Shodh Sari-An International Multidisciplinary Journal

2026, Vol. 05, Issue 02, 145-154
ISSN: 2959-1376
DOI: https://doi.org/10.59231/SARI7919

Comparative Analysis of Regression Models for Predicting Student Satisfaction in AI-Assisted Learning

Kumar, Gulshan

Research Scholar (Department of Computer Science), Punjabi University, Patiala

Abstract

The effect of growth of technology can be seen through the rapid increase in the use of AI in every sphere of life. AI has become widely integrated into almost every domain of human activity, including the field of education, where its presence has increased significantly. Consequently, evaluating the level of student satisfaction towards AI assisted learning has emerged as an important area of research. For the study of this paper a dataset is used. The title of the dataset is "AI Assistant Usage in Student Life." This dataset has the record of 10,000 students which is showing the AI interaction session with various attributes. The objective of this paper is to predict Student satisfaction by using machine learning regression techniques. The dataset was processed using Python in Jupyter notebook. The dataset was divided into two parts i.e. training and testing sets in 75:25 ratio. On the dataset four regression algorithms were applied. This research paper shows that the regression-based machine learning models can analyze and predict student satisfaction effectively. The overall study supports the growing role of AI in education because the result shows that most of the students are satisfied after using AI assistants. This paper also tries to provide comparative study of all the regression models applied on the dataset. For this purpose, a confusion matrix has been generated which is divided into two categories labeled as satisfied and not satisfied. This study helps in making decisions about the use of AI in education.

Keywords: Artificial Intelligence, Machine Learning, Student Satisfaction, Regression Analysis, Confusion Matrix.

I. Introduction

The increasing growth of AI has introduced various possibilities in different fields including education. In this era AI powered assistants for instance visual tutors, chatbots and other automated systems have taken an

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

inseparable place in the life of students. Students are using these AI tools for various tasks such as studying, writing, making assignments and preparing projects [6]. Students are using these tools rapidly because of easy availability of internet and AI tools and due to this it is becoming important part of learning environment.

There is no doubt that the use of AI has increased greatly, but it is important to understand whether students are actually satisfied with AI assistants. Just using AI does not always mean that students are benefiting from it. Therefore, it becomes necessary to study how reliable and helpful AI tools really are. The study of satisfaction level of students is also important to know because if they really help the students then the ways should be found to use AI tools in better way [8]. But if these tools provide confusion or wrong answers then the use of AI tools should be restricted.

Earlier the student satisfaction was measured through surveys, feedback forms or interviews etc. No doubt these methods provide good results but these are time consuming too. Now a days with the availability of datasets, it has become possible and easy to analyze the behavior of students and their satisfaction level using data driven and machine learning. This study is trying to analyze the satisfaction of students by using dataset named "AI Assistant Usage in Student Life" [7]. The dataset is showing the information about student AI interaction and various details such as discipline, session duration, number of prompts asked, repeat usage and student satisfaction rating etc.

Regression is generally used technique used for predicting continuous value based on multiple independent variables. In relation to this study regression is helping in estimating the strength of each factor [9]. For the accomplishment of this study, multi linear regression, ridge regression, lasso regression and decision tree regression has been used. Every method has different pros and cons. After applying different regression algorithms, a comparative study has been done. Under this study MSE and R^2 score are used to find out how closely the predicted values match the actual satisfaction rating. This study helps in decision making in educational institutions in relation to the use of AI tools by knowing whether students are satisfied or not. A confusion matrix is formed to evaluate the efficiency of the models by

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

classifying students into two categories named satisfied and unsatisfied groups. The reason for working on this research paper is the increasing dependence of students on AI based tools for education purpose. Students are using AI tools for various academic activities because of which it becomes essential to know the level of accuracy and satisfaction of users.

Objectives:

a. Predicting student satisfaction with AI Assistants using Machine Learning Regression Model
b. Comparison of different Regression Algorithms in terms of reliability
c. Evaluate Regression outcomes with the help of Confusion matrix

Impact of the Study:

The impact of this research is significant for educational institutions. Instead of relying only on traditional methods such as surveys, feedback forms, and interviews, institutions can use machine learning techniques to analyze large datasets and measure student satisfaction more efficiently. This approach saves time and provides more objective and scalable evaluation. The findings of this study help institutions understand whether AI tools are improving the learning experience or creating confusion due to inaccurate or incomplete responses. Based on such analysis, administrators can make informed decisions regarding the proper implementation, regulation, or improvement of AI-based systems in academic environments.

Overall, the impact of this research lies in offering a reliable, data-driven framework for evaluating student satisfaction with AI assistants. It supports educational institutions, AI developers, and researchers in improving the quality, reliability, and effectiveness of AI-powered learning systems while ensuring that student learning remains the central focus.

II. Literature Review

As more people are using AI in different fields, the research interest in this issue is also increasing. The researchers are trying to explore that how AI tools are influencing the student performance, behavior and level of satisfaction. This segment of this paper is showing the important work done in relation to the use of AI and Machine Learning in education.

Early studies initially were focusing on the AI based tutoring system in improving the understanding of students. Those studies

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

showed that AI provides instant help to students, which was helping in providing instant solution to the problems [1], [2]. The studies reported that students using AI were more engaged in task.

After those researchers started analyzing the evaluation of student's behavior in digital learning environment by using data driven approaches. The findings of those studies highlighted the importance of usage of behavior in understanding student's experience [3], [4].

Then due to the immense growth of AI and ML, various Machine Learning methods have been widely adopted for predicting academic performance and student engagement. Regression models, classification algorithms and other ML techniques are used to analyze the datasets. Many researchers argued that ML models outperform traditional statistical methods in predictive accuracy when large datasets are available [9]. The previous studies examined the effectiveness of AI tools. Those studies found that AI assistants are helpful for various tasks for instance answering the questions, providing help in writing and making assignments etc. In today's era the emphasis is shifting to evaluate the level of satisfaction of students using AI tools [5], [6].

III. Methodology

This experiment was conducted on a dataset which was acquired from Kaggle and uploaded the dataset file on Jupyter Notebook, which provides Python environment. For this purpose, required Python libraries such as numpy, pandas, scikit learn, matplotlib were imported. The dataset titled "AI Assistant Usage in student life" containing 10000 records [7]. The dataset contained both categorical and numerical variables which require appropriate transformation to make it suitable for regression-based prediction task.

❖ Dataset Description

Key Attributes:

- *Student Level* (Undergraduate, Postgraduate, etc.)
- *Discipline* (Engineering, Psychology, Arts, etc.)
- *Session Length Min*
- *Total Prompts*
- *Task Type* (Coding, Writing, Studying, etc.)
- *AI Assistance Level*
- *Final Outcome*
- *Used Again* (Boolean)

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

- Target Variable: *Satisfaction Rating*

The aim is to predict **student satisfaction** via numerous regression models.

❖ **Data Pre-processing**

- Feature–Target Separation:

a. Independent variables: All attributes except *Satisfaction Rating*

b. Dependent variable: *Satisfaction Rating*

- Encoding of Categorical Attributes:

a. Categorical variables were transformed using **One-Hot Encoding** to convert them into numerical format suitable for regression algorithms.

- Train–Test Split:

The dataset was divided into:

- **75% Training Set**
- **25% Testing Set**

❖ **Regression Models Implemented**

Following regression algorithms were applied:

1. **Multiple Linear Regression (MLR):** MLR Model predicts a dependent variable using two or more independent variables,

Algorithm:

Input: Dataset D with features X and target y

Output: Predicted satisfaction values y_pred

1. Load dataset D
2. Separate features X and target variable y
3. Identify categorical and numerical attributes
4. Apply One-Hot Encoding to categorical attributes
5. Split data into training set (75%) and testing set (25%)
6. Initialize Linear Regression model
7. Train model using training data (X_train, y_train)
8. Predict output using test data: y_pred = model.predict(X_test)
9. Compute MSE and R2 Score
10. Return evaluation results

2. **Ridge Regression:** Ridge Regression Model is a kind of regularized linear regression that adds an L2 penalty (sum of squared coefficients) to the loss function, efficiently lessening coefficients to avoid overfitting as well as reduce multicollinearity.

Algorithm:

1. Load dataset D
2. Preprocess data (encoding + train-test split)
3. Initialize Ridge model with regularization parameter alpha
4. Train model using X_train and y_train
5. Predict: y_pred = model.predict(X_test)

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

6. Compute MSE and R2 Score
7. Return results

3. **Lasso Regression:** Lasso Regression Model stands for Least Absolute Shrinkage and Selection Operator is a machine learning technique that enhances linear regression by adding an L1 penalty, shrinking less important feature coefficients towards zero, which performs automatic feature selection and regularization to prevent overfitting and simplify models, making them more interpretable, especially with high-dimensional data.

Algorithm:

1. Load dataset D
2. Apply preprocessing and encoding
3. Initialize Lasso model with alpha parameter
4. Train model on training data
5. Generate predictions on test data
6. Evaluate using MSE and R2 Score
7. Identify reduced features (coefficients shrunk to zero)
8. Return results

4. **Decision Tree Regression:** A non-linear model that partitions the feature space into decision rules for predicting satisfaction.

Algorithm:

1. Load dataset D
2. Perform preprocessing and train-test split
3. Initialize Decision Tree Regressor
4. Train decision tree on X_train and y_train
5. Predict satisfaction for X_test
6. Compute MSE and R2 Score
7. Return evaluation results

Each model was trained using the same training dataset and tested using the same test dataset for fair comparison.

❖ **Model Evaluation Metrics**

To evaluate the performance of regression models, the following metrics were used:

1. **Mean Squared Error (MSE):** Measures the average squared difference between predicted and actual satisfaction ratings.
2. **Coefficient of Determination (R^2 Score):** Indicates the proportion of variance in satisfaction explained by the model.

❖ **Confusion Matrix**

Since regression models produce continuous outputs, a confusion matrix cannot be directly applied. Therefore, the regression output was converted into a binary classification problem:

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

- **Satisfied (1):** Satisfaction Rating ≥ 3.0
- **Not Satisfied (0):** Satisfaction Rating < 3.0

Using this threshold-based discretization:

- A **confusion matrix** was generated.
- **Classification accuracy** was computed to evaluate predictive consistency in categorical terms.

This approach enables both numerical and categorical interpretation of model performance.

Algorithm:

Input: Actual values y_test, Predicted values y_pred, Threshold T

Output: Confusion Matrix CM, Accuracy

1. Convert y_test into binary classes:
   if y_test >= T → class = 1 (Satisfied)
   else → class = 0 (Not Satisfied)
2. Convert y_pred into binary classes using same threshold T
3. Compute confusion matrix CM between actual and predicted classes
4. Calculate classification accuracy
5. Return CM and accuracy

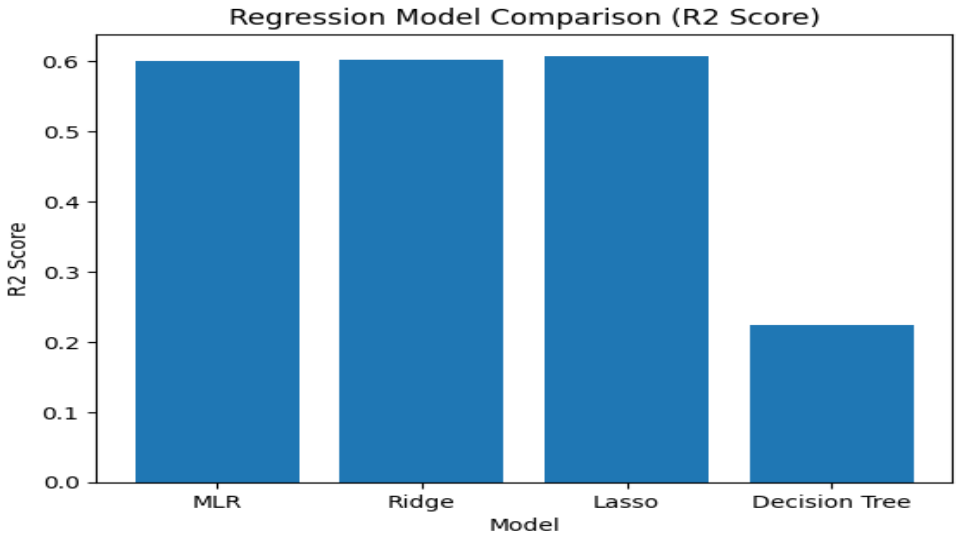
IV. Results and Discussion

❖ Regression performance analysis

On the dataset, four regression models were applied, named Multiple Linear Regression, Ridge Regression, Lasso Regression and Decision Tree Regression. The performance was evaluated by using MSE and R^2 score. The outcome shows that Multiple Linear Regression and Ridge Regression achieved stable performance with low prediction error and good R^2 values. This clearly defines that student satisfaction has strong relationship with the key attributes of dataset.

Lasso Regression produces slightly higher error. But by reducing the weights of less important variables to zero, Lasso Regression simplified the model. This makes the Lasso more useful for identifying important attributes and provides better results. Decision Tree Regression caught non-linear patterns. While it showed good performance on training data but its performance on test data was less consistent and over-fitted. Comparative study shows that decision tree Regression is less reliable than linear regression models.

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

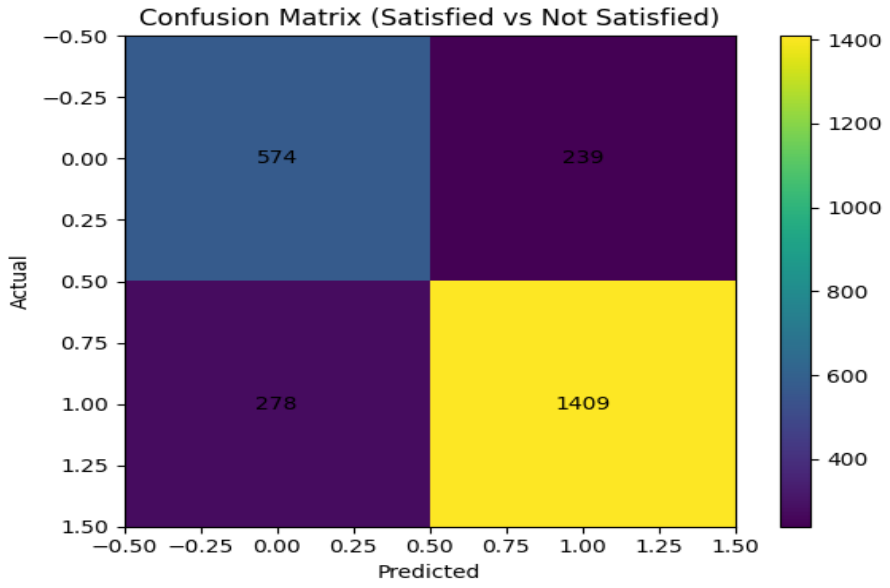


❖ Confusion Matrix

In order to warrant practical interpretation, the continuous satisfaction rating was converted into two categories: Satisfied and Not Satisfied. A confusion matrix was generated. It showed that the model correctly classified satisfied students which indicate that regression model is effectual in identifying user experience. The dual evaluation approach provides both numerical accuracy and practical insights, making the model more useful for education.

Outcome of Confusion Matrix:

Accuracy: 0.7932



@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

V. Conclusion

This study demonstrates that Machine Learning regression model can predict student satisfaction using AI Assistants. Linear model give transparency. Educational stakeholders can use Machine Learning Regression techniques for decision making. The findings of this paper support the mounting role of AI in education and also describe the importance of data driven evaluation because it can help the educators, administrators etc. in improving AI tools to better support the students which can help in improving learning experience.

References:

[1] Alina-Elena, O. (2025, June 1). *The use of AI in education: Positive aspects and potential risk* [Paper presentation]. International Conference Knowledge-Based Organization.

[2] Lampou, R. (2023, August 18). The integration of artificial intelligence in education: Opportunities and challenges. *Review of Artificial Intelligence in Education*, *4*. https://doi.org/10.37497/rev.artif.intell.educ.v4i00.15

[3] Anggraeni, A., Toruan, J. B. G. L., & Karsen, M. (2025, August 28). *The use of AI in education: A systematic literature review* [Paper presentation]. International Conference on [Conference Name Missing].

[4] Dziubanovska, N., & Maslii, V. (2025). The impact of AI integration on the formation of students' critical thinking in the modern educational process. *West Ukrainian National University*.

[5] Roser, M. (2022). *The brief history of artificial intelligence: The world has changed fast – what might be next?* Our World in Data. https://ourworldindata.org/brief-history-of-ai

[6] Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, Article 124167. https://doi.org/10.1016/j.eswa.2024.124167

[7] Ayeshasal89. (2024). *AI assistant usage in student life - synthetic* [Data set]. Kaggle. https://www.kaggle.com/datasets/ayeshasal89/ai-assistant-usage-in-student-life-synthetic

[8] Compilatio. (2025, May 10). *Artificial intelligence in education: Opportunities and challenges in 2025*.

@2026 International Council for Education Research and Training
ISSN: 2959-1376
2026, Vol. 05, Issue 02, 145-154
DOI: https://doi.org/10.59231/SARI7919

https://www.compilatio.net/en/blog/ai-in-education

[9] Verma, B. K., Srivastava, N., & Bharti, A. K. (2025, February 27). *Towards smart education systems: A hybrid data mining and machine learning approach to predicting academic outcomes*. IEEE InTech. https://doi.org/10.1109/InTech64186.2025.11198559

[10] Rani, B. T. (2024). Artificial intelligence tools in learning English language and teaching: How can AI be used for language learning. *Edumania-An International Multidisciplinary Journal*, *2*(4), 230–234. https://doi.org/10.59231/edumania/9085

Received: Feb 20, 2026
Accepted: Mar 11, 2026
Published: Apr 01, 2026