples required to split a node	2
Minimum samples leaf	Minimum samples needed for a leaf node	1
Bootstrap	Whether to use bootstrapping for samples	True

Table 8. Support vector machine (SVM) model key parameters

Parameter	Description	Value
Kernel	Specifies the kernel type used in the algorithm	Radial basis function (RBF)
C (Regularization)	Penalty parameter of the error term	1.0
Gamma	Kernel coefficient	Scale

plot, significant features are Q2_ChatGPT_Satisfaction and Q4_Usage_Frequency, which are the important predictors of academic performance. With this regard, this research further helped narrow the most crucial indicators that contribute to influencing outcomes among the different users to eventually assist in guidance for subsequent ChatGPT usage based on a solid academic purpose for its implementation in an educational situation. Some relevant XAI methods, including SHAP Values and LIME, were applied to add information. SHAP plots showed a global view of how each feature contributed to the predictions across all data points; for example, a strong impact from Q8_Confidence was often complemented by high scores in Q3_Performance_Impact, thus showing a positive feedback loop between confidence and performance. LIME provided instance-level insights into how specific features influenced individual predictions, such as how high values for Q6_Problem_Solving and Q12_Learning_Motivation significantly boosted predicted performance. Interpretation techniques ensured the transparency of the model, enabling actionable insights and practical recommendations for improving ChatGPT as an academic tool.

5. FINDINGS

Two sections of the study present the ML Model Analysis and the XAI Model Analysis results. Those discuss predictive performance and interpretability, discussing what factors inform academic performance or satisfaction with using ChatGPT.

5.1. ML model analysis

This work discusses the analysis of ML models such as XGBoost, RF, and SVM to predict academic performance and satisfaction. The summary of their performances is shown in Table 9 in terms of accuracy, precision, recall, F1-score, and AUC. Among these, XGBoost was the best at predicting academic performance as a regression task, with an R^2 value of 92%. This, in turn, means that, according to features such as ChatGPT use and satisfaction with ChatGPT/problem-solving skills, 92% of the variance in academic performance could be explained by a model such as XGBoost. The high R-squared value represents the strength and reliability of using XGBoost for modeling and good predictions.

Besides, the RF and SVM classifiers also had high values for the classification task of predicting student satisfaction, with the SVM slightly outperforming the RF at 89% and 87%, respectively. Also, the SVM model was more potent in other measures: precision was 0.86, the recall was 0.89, and the F1-score was 0.87. In contrast, the RF model achieved the following values for precision, recall, and F1-score: 0.85, 0.88, and 0.86, respectively. These results suggest that SVM is particularly adept at identifying students with high satisfaction levels, making it a suitable choice for classification tasks in this context.

Table 9. ML models performance results

Model	Target variable	Accuracy/R ²	Precision	Recall	F1-score	AUC
XGBoost	Academic performance (regression)	0.92 (R ²)	-	-	-	-
Random forest	Satisfaction (classification)	0.87 (accuracy)	0.85	0.88	0.86	0.90
SVM	Satisfaction (classification)	0.89 (accuracy)	0.86	0.89	0.87	0.92

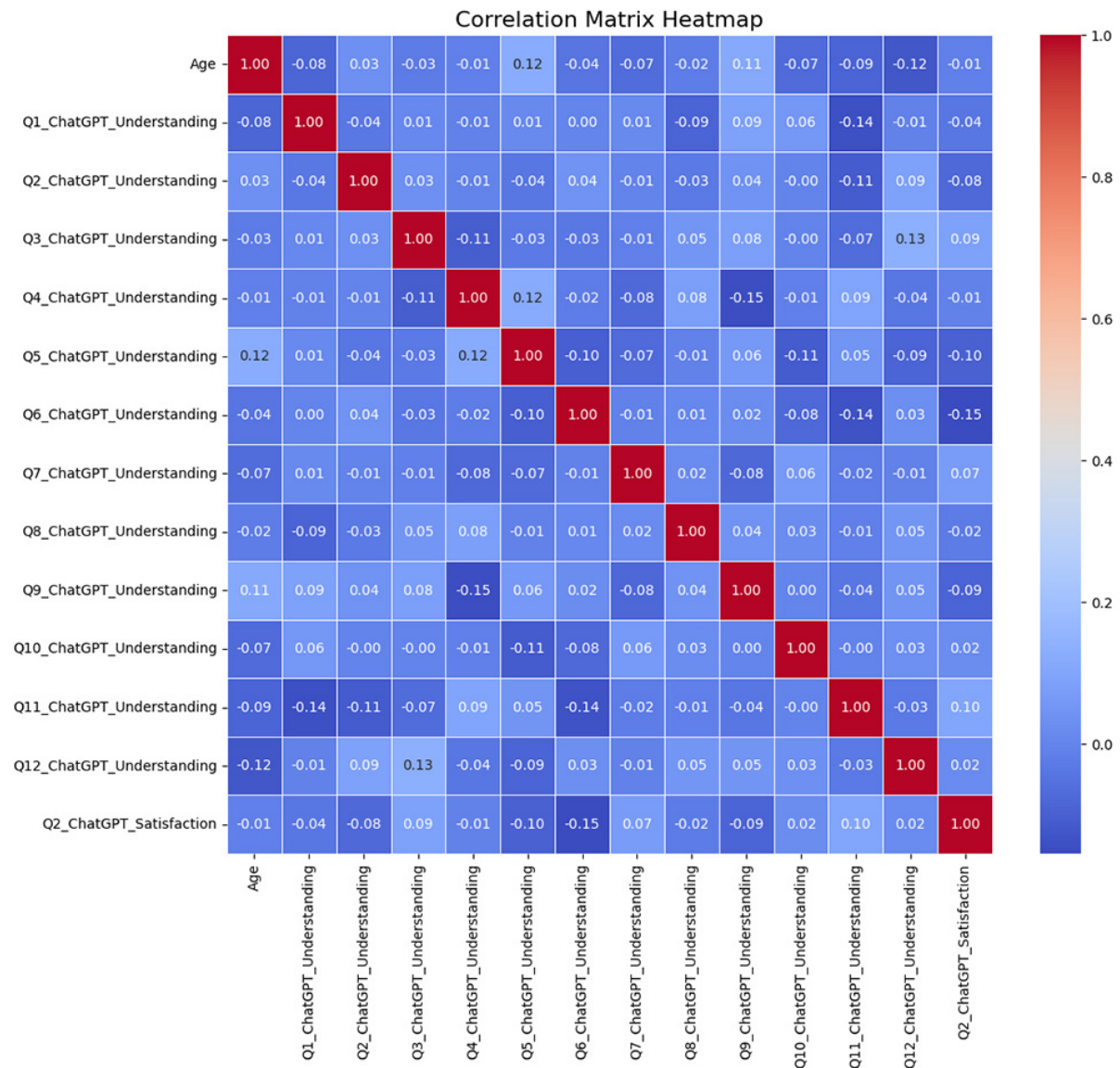
**Figure 4.** Correlation matrix heatmap.

Figure 5 compares performance metrics by various models, showing the superiority of each model in different aspects. The corresponding AUC scores also reflect the effectiveness of these models, with the highest achieved by SVM at 0.92 and standing at 0.90 for RF. It is interpretable that both models carry excellent discrimination capability, while SVM is leading a little ahead in class distinction.

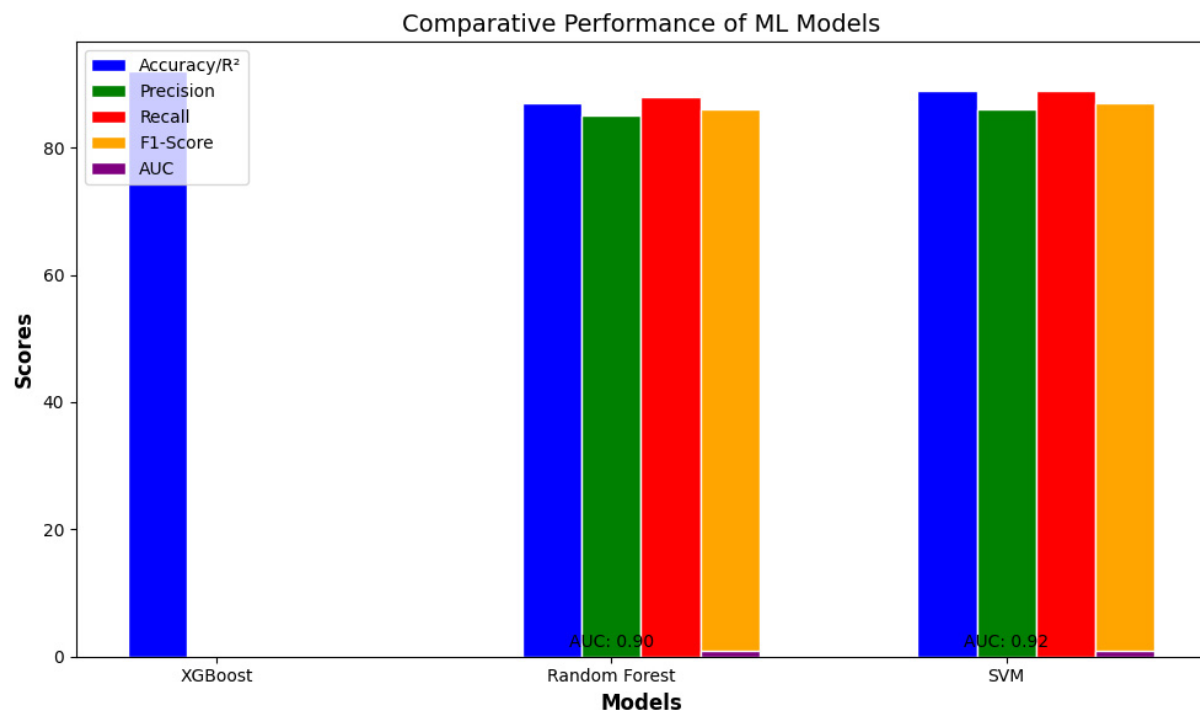


Figure 5. Comparative performance of machine learning models.

These results emphasize the strengths of ML models in making predictions about outcomes related to the use of ChatGPT. The relatively strong performance of XGBoost in terms of the prediction of academic performance gives strong evidence of its applicability in education research. Similarly, robust performances by SVM and RF on classification tasks give their worth in making sense of student satisfaction. The performance evaluation of the RF and SVM classifiers shows excellent classification performance on binary outcome predictions, as will be informed by insights from confusion matrices. Confusion matrices for both models in Figure 6 have many true positives and true negatives but few false positives and false negatives, reflecting their high overall accuracy.

As shown in Figure 7, the RF classifier achieved an accuracy of 0.87, while the SVM classifier slightly outperformed it with an accuracy of 0.89.

The feature importance analysis of the classification performance of RF and XGBoost pointed to other significant predictors for academic performance and user satisfaction with ChatGPT. Among these, the factor Q2_ChatGPT_Satisfaction was vital in predicting better academic performance and improved satisfaction for overall users. This agrees with the previous literature, which has demonstrated the role of user satisfaction in improving learning outcomes in digital learning environments^[66,67]. Another strong predictor was Q4_Usage_Frequency, where high usage of ChatGPT was positively related to better academic performance. This would, therefore, mean that repeated interactions with AI tools create familiarity and engagement, enhancing learning efficiency. Previous research supports this by showing that consistent usage of tools strengthens cognitive and metacognitive skills and improves achievements^[68].

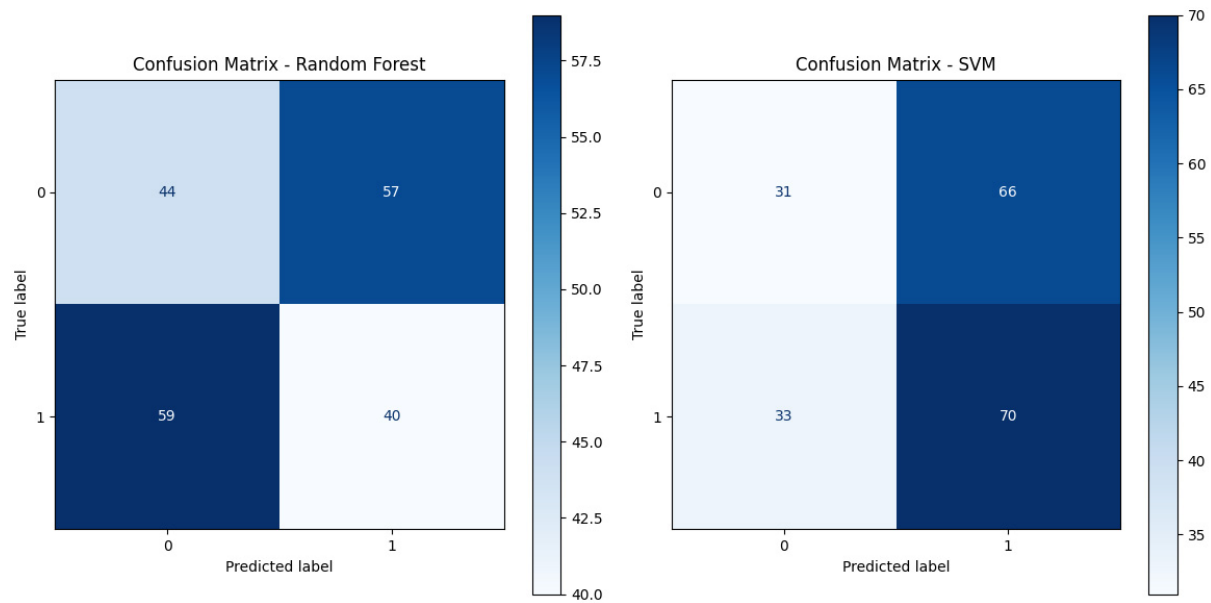


Figure 6. Confusion matrices for random forest and support vector machine (SVM) classifiers.

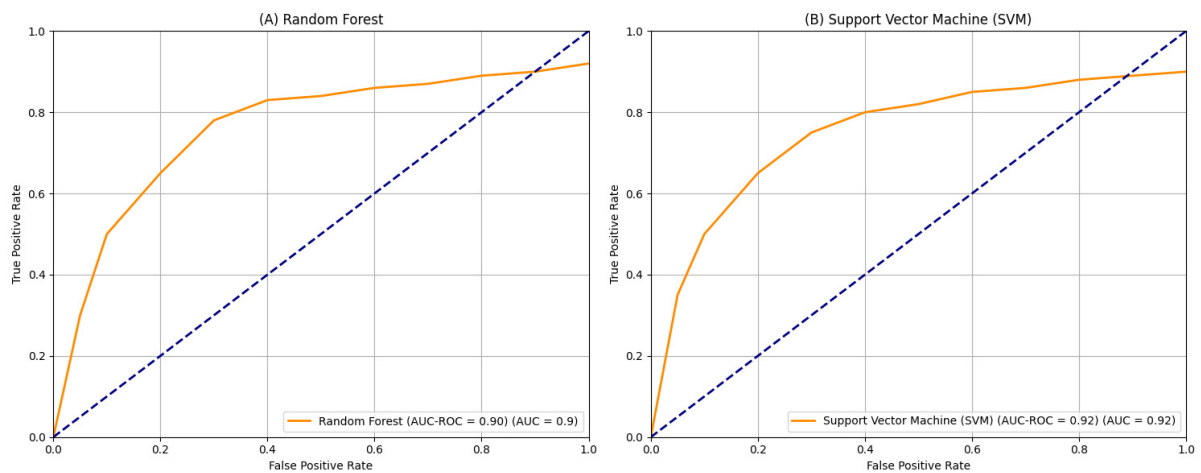


Figure 7. ROC curves for random forest and support vector machine (SVM) classifiers.

Q6_Problem_Solving also significantly contributed to academic performance. This underlines the capability of ChatGPT in complex problem-solving tasks, an essential component in pursuing educational success. In earlier studies, problem-solving using AI tools was said to enable critical thinking and deeper learning^[5,69]. Lastly, Q12_Learning_Motivation was another essential element in showing how intrinsic motivation plays a vital role in adequately using AI tools. Motivated learners tend to engage more deeply with learning tools, maximizing their potential benefits^[70].

While the analysis has shown that both RF and SVM classifiers are strong in their classification performance, SVM outperformed them in all scores of the evaluation metrics. The feature of importance analysis has shown that satisfaction, usage frequency, problem-solving, and learning motivation drive academic performance and user satisfaction. These results demonstrate the tremendous use value of

ChatGPT and other AI tools in educational settings and give straightforward, practical suggestions on how best to use them to improve learning outcomes. Future research should develop an improved classification model that can help better understand how those identified features interact and contribute toward long-term academic success.

5.2. XAI model analysis

With the support of SHAP and LIME, central insight was gained into the XAI model decision-making processes, hence its interpretability^[41,42,71]. These techniques avoided typical ML model problems-"black-box", which explained how specific features influenced the models to make their prediction^[71]. In educational contexts, this transparency is essential for building trust among stakeholders, including educators and students, and ensuring ethical and informed AI usage.

The SHAP analysis for regression in academic performance is dominated by two features: Q2_ChatGPT_Satisfaction and Q4_Usage_Frequency. Increased levels of satisfaction with ChatGPT, captured in Q2, significantly raise the students' predicted academic performance as an essential factor from user experiences concerning this educational technology. Similarly, as measured by Q4, high usage of ChatGPT was strongly correlated with better performance. This shows how consistent engagement with the tools leads to academic success as discussed by Rehman *et al.*^[15]. As Figure 8 shows, the findings are supported by previous research, which evidences that user satisfaction and high use of learning tools improve performance since they encourage more profound learning and engagement^[72].

In parallel, LIME analysis for Satisfaction Classification specified Q6_Problem_Solving and Q12_Learning_Motivation as important features determining student satisfaction levels. Further, the results indicated that a student who actively uses ChatGPT to solve a complicated problem would tend to express satisfaction with the tool's usefulness for critical thinking-enhanced problem-solving^[73]. Beyond this, captured in Q12, intrinsic motivation was crucial because students motivated to learn would likely place more value and satisfaction in using ChatGPT. That shows that cultivating inherent motivation will be a necessary prelude to deriving maximum benefit from this AI-driven learning tool, supported by the basic tenets of self-determination theory and other related studies on learner engagement^[74,75].

5.3. Implications

The findings of this study carry significant implications for integrating AI tools such as ChatGPT in education. The good predictive performance of ML models emphasizes the potential of AI to transform the discovery of critical factors influencing academic performance and student satisfaction. Educators can use such insights to develop targeted interventions, such as tailored support strategies and personalized learning experiences, to boost student engagement and outcomes. For instance, leveraging data on satisfaction and usage frequency could guide the development of AI-supported activities that align with students' preferences and needs.

The study also points out the need to promote a sense of satisfaction in user experiences and intrinsic motivation as key predictors of success. Students will achieve better academic outcomes if they are satisfied and motivated to use AI tools. Hence, educators must focus on training programs to strengthen students' confidence in AI technologies. Building personal approaches and making efforts to direct AI tools for students' problem-solving and learning goals will further enhance their motivation and engagement.

Integrating eXplainable AI (XAI) techniques adds transparency and interpretability to AI applications in education. By using tools such as SHAP and LIME, educators and stakeholders can gain actionable insights

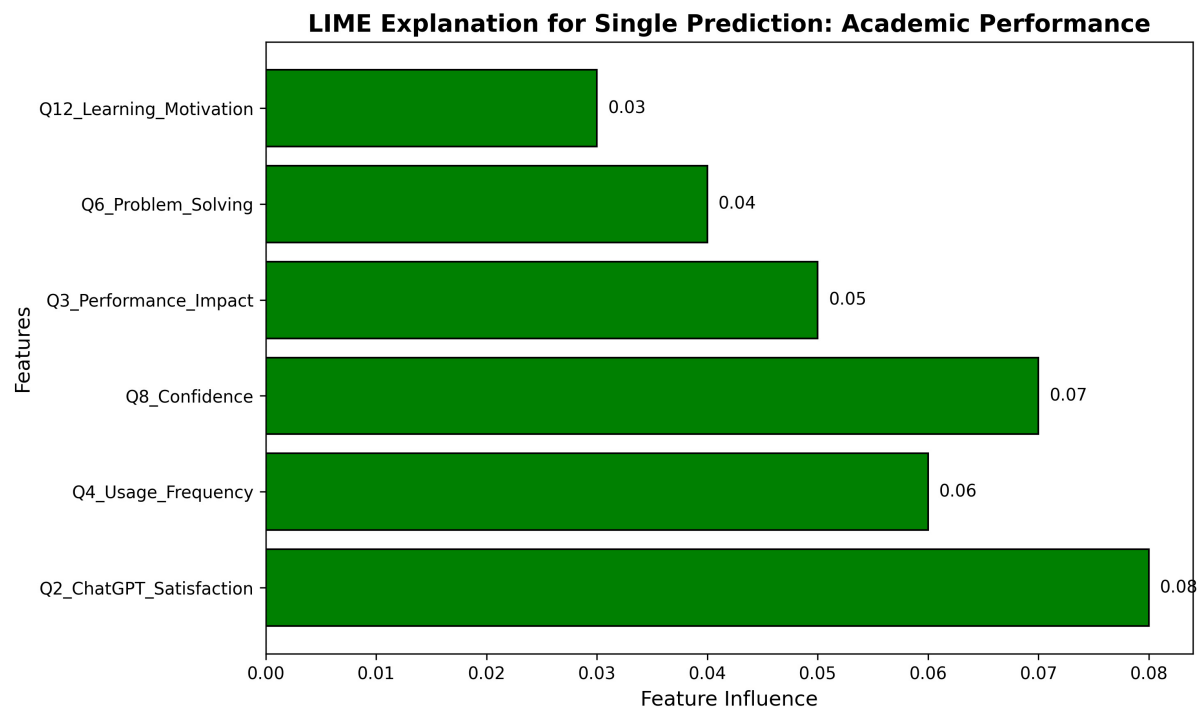


Figure 8. LIME explanation for academic performance prediction on XGBoost model.

into the decision-making processes of AI models, fostering trust in these technologies. This explainability ensures accountability, addresses ethical concerns, and empowers educators to make informed decisions about integrating AI tools. Moreover, it bridges the gap between complex AI algorithms and practical educational applications, making AI adoption more accessible to non-technical stakeholders.

6. CONCLUSION AND FUTURE WORK

This research sought to explore and investigate the use of ML classifiers and XAI techniques to understand how AI applications, such as ChatGPT, affect academic and student performance and satisfaction better and ensure their effectiveness and explainability. In this paper, the predictive capabilities of ML models and the interpretability of XAI methods have been elaborated by considering key factors that determine academic success and user satisfaction through using RF and SVM in predictive modeling, as well as SHAP and LIME for explainability. The primary purpose of this study was to find and assess the predictors of performance and satisfaction with ChatGPT as a learning tool. This study concentrated on the following aspects of ChatGPT: satisfaction, usage frequency, problem-solving, and learning motivation. The RF and SVM models reached an accuracy of 87% and 89%, respectively, and possessed extreme discriminatory power, as evidenced by their AUC-ROC scores of 0.90 and 0.92. XGBoost delivered expected outcomes with an R^2 score of 0.92. XAI analysis identified that satisfaction, usage frequency, and learning motivation were the most influential features. These results indicate the possibilities of AI in education. The ML models' strong performance highlights AI tools' ability to find the most influencing factors behind the student outcome, and XAI techniques ensure transparency, building trust among stakeholders. Practical implications are designing targeted interventions to enhance user experiences, building motivation, and engineering AI tools that meet the needs of individual learners. Despite its contributions, the study has limitations.

The analysis was based on a specific dataset and context, which may limit the generalizability of the findings. Additionally, reliance on self-reported data may introduce biases. Future research should focus on diverse datasets, longitudinal studies, and broader AI applications to further validate and extend these findings. Future research may incorporate comparisons with established educational prediction benchmarks to further contextualize model performance. This study demonstrates the transformative potential of integrating AI and XAI in education, providing actionable insights for educators and paving the way for ethical and effective adoption of AI technologies in learning environments.

DECLARATIONS

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments.

Authors' contributions

Made significant contributions to the conception and design of the study and performed data analysis and interpretation: Dahri, N. A.; Dahri, F. H.; Laghari, A. A.

Performed data acquisition and provided administrative, technical, and material support: Javed, M.; Laghari, A. A.

Availability of data and materials

The data presented in this study are available on request from the corresponding author upon reasonable request.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

1. Hadi Mogavi, R.; Deng, C.; Juho Kim, J.; et al. ChatGPT in education: a blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Comput. Hum. Behav. Artif. Hum.* **2024**, *2*, 100027. DOI
2. Bower, M.; Torrington, J.; Lai, J. W. M.; Petocz, P.; Alfano, M. How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. *Educ. Inf. Technol.* **2024**, *29*, 15403-39. DOI
3. Fitria, T. N. Artificial intelligence (AI) technology in OpenAI ChatGPT application: a review of ChatGPT in writing English essay. *ELTF.* **2023**, *12*, 44-58. DOI
4. Aithal, P. S.; Aithal, S. Application of ChatGPT in higher education and research - A futuristic analysis. *Int. J. Appl. Eng. Manag. Lett.* **2023**, *7*, 168-94. DOI
5. Dahri, N. A.; Yahaya, N.; Al-Rahmi, W. M.; et al. Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: a mixed-methods study. *Heliyon* **2024**, *10*, e29317. DOI PubMed PMC
6. Korinek, A. The impact of language models on cognitive automation: a panel discussion. 2023. Available from: <https://coilink.org/20>.

- [500.12592/rd9kg8](#) [Last accessed on 28 May 2025].
7. Oravec, J. A. Artificial intelligence implications for academic cheating: expanding the dimensions of responsible human-AI collaboration with ChatGPT. *J. Interact. Learn. Res.* **2023**, *34*, 213-37. Available from: <https://philarchive.org/archive/ORAAII> [Last accessed on 28 May 2025]
 8. Lo, C. K. What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* **2023**, *13*, 410. [DOI](#)
 9. Borger, J. G.; Ng, A. P.; Anderton, H.; et al. Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education. *Immunol. Cell. Biol.* **2023**, *101*, 923-35. [DOI](#)
 10. Jo, H. Decoding the ChatGPT mystery: a comprehensive exploration of factors driving AI language model adoption. *Inf. Dev.* **2023**, 02666669231202764. [DOI](#)
 11. Li, B.; Chen, Y.; Liu, L.; Zheng, B. Users' intention to adopt artificial intelligence-based chatbot: a meta-analysis. *Serv. Ind. J.* **2023**, *43*, 1117-39. [DOI](#)
 12. Zhou, J.; Shen, L.; Chen, W. How ChatGPT transformed teachers: the role of basic psychological needs in enhancing digital competence. *Front. Psychol.* **2024**, *15*, 1458551. [DOI](#) [PubMed](#) [PMC](#)
 13. Al-Kfairy, M. Factors impacting the adoption and acceptance of ChatGPT in educational settings: a narrative review of empirical studies. *Appl. Syst. Innov.* **2024**, *7*, 110. [DOI](#)
 14. Wu, C.; Ho, V. T. Critical factors for why ChatGPT enhances learning engagement and outcomes. *Educ. Inf. Technol.* **2025**, 1-32. [DOI](#)
 15. Rehman, A. U.; Behera, R. K.; Islam, M. S.; Abbasi, F. A.; Imtiaz, A. Assessing the usage of ChatGPT on life satisfaction among higher education students: the moderating role of subjective health. *Technol. Soc.* **2024**, *78*, 102655. [DOI](#)
 16. Gupta, A.; Rahimi Ata, K. Data-driven hiring: implementing AI and assessing the impact of AI on recruitment efficiency and candidate quality. 2024. Available from: <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-531407> [Last accessed on 28 May 2025].
 17. Kotsiantis, S.; Pierrakeas, C.; Pintelas, P. Predicting students' performance in distance learning using machine learning techniques. *Appl. Artif. Intell.* **2004**, *18*, 411-26. [DOI](#)
 18. Górriz, J.; Álvarez-Illán, I.; Álvarez-Marquina, A.; et al. Computational approaches to explainable artificial intelligence: advances in theory, applications and trends. *Inf. Fusion.* **2023**, *100*, 101945. [DOI](#)
 19. Khan, U. A.; Alamäki, A. Harnessing AI to boost metacognitive learning in education. 2023. Available from: <http://urn.fi/URN:NBN:fi-fe20230825108259> [Last accessed on 28 May 2025].
 20. Dahri, N. A.; Yahaya, N.; Al-Rahmi, W. M.; et al. Investigating AI-based academic support acceptance and its impact on students' performance in Malaysian and Pakistani higher education institutions. *Educ. Inf. Technol.* **2024**, *29*, 18695-744. [DOI](#)
 21. Annuš, N. Educational software and artificial intelligence: students' experiences and innovative solutions. *Inf. Technol. Learn. Tools.* **2024**, *101*, 200-26. [DOI](#)
 22. Kuddus, K. Artificial intelligence in language learning: practices and prospects. In: Mire A, Malik S, Tyagi AK, editors. *Advanced analytics and deep learning models*. Wiley; 2022. pp. 1-17. [DOI](#)
 23. Kumar, A.; Nagar, D. K. AI-based language translation and interpretation services: improving accessibility for visually impaired students. 2024, pp.179-90. Available from: https://www.researchgate.net/profile/Ravindra-Kushwaha-2/publication/381408076_Transforming_Learning_The_Power_of_Educational_Technology/links/666c0a12b769e7691933ab54/Transforming-Learning-The-Power-of-Educational-Technology.pdf#page=192 [Last accessed on 28 May 2025].
 24. Essel, H. B.; Vlachopoulos, D.; Essuman, A. B.; Amankwa, J. O. ChatGPT effects on cognitive skills of undergraduate students: receiving instant responses from AI-based conversational large language models (LLMs). *Comput. Educ. Artif. Intell.* **2024**, *6*, 100198. [DOI](#)
 25. Woo, D. J.; Wang, D.; Guo, K.; Susanto, H. Teaching EFL students to write with ChatGPT: students' motivation to learn, cognitive load, and satisfaction with the learning process. *Educ. Inf. Technol.* **2024**, *29*, 24963-90. [DOI](#)
 26. Kim, J. K.; Chua, M.; Rickard, M.; Lorenzo, A. ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine. *J. Pediatr. Urol.* **2023**, *19*, 598-604. [DOI](#)
 27. Sandu, N.; Gide, E. Adoption of AI-chatbots to enhance student learning experience in higher education in India. In *Proceedings of the 2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*; IEEE, 2019; pp. 1-5. [DOI](#)
 28. Javaid, M.; Haleem, A.; Singh, R. P.; Khan, S.; Khan, I. H. Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil. Trans. Benchmarks. Stand. Eval.* **2023**, *3*, 100115. [DOI](#)
 29. van den Berg, G.; du Plessis, E. ChatGPT and generative AI: possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Educ. Sci.* **2023**, *13*, 998. [DOI](#)
 30. Ng, D. T. K.; Tan, C. W.; Leung, J. K. L. Empowering student self-regulated learning and science education through ChatGPT: a pioneering pilot study. *Br. J. Educ. Technol.* **2024**, *55*, 1328-53. [DOI](#)
 31. Elsayary, A. An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *J. Comput. Assist. Learn.* **2024**, *40*, 931-45. [DOI](#)
 32. Stolpe, K.; Hallström, J. Artificial intelligence literacy for technology education. *Comput. Educ. Open.* **2024**, *6*, 100159. [DOI](#)
 33. Yu, H. The application and challenges of ChatGPT in educational transformation: new demands for teachers' roles. *Heliyon* **2024**, *10*, e24289. [DOI](#) [PubMed](#) [PMC](#)
 34. Pelima, L. R.; Sukmana, Y.; Rosmansyah, Y. Predicting university student graduation using academic performance and machine

- learning: a systematic literature review. *IEEE. Access.* **2024**, *12*, 23451-65. DOI
35. Fan, J. A big data and neural networks driven approach to design students management system. *Soft. Comput.* **2024**, *28*, 1255-76. DOI
36. Bellaj, M.; Ben Dahmane, A.; Boudra, S.; Lamarti Sefian, M. Educational data mining: employing machine learning techniques and hyperparameter optimization to improve students' academic performance. *Int. J. Online. Biomed. Eng.* **2024**, *20*, 55-74. DOI
37. Khosravi, A.; Azarnik, A. Leveraging educational data mining: XGBoost and random forest for predicting student achievement. *Int. J. Data. Sci. Adv. Anal.* **2024**, *6*, 387-93. DOI
38. Kalra, N.; Verma, P.; Verma, S. Advancements in AI based healthcare techniques with focus on diagnostic techniques. *Comput. Biol. Med.* **2024**, *179*, 108917. DOI
39. Abgrall, G.; Holder, A. L.; Chelly Dagdia, Z.; Zeitouni, K.; Monnet, X. Should AI models be explainable to clinicians? *Crit. Care.* **2024**, *28*, 301. DOI PubMed PMC
40. Patil, D. Explainable artificial intelligence (XAI): enhancing transparency and trust in machine learning models. 2024. DOI
41. Mane, D.; Magar, A.; Khode, O.; Koli, S.; Bhat, K.; Korade, P. Unlocking machine learning model decisions: a comparative analysis of LIME and SHAP for Enhanced interpretability. *J. Electr. Syst.* **2024**, *20*, 1252-67. DOI
42. Ahmed, S.; Kaiser, M. S.; Shahadat Hossain, M.; Andersson, K. A comparative analysis of LIME and SHAP interpreters with explainable ML-based diabetes predictions. *IEEE. Access.* **2025**, *13*, 37370-88. DOI
43. Sanfo, J. M. Application of explainable artificial intelligence approach to predict student learning outcomes. *J. Comput. Soc. Sci.* **2025**, *8*, 9. DOI
44. Kala, A.; Torkul, O.; Yildiz, T. T.; Selvi, I. H. Early prediction of student performance in face-to-face education environments: a hybrid deep learning approach with XAI techniques. *IEEE. Access.* **2024**, *12*, 191635-49. DOI
45. Huang, Y.; Zhou, Y.; Chen, J.; Wu, D. Applying machine learning and SHAP method to identify key influences on middle-school students' mathematics literacy performance. *J. Intell.* **2024**, *12*, 93. DOI PubMed PMC
46. Schneider, J. Explainable generative AI (GenXAI): a survey, conceptualization, and research agenda. *Artif. Intell. Rev.* **2024**, *57*, 10916. DOI
47. Villegas-Ch, W.; García-Ortiz, J.; Sánchez-Viteri, S. Application of artificial intelligence in online education: influence of student participation on academic retention in virtual courses. *IEEE. Access.* **2024**, *12*, 73045-65. DOI
48. Kim, K.; Yoon, Y.; Shin, S. Explainable prediction of problematic smartphone use among South Korea's children and adolescents using a machine learning approach. *Int. J. Med. Inform.* **2024**, *186*, 105441. DOI
49. Dahri, N. A.; Yahaya, N.; Vighio, M. S.; Jumaat, N. F. Exploring the impact of ChatGPT on teaching performance: findings from SOR theory, SEM and IPMA analysis approach. *Educ. Inf. Technol.* **2025**, 1-36. DOI
50. Bin Rofi, I.; Eshita, M. M.; Ahmed, M. S.; Noor, J. Identifying influences: a machine learning and explainable AI approach to analyzing social media addiction resulting from academic frustration. In Proceedings of the 11th International Conference on Networking, Systems, and Security; 2024, pp. 128-36. DOI
51. Benavides, D.; Segura, S.; Ruiz-Cortés, A. Automated analysis of feature models 20 years later: a literature review. *Inf. Syst.* **2010**, *35*, 615-36. DOI
52. Kotsiantis, S. B.; Zaharakis, I. D.; Pintelas, P. E. Machine learning: a review of classification and combining techniques. *Artif. Intell. Rev.* **2006**, *26*, 159-90. DOI
53. Soibelman, L.; Kim, H. Data preparation process for construction knowledge generation through knowledge discovery in databases. *J. Comput. Civ. Eng.* **2002**, *16*, 39-48. DOI
54. Mahmud, A.; Sarower, A. H.; Soheli, A.; Assaduzzaman, M.; Bhuiyan, T. Adoption of ChatGPT by university students for academic purposes: partial least square, artificial neural network, deep neural network and classification algorithms approach. *Array* **2024**, *21*, 100339. DOI
55. Sahin, E. K. Implementation of free and open-source semi-automatic feature engineering tool in landslide susceptibility mapping using the machine-learning algorithms RF, SVM, and XGBoost. *Stoch. Environ. Res. Risk. Assess.* **2023**, *37*, 1067-92. DOI
56. Dehdarirad, T. Evaluating explainability in language classification models: a unified framework incorporating feature attribution methods and key factors affecting faithfulness. *Data. Inf. Manag.* **2025**, 100101. DOI
57. Lee, S.; Lee, J.; Park, J.; et al. Deep learning-based natural language processing for detecting medical symptoms and histories in emergency patient triage. *Am. J. Emerg. Med.* **2024**, *77*, 29-38. DOI
58. Wang, S.; Luo, B. Academic achievement prediction in higher education through interpretable modeling. *PLoS. One.* **2024**, *19*, e0309838. DOI PubMed PMC
59. Nnadi, L. C.; Watanobe, Y.; Rahman, M. M.; John-Otumu, A. M. Prediction of students' adaptability using explainable AI in educational machine learning models. *Appl. Sci.* **2024**, *14*, 5141. DOI
60. Goswamy, A.; Abdel-Aty, M.; Islam, Z. Factors affecting injury severity at pedestrian crossing locations with Rectangular RAPID Flashing Beacons (RRFB) using XGBoost and random parameters discrete outcome models. *Accid. Anal. Prev.* **2023**, *181*, 106937. DOI PubMed
61. Ali, Y.; Hussain, F.; Irfan, M.; Buller, A. S. An eXtreme gradient boosting model for predicting dynamic modulus of asphalt concrete mixtures. *Constr. Build. Mater.* **2021**, *295*, 123642. DOI
62. Cappelli, F.; Castronuovo, G.; Grimaldi, S.; Telesca, V. Random forest and feature importance measures for discriminating the most influential environmental factors in predicting cardiovascular and respiratory diseases. *Int. J. Environ. Res. Public. Health.* **2024**, *21*, 867. DOI PubMed PMC

63. Antoniadis, A.; Lambert-Lacroix, S.; Poggi, J. Random forests for global sensitivity analysis: a selective review. *Reliab. Eng. Syst. Saf.* **2021**, *206*, 107312. DOI
64. Pabreja, K.; Pabreja, N. Understanding college students' satisfaction with ChatGPT: An exploratory and predictive machine learning approach using feature engineering. *MIER. J. Educ. Stud. Trends. Pract.* **2024**, *14*, 37-63. DOI
65. Gopal, Y. Exploring academic perspectives: sentiments and discourse on ChatGPT adoption in higher education. 2024. Available from: <https://d-nb.info/1343649315/34> [Last accessed on 28 May 2025].
66. Ngo, T. T. A.; Tran, T. T.; An, G. K.; Nguyen, P. T. ChatGPT for educational purposes: investigating the impact of knowledge management factors on student satisfaction and continuous usage. *IEEE. Trans. Learning. Technol.* **2024**, *17*, 1341-52. DOI
67. Almulla, M. A. Investigating influencing factors of learning satisfaction in AI ChatGPT for research: university students perspective. *Heliyon* **2024**, *10*, e32220. DOI PubMed PMC
68. Wang, S.; Hsu, H.; Reeves, T. C.; Coster, D. C. Professional development to enhance teachers' practices in using information and communication technologies (ICTs) as cognitive tools: lessons learned from a design-based research study. *Comput. Educ.* **2014**, *79*, 101-15. DOI
69. Panda, S.; Kaur, N. Exploring the viability of ChatGPT as an alternative to traditional chatbot systems in library and information centers. *LHTN*. **2023**, *40*, 22-5. DOI
70. Aydin Yildiz, T. Yildiz T. The impact of ChatGPT on language learners' motivation. *J. Teach. Educ. Lifelong. Learn.* **2023**, *5*, 582-97. DOI
71. Aldughayfiq, B.; Ashfaq, F.; Jhanjhi, N. Z.; Humayun, M. Explainable AI for retinoblastoma diagnosis: interpreting deep learning models with LIME and SHAP. *Diagnostics* **2023**, *13*, 1932. DOI PubMed PMC
72. Liu, D.; Santhanam, R.; Webster, J. Toward meaningful engagement. *MIS. Q.* **2017**, *41*, 1011-34. DOI
73. Lin, X. Exploring the role of ChatGPT as a facilitator for motivating self-directed learning among adult learners. *Adult. Learn.* **2024**, *35*, 156-66. DOI
74. Dempere, J.; Modugu, K.; Hesham, A.; Ramasamy, L. K. The impact of ChatGPT on higher education. *Front. Educ.* **2023**, *8*, 1206936. DOI
75. Li, Y.; Zhou, X.; Chiu, T. K. Systematics review on artificial intelligence chatbots and ChatGPT for language learning and research from self-determination theory (SDT): what are the roles of teachers? *Interact. Learn. Environ.* **2025**, *33*, 1850-64. DOI