Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 4, Issue 4, 2026

AN IN-DEPTH STUDY ON STUDENTS' PERFORMANCE EVALUATION USING MULTIPLE MACHINE LEARNING CLASSIFIERS AND DATA ANALYTICS APPROACHES

Rabia Hassan[*1], Muhammad Sajid Maqbool[*2], Dr. Naeem Asalam[3], Ariba Afzal[4], Hira Saleem[5], Muqadas Nadeem[6]

[*1,*2,3,5]*Department of Computer Science, NFC Institute of Engineering and Technology, Multan*
[4]*Department of Computer Science, Government College and University, Faisalabad*
[6]*Department of Computer Science, Emerson University, Multan, Pakistan*

[*1]rabialqp555@gmail.com, [*2]sajid.maqbool@nfciet.edu.pk, [3]naeem.aslam@nfciet.edu.pk, [4]ariba16170@gmail.com, [5]hira.saleem@nfciet.edu.pk, [6]nmuqadas587@gmail.com

DOI: https://doi.org/10.5281/zenodo.19494772

Keywords
Dropout prediction, Academic success prediction, Machine learning, Hist Gradient Boosting, performance prediction.

Article History
Received: 11 February 2026
Accepted: 21 March 2026
Published: 06 April 2026

Corresponding Author: *
Rabia Hassan
Muhammad Sajid Maqbool

Abstract
Over the years, institutes have shown interest in improving the quality of education. To enhance education quality, it is important to predict whether a student is at risk of dropping out and to identify if they will achieve high or low scores. High dropout rates and low academic performance not only affect individual students' futures but also impact the institute's reputation. Improving academic performance and student retention involves collecting relevant data to identify students at risk. We use a Kaggle dataset in our research for prediction purposes. The data collected from students includes background information, academic records, behavioral data, and institutional details. Different preprocessing methods are used to improve data quality and validate models that predict student dropout and performance at an early stage. Various machine learning algorithms, such as Decision Tree, Random Forest, Logistic Regression, and Support Vector Machines, are used to evaluate models and analyze student datasets that produce overall results in academic performance. We evaluate all models on both training and testing data and also calculate their accuracy, comparing all model results for better prediction. By applying the Hist Gradient Boosting Classifier, accuracy reaches 90.7%. It is an updated version of the Gradient Boosting algorithm that enhances the efficiency of Gradient Boosting, making it suitable for large datasets. The purpose of the label encoder is to encode categorical features into numerical values. Using the label encoder, accuracy reaches 96.8%, which shows the best performance.

1. Introduction

Education is very important in every field of life; without it, no one can succeed. Education is important because it develops problem-solving, communication, and critical-thinking skills, which lead to better career opportunities. In previous years, we focused on our education to improve the quality, and now, in recent years, we have enhanced the quality of education by using different artificial intelligence techniques and methods. Predicting student performance at an early stage, whether a student is at risk of dropout, there are two possible outcomes of prediction: either the student drops out or remains in session, and also observes student performance in grades, attendance, and other academic activities, achieving a high score or not [1]. Predictions about student performance at an early stage can lead to better results. Predicting academic success means how students' performance in studies, for example, test scores, attendance in class, etc. Students enrolled in an institute from different backgrounds and academic performances. We cannot ensure that a student can only achieve a high score in the academic year, or continue or leave their studies. We use prediction at an early stage to enhance the quality of education. Students dropping out or poor performance is a major problem worldwide. High dropout rates and low academic performance not only affect individual students' futures but also impact the institute in terms of reputation [2] [3]. We use different attributes for prediction by applying ML techniques to gather a huge amount of data, such as grades, attendance, financial background, and educational background. All these attributes help in finding better results for students [4]The core objective to predict students' performance at an early stage leads to enhancing retention and good results in studies. We want to identify whether a student is at risk of dropping out or not, collect student-related data such as personal information, institutional data, academic information, behavioral and educational background [4]. The main objective to predict students early to reduce dropout rates and improve student academic success with the help of different teaching strategies that individual can easily to understand. We explore our prediction using different AI techniques in which we collect online various data that is related to students for accurate prediction. ML predict factors that affecting student performance. Machine learning (ML) used to identify student performance during academic time.

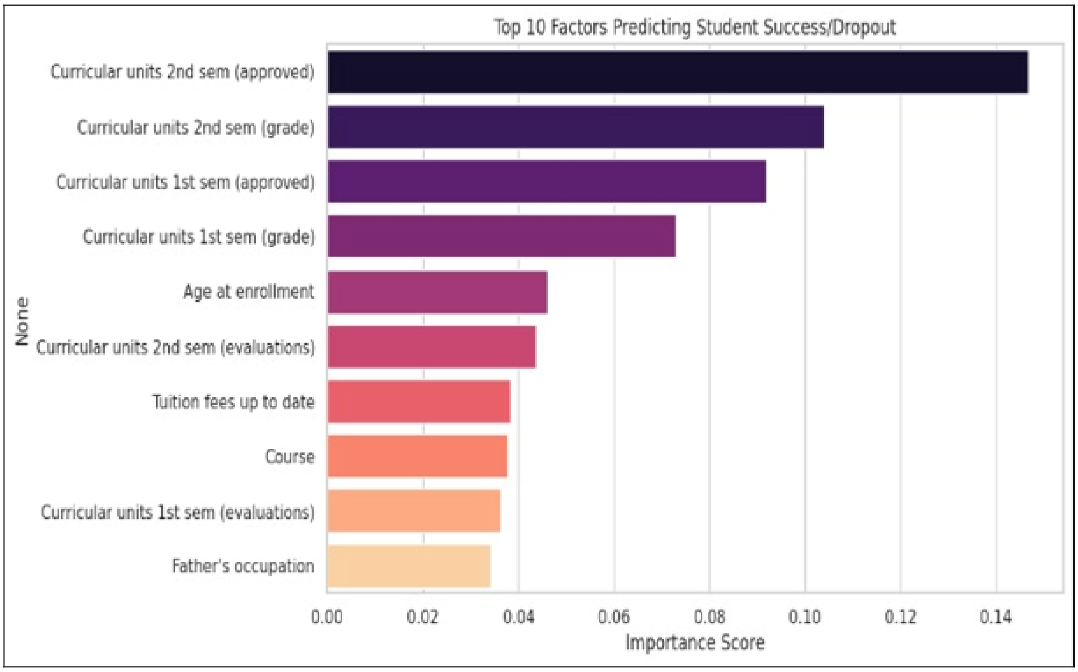

Figure 1: Factors that predict students' performance

Figure 1 represents the main factors that clearly define students' success and dropout performance. We use various algorithms such as Decision Tree, Random Forest, Logistic Regression, Support Vector Machine, and different Kaggle sets. In a decision tree, we use a classification method in which an attribute that best describes student features breaks the data into further classes that are easy to interpret and observe. This algorithm is used to handle small and complex data to quickly take decision in a dataset. Random Forest is an expansion of the decision tree and provides high accuracy. It is used to handle missing values in data for prediction. If small changes are found in data, the majority no. of data remains the same. In Logistic Regression, a classification method is used to predict the probability of student dropout or retention during an academic session. The support vector machine best describes the classification results [5].

2. Literature review

In past educational prediction relies on observation with statistical analysis, not reliable attach limited outcomes students' performance. Innovations with time improve prediction qualities [6] Mostly researchers used educational data mining (EDM) relies on data mining and machine learning algorithms to analyze student dropout and academic performance. The goal of algorithms to early predict student dropout risk and enhance retention or academic performance [7]. Compare different techniques results and find which one produce good quality outcome about prediction. In educational institute predict data based on student's grades, course, attendance, academic history with support of techniques such as decision trees, logistic regression, random forests, support vector machine to find accurate results [8] There are many purposes in research to identify early student at risk of dropout or achieving low grades. In previous research, researcher used deep learning models that predict student academic performance with accuracy. The model evaluation based on accuracy, precision, recall. In previous studies researcher used 70% survey work that rely on student grades and only 10% prediction data based on curriculum that generates results [9]. Researcher compared machine learning algorithms that predict student performance with high accuracy results. While predicting model transparency and accuracy are significant in executing the models. Previous researches model achieved AUC-ROC score of 0.90, indicating a high degree of separability between the 'Dropout' and 'Retain' classes [5]. This paper focus on machine learning algorithms that produce accuracy and efficiency in student performance that relates student retain or dropout on the basis of their results. In another research Cat Boost is considered the best technique but it achieved AUC-ROC score of 0.77 [10]

Many researchers used data balancing methods with ML classifiers to measure prediction gain accuracy level 85.3%. Different models show different accuracy outcomes [11]

In 2022 dropout students' prediction by using ML techniques had a success level 93% with four years collection of data that used as a dataset, less than 50% of data was used for early prediction [12]

3. Methodology

Our proposes approach for prediction of students' dropout and academic success shown in flowchart. Our methodology starts with the collection of student related data that is collected from different areas then preprocess data for cleaning. Further data is divided into smaller parts that used for training and testing. To train the data different Machine learning algorithms used. All machine learning algorithm used in testing 20% and training 80% of data for prediction. Further all different results stored in a single place(repository) and also compare the results which model predict better results on the base of testing data [7]. After validation or evaluation final results generate on the basis of previous process that we follow for prediction students' dropout and academic success.

1. Collection of data
2. Data preprocessing
3. Division of data
4. Machine learning techniques used
5. Prediction of best model
6. Validation and evaluation
7. Final results

The steps of proposed methodology are given below in detail:

3.1. Collection of data:

Initial step is the data collection that is essentials required for analysis in whole process. Data is collected from different areas about students' performance in academic session. Student data that strongly related with their academic performance. During collection of data previous academic performance history matters along with current performance such as current attendance, marks, performance in class, assignment work these are the preliminary investigation data used in prediction. Further educational and financial family background data also collected [13].

3.2. Data Preprocessing

After data collected from different resources perform preprocessing on data to clean from inconsistencies and unclear information that further used in prediction about student dropout and academic success.

3.3. Division of data

When preprocessing is done data is divided into two parts training and testing for model training and evaluation. Training and testing refer to the process of dividing the collection of data into two parts a training set and a testing test we use 80 percent data for training models and 20 percent data is used in testing the machine learning models.

3.4. Machine learning techniques used:

Prediction either student will dropout or retain in a academic session we use Logistic regression in which binary classification (dropout/retain) on the basis of common attributes of student academics used. ML also provide a visual presentation of binary classification data for easily prediction which is known as Decision trees that shows workflow of results. We use multiple flowcharts of students' performance. prediction and make a clear comparison of flowchart result therefore interpret its result in a single place for good results on the basis of classification and academic data prediction this ML techniques is Random forest. Support vector machine used both in small and large amount of data. It is used in classification where student is at risk of dropout or retain in performance.

3.5. Prediction of best model

Compare the performance of implemented models by using their relevant techniques we made prediction on the basis of their results which model produce better outcomes.

3.6. Validation and evaluation:

Use data for testing in which we can assess the effectiveness of the model results after evaluate or validate its outcomes.

3.7. Final results:

After validation model results final accurate results found about students will dropout or succeed in their academic session.

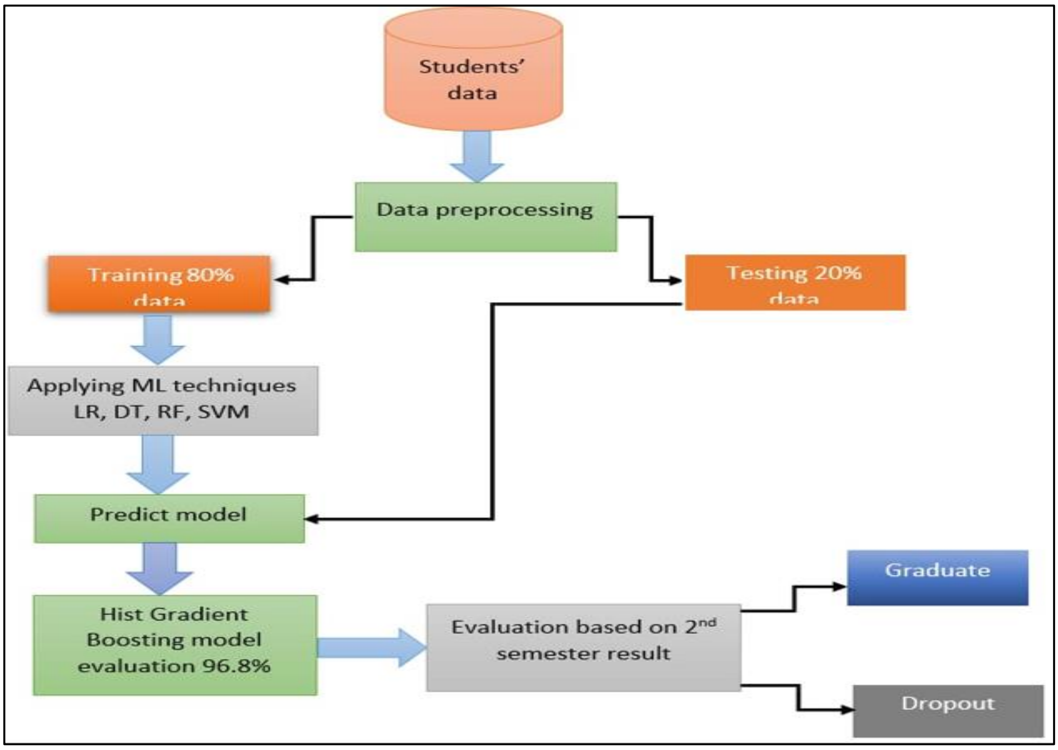

Figure 2 Workflow of the study

4. Dataset description:

Selecting Datasets: The dataset is downloaded from Kaggle dataset database(repository) containing 4425 rows and 35 columns(features). Every feature will be fetched using repository where dataset kept for analysis. Application mode, Application order, Marital status, Daytime/evening, Course, Previous qualification, Nationality, Mother qualification, Father qualification, Mother occupation, Father occupation, Tuition fees up to date, Displaced, Gender, Scholarship holder, Age at enrollment, Curricular units 1st sem (credited), Debtor, International, Curricular units 1st sem (enrolled), Educational special needs, Curricular units 1st sem (evaluations), Curricular units 1st sem (approved), Curricular units 1st sem (grades), Curricular units 1st sem (without grades), Curricular units 2nd sem (credited), Curricular units 2nd sem (enrolled), Curricular units 2nd sem (evaluations), Curricular units 2nd sem (approved), Curricular units 2nd sem (grades), Curricular units 2nd sem (without grades), Unemployment rate, Inflation rate, GDP, Target. In these features some values are binary, numeric and some string value.

Table 1 Description of Dataset Features

Features	Type	Values
Marital status	Binary	Single/Married
Application mode	Numeric	0 to 20
Application order	Numeric	0 to 10
Course	Categorical	Different values
Daytime/evening	Binary	0 and 1
Previous qualification	Categorical	High, Low, Average
Nationality	Categorical	Different values
Mother's qualification	Categorical	Different values
Father's qualification	Categorical	Different values
Mother's occupation	Categorical	Different values
Father's occupation	Categorical	Different values
Displaced	Categorical	Different values
Educational special	Categorical	High, Low, Average
Debtor	Categorical	High, Low, Average
Tuition fees up to date	Categorical	Different values
Gender	Binary	Male/Female
Scholarship holder	Binary	Yes/No
Age at enrollment	Numeric	Different values
International	Binary	Yes/No
Curricular units 1st sem (credited)	Numeric	Different values
Curricular units 1st sem (enrolled)	Numeric	Different values
Curricular units 1st sem (evaluations)	Numeric	Different values
Curricular units 1st sem (approved)	Numeric	Different values
Curricular units 1st sem (grades)	Categorical	Different values
Curricular units 1st sem (without grades)	Numeric	Different values

Curricular units 2nd sem (credited)	Numeric	Different values
Curricular units 2nd sem (enrolled)	Numeric	Different values
Curricular units 2nd sem (evaluations)	Numeric	Different values
Curricular units 2nd sem (approved)	Numeric	Different values
Curricular units 2nd sem (grades)	Categorical	Different values
Curricular units 2nd sem (without grades)	Numeric	Different values
Unemployment rate	Numeric	Different values
Inflation rate	Numeric	Different values
Target	Categorical	Enrolled/Dropout/Success

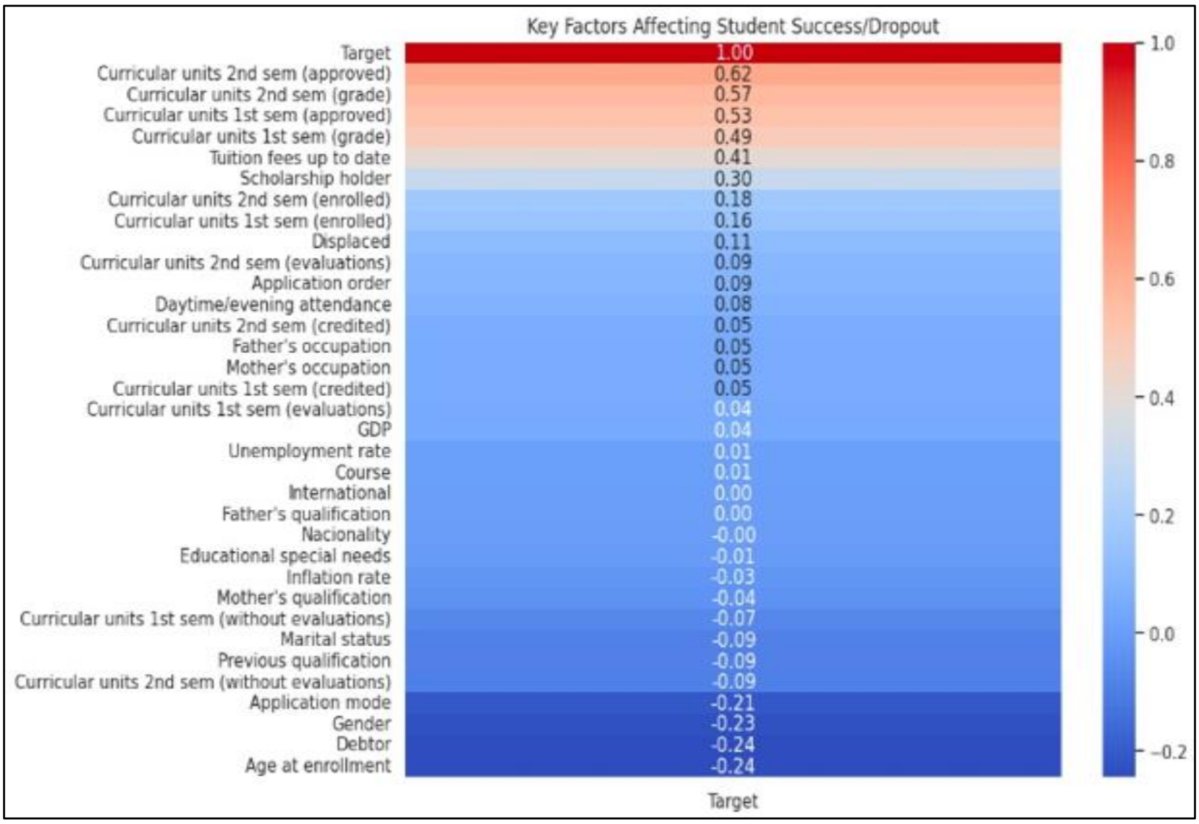

Figure 3 Features of the used Dataset

To accurately process this dataset on machine learning models, we label its features [7]

Table 2 Gender labeling

Old features	New labels
Male	0
Female	1

All the binary type features, such as Marital status, Daytime/evening, Gender, International, and scholarship holder, are labeled as 0 and 1.

Table 3 Previous qualification labeling

Old features	New labels
High	0
Low	1
Average	2

All the features that consist of three values can be labeled as 0, 1 and 2 as shown above qualification labeling table

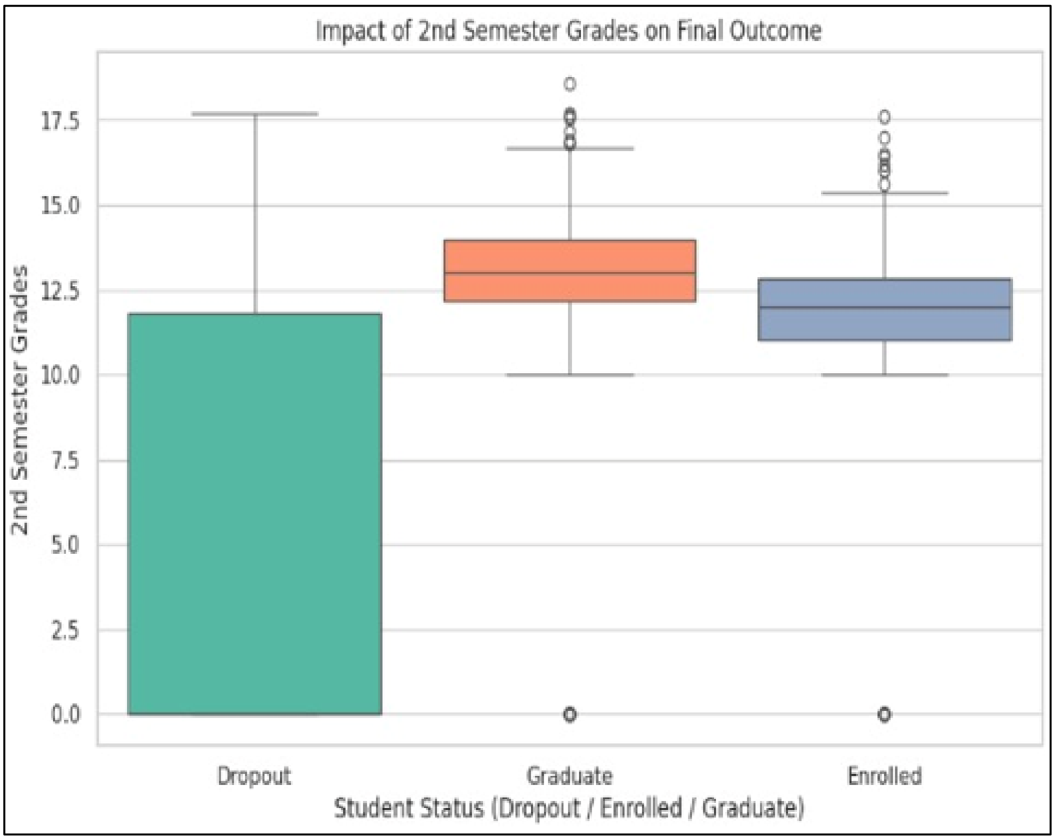

Figure 4 Dropout vs Retention

This Boxplot graph figure 4 shows those students who gained high grades in 2^{nd} semester having maximum chances of graduation while others often can be dropout in academic session.

After selecting dataset, next step is division of dataset that trained dataset for further testing this dataset.

Table 4 Dataset division for training and testing

	%age	No. of records
Training dataset	80%	3540
Testing dataset	20%	885
Total	100%	4425

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
X = df.drop(columns=['Target'])
y = le.transform(df['Target'])
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=42, stratify=y)
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

print("Data successfully split and scaled.")
[5]
Data successfully split and scaled.
```

Figure 5 Splitting dataset

Dataset having no null values, now dataset is divided into two parts: we use 80% of data for training and remaining 20% used for testing. Above figure 5 shown that dataset successfully split.

4.1. Selecting ML models:

Here we select four machine learning (ML) algorithms such as Logistic regression (LR), Decision tree (DT), Random forest (RF), Support vector machine (SVM) to check its accuracy that best describe model for predicting results. The following figure 6 clearly shows model accuracy results.

```
models = {
    "Logistic Regression": LogisticRegression(max_iter=1000),
    "Decision Tree": DecisionTreeClassifier(random_state=42),
    "Random Forest": RandomForestClassifier(n_estimators=100, random_state=42),
    "SVM": SVC(probability=True)
}

results = {}
for name, model in models.items():
    model.fit(X_train_scaled, y_train)
    y_pred = model.predict(X_test_scaled)
    acc = accuracy_score(y_test, y_pred)
    results[name] = acc
    print(f"{name} Accuracy: {acc*100:.2f}%")

res_df = pd.DataFrame(list(results.items()), columns=['Model', 'Accuracy']).sort_values(by='Accuracy', ascending=False)
plt.figure(figsize=(10, 5))
sns.barplot(x='Accuracy', y='Model', data=res_df, palette='viridis')
plt.title('Comparison of ML Models for Student Prediction')
plt.show()

Logistic Regression Accuracy: 75.82%
Decision Tree Accuracy: 68.59%
Random Forest Accuracy: 77.97%
SVM Accuracy: 75.93%
/tmp/ipykernel_2348/1756015129.py:27: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x='Accuracy', y='Model', data=res_df, palette='viridis')
```

Figure 6 Model accuracy results

Comparison of ML models for predicting students' performance shows its accuracy by using bar graph depict as:

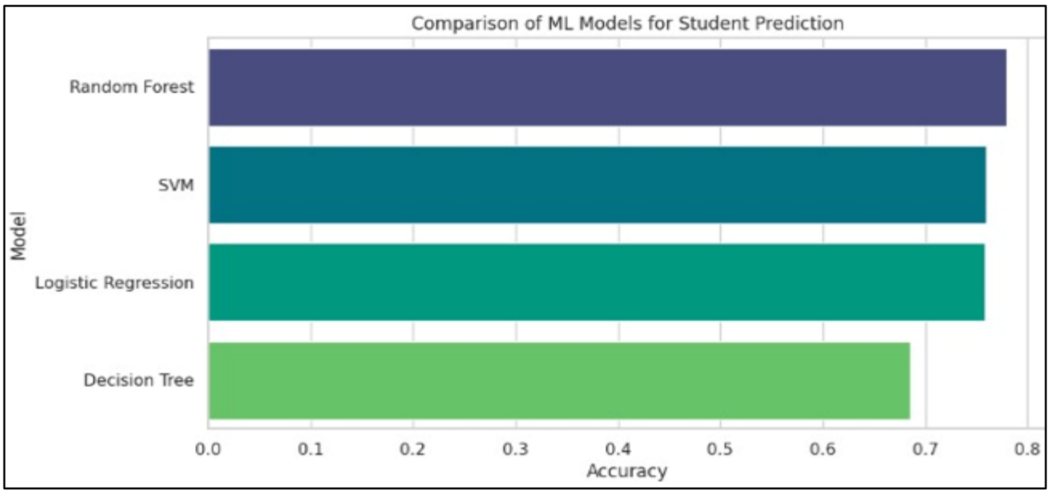

Figure 7 Comparing ML models

The figure 7 shows trained model accuracy reach at 75%. These results are not appropriate for students' performance evaluation. Again, optimize model with 20:80 ratio to improve results.

By applying Hist Gradient Boosting Classifier, accuracy reach at 90.7%. It is updated version of Gradient Boosting algorithm enhance efficiency of Gradient Boosting that suitable for large datasets. By using label encoder, accuracy reach 96.8% which shows best position. Purpose of label encoder encodes categorical features into numerical values. Visualization of model evaluation that fit or not shown in following graph:

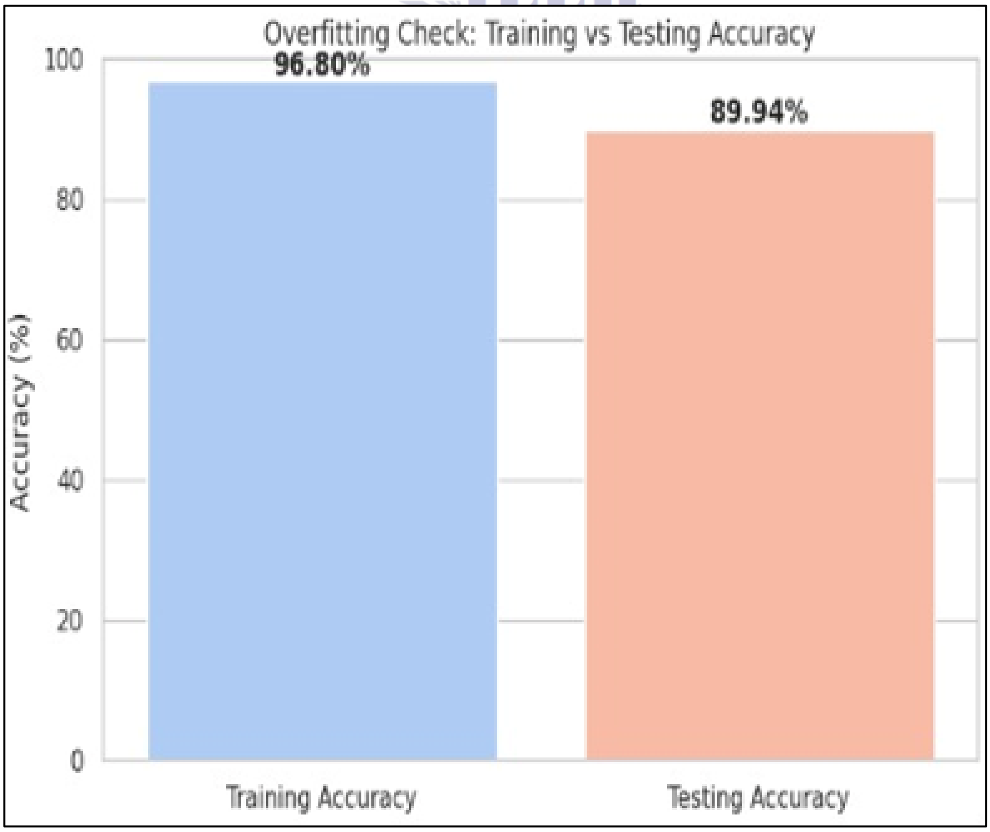

Figure 8 Data accuracy using 20:80 ratio

4.2. Model validation:

The model achieved an outstanding AUC-ROC score of 0.95, indicating a high degree of separability between the 'Dropout' and 'Graduate' classes. The Confusion Matrix further validates this performance, showing that the model correctly identified 233 dropout cases and 420 graduate cases, with a very low false-positive rate [19]. This demonstrates the robustness of the proposed framework in predicting student outcomes. Parameters used to measure models by using accuracy, precision and recall etc. A system will be generated for prediction about students' by using best model after validation.

4.3. Results:

Here is a graph which explain all factors that directly or indirectly influence on students' performance. Figure 9 shows all impact factors about their performance level.

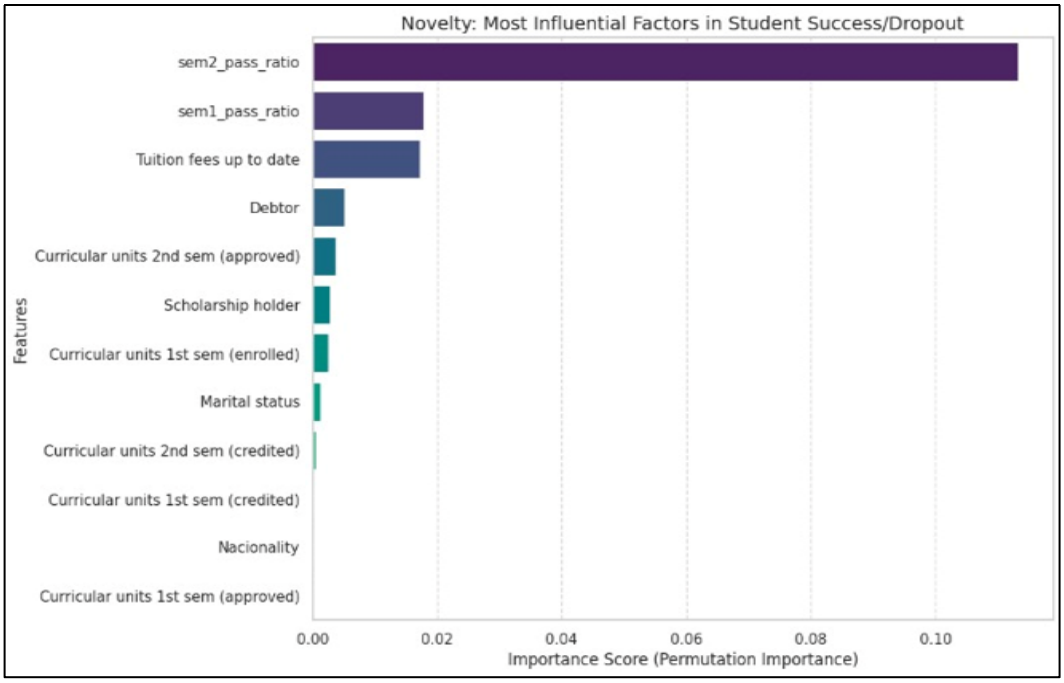

Figure 9 Impact factor on academic performance

This graph shows students retained in academic session and their success level depends on results of 2nd semester.

5. Conclusion:

In this research paper we want to check students' performance level in their academics by using machine learning (ML) techniques. By selecting models RF, DT, LR and SVM compare their results for predicting results. The model achieved an outstanding AUC-ROC score of 0.95, indicating a high degree of separability between the 'Dropout' and 'Graduate' classes. By applying Hist Gradient Boosting Classifier, accuracy reach at 90.7%. Again, optimize model with 20:80 ratio to improve results. By using label encoder, accuracy reach 96.8% which shows best position. In future our main focus will be on success level of students' and retention of students' in academic session, by using advanced ML models we conclude accurate and powerful prediction results that showed long time impact. Combining ML models expected for prediction of students' performance.

6. REFERENCES

W. F. Wan Yaacob, N. Mohd Sobri, S. A. M. Nasir, W. F. Wan Yaacob, N. D. Norshahidi, and W. Z. Wan Husin, "Predicting Student Drop-Out in Higher Institution Using Data Mining Techniques," in Journal of Physics: Conference Series, Institute of Physics Publishing, Jun. 2020. doi: 10.1088/1742-6596/1496/1/012005.

S. A. SULAK and N. KOKLU, "Predicting Student Dropout Using Machine Learning Algorithms," Intelligent Methods in Engineering Sciences, Jan. 2025, doi: 10.58190/imiens.2024.103.

A. J. Kehinde, A. E. Adeniyi, R. O. Ogundokun, H. Gupta, and S. Misra, "Prediction of Students' performance with Artificial Neural Net-work using Demographic Traits." [Online]. Available: https://www.kaggle.com/dipam7/student

M. M. Hussain, S. Akbar, S. A. Hassan, M. W. Aziz, and F. Urooj, "Prediction of Student's Academic Performance through Data Mining Approach," Journal of Informatics and Web Engineering, vol. 3, no. 1, pp. 241–251, Feb. 2024, doi: 10.33093/jiwe.2024.3.1.16.

S. Albugami, H. Almagrabi, and A. Wali, "From Data to Decision: Machine Learning and Explainable AI in Student Dropout Prediction," IBIMA Business Review, vol. 2024, 2024, doi: 10.5171/2024.246301.

Syeda Qanitah Naqvi, Syeda Rabail Zahra, Muhammad Sajid Maqbool, Muqadas Nadeem, Hira Saleem, & Mahnoor Zaman. (2026). AN AUTOMATED AND ARTIFICIAL INTELLIGENCE-BASED SYSTEM FOR THE DIAGNOSIS OF SKIN CANCER. *Policy Research Journal*, 4(4), 58–72. Retrieved from https://policyrj.com/1/article/view/1769.

J. Bhat and Y. Jayaram, "Predictive Analytics for Student Retention and Success Using AI/ML," International Journal of Artificial Intelligence, Data Science, and Machine Learning, vol. 4, pp. 121–131, 2023, doi: 10.63282/3050-9262.ijaidsml-v4i4p114.

Sarim Javed, Muhammad Sajid Maqbool, Dr. Naeem Aslam, Muhammad Haseeb Ur Rehman, Muqadas Nadeem, & Hira Saleem. (2026). HIGH ACCURACY INTRUSION DETECTION IN IOT VIA HYBRID ML DL MODELS. *Policy Research Journal*, 4(4), 73–85. Retrieved fromhttps://policyrj.com/1/article/view/1770.

N. Aslam, K. Abid, N. Aslam, M. Fuzail, M. S. Maqbool, and K. Sajid, "An Efficient Deep Learning Approach for Prediction of Student Performance Using Neural Network," VFAST Transactions on Software Engineering, vol. 11, no. 4, pp. 67–79, Oct. 2023, doi: 10.21015/vtse.v11i4.1647.

J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," Feb. 01, 2020, MDPI AG. doi: 10.3390/app10031042.

A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," Jan. 01, 2021, MDPI AG. doi: 10.3390/app11010237.

M. Rebelo Marcolino et al., "Student dropout prediction through machine learning optimization: insights from moodle log data," Sci. Rep., vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-93918-1.

N. Mduma, "Data Balancing Techniques for Predicting Student Dropout Using Machine Learning," Data (Basel)., vol. 8, no. 3, Mar. 2023, doi: 10.3390/data8030049.

H. Dasi and S. Kanakala, "International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Student Dropout Prediction Using Machine Learning Techniques." [Online]. Available: www.ijisae.org

J. H. Guanin-Fajardo, J. Guaña-Moya, and J. Casillas, "Predicting Academic Success of College Students Using Machine Learning Techniques," Data (Basel)., vol. 9, no. 4, Apr. 2024, doi: 10.3390/data9040060.

E. Kalita et al., "Predicting student academic performance using Bi-LSTM: a deep learning framework with SHAP-based interpretability and statistical validation," Front. Educ. (Lausanne)., vol. 10, 2025, doi: 10.3389/feduc.2025.1581247.

Khan, M. M. R. (2025). Recent advances in breast cancer detection: A review on imaging and AI techniques. ICCK Proceedings.

Maqbool, M. S., Hanif, I., Iqbal, S., Basit, A., & Shabbir, A. (2023). Optimized feature extraction and cross-lingual text reuse detection using ensemble machine learning models. Journal of Computing & Biomedical Informatics, 5(01), 26-40.

Abid, K., Aslam, N., Fuzail, M., Maqbool, M. S., & Sajid, K. (2023). An efficient deep learning approach for prediction of student performance using neural network. VFAST Transactions on Software Engineering, 11(4), 67-79.

Kanwal, F., Abid, M. K., Maqbool, M. S., Aslam, N., & Fuzail, M. (2023). Optimized classification of cardiovascular disease using machine learning paradigms. VFAST Transactions on Software Engineering, 11(2), 140-148.

Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). Understanding Urban Expansion Through Multi-Temporal Satellite Data Analysis. Kashf Journal of Multidisciplinary Research, 2(09), 252-273.

Hasnain, M. A., Ali, S., Malik, H., Irfan, M., & Maqbool, M. S. (2023). Deep learning-based classification of dental disease using x-rays. Journal of Computing & Biomedical Informatics, 5(01), 82-95.

Basit, A., Hanif, I., Maqbool, M. S., Qayyum, W., Hasnain, M. A., & Nazeer, R. (2023). Cross-lingual information retrieval in a hybrid query model for optimality. Journal of Computing & Biomedical Informatics, 5(01), 130-141.

Hasnain, M. A., Ali, Z., Maqbool, M. S., & Aziz, M. (2024). X-ray image analysis for dental disease: A deep learning approach using efficientnets. VFAST Transactions on Software Engineering, 12(3), 147-165.

Rafiqee, M. M., Qaiser, Z. H., Fuzail, M., Aslam, N., & Maqbool, M. S. (2023). Implementation of efficient deep fake detection technique on videos dataset using deep learning method. Journal of Computing & Biomedical Informatics, 5(01), 345-357.

Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. Kashf Journal of Multidisciplinary Research, 2(12), 39-57.

Muhammad Noman, Muhammad Sajid Maqbool, Dr. Naeem Aslam, Muqadas Nadeem, Hira Saleem, & Hanzla. (2026). SLEEP DISORDER SCORING AUTOMATED USING ADVANCED DATA SCIENCE AND MACHINE LEARNING TECHNIQUES. Policy Research Journal, 4(3), 853–868. Retrieved from https://policyrj.com/1/article/view/1713

Zainab Naveed, Rubaina Nazeer, Muhammad Sajid Maqbool, Dr. Naeem Aslam, Hira Saleem, & Muqadas Nadeem. (2026). AN END-TO-END ORTHOPEDIC DISEASE IMAGE CLASSIFICATION SYSTEM USING CONVOLUTIONAL NEURAL NETWORKS. Policy Research Journal, 4(3), 837–852. Retrieved from https://policyrj.com/1/article/view/1712

Mahnoor Zaman, Nosheen Fatima, Muhammad Sajid Maqool, Dr. Naeem Aslam, Rubaina Nazeer, & Hira Saleem. (2026). INGREDINET: INTELLIGENT CNN FOR FOOD INGREDIENT RECOGNITION AND CLASSIFICATION. Policy Research Journal, 4(3), 789–805.

Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). Understanding Urban Expansion Through MultiTemporal Satellite Data Analysis. Kashf Journal of Multidisciplinary Research, 2(09), 252-273.

M. A., Ali, Z., Maqbool, M. S., & Aziz, M. (2024). X-ray image analysis for dental disease: A deep learning approach using efficientnets. VFAST Transactions on Software Engineering, 12(3), 147-165.

Abid, K., Aslam, N., Fuzail, M., Maqbool, M. S., & Sajid, K. (2023). An efficient deep learning approach for prediction of student performance using neural network. VFAST Transactions on Software Engineering, 11(4), 67-79.

R. Paul, S. Sarker, H. El Aouifi, S. Hussain, A. K. Baruah, and S. Gaftandzhieva, "Analyzing dropout of students and an explainable prediction of academic performance utilizing artificial intelligence techniques," Front. Educ. (Lausanne)., vol. 10, 2025, doi: 10.3389/feduc.2025.1698505.

T. R. Noviandy et al., "Machine Learning for Early Detection of Dropout Risks and Academic Excellence: A Stacked Classifier Approach," Journal of Educational Management and Learning, vol. 2, no. 1, pp. 28–34, May 2024, doi: 10.60084/jeml.v2i1.191.

A. Villar and C. R. V. de Andrade, "Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study," Discover Artificial Intelligence, vol. 4, no. 1, Dec. 2024, doi: 10.1007/s44163-023-00079-z.

S. C. Matz, C. S. Bukow, H. Peters, C. Deacons, and C. Stachl, "Using machine learning to predict student retention from socio-demographic characteristics and app-based engagement metrics," Sci. Rep., vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-32484-w.

M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40561-022-00192-z.

K. Oqaidi, S. Aouhassi, and K. Mansouri, "Towards a Students' Dropout Prediction Model in Higher Education Institutions Using Machine Learning Algorithms," International Journal of Emerging Technologies in Learning, vol. 17, no. 18, pp. 103–117, 2022, doi: 10.3991/ijet.v17i18.25567.