Spectrum of Engineering Sciences

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

A ROBUST DEEP LEARNING FRAMEWORK FOR PREDICTING ACADEMIC SUCCESS USING ADVANCED NEURAL ARCHITECTURES

Sundas Israr[1], Muhammad Sajid Maqbool[*2], Dr. Israr Hanif[3], Muqadas Nadeem[4], Abdul Basit[5], Aiman Ali Batool[6]

[1]*Department of Software Engineering, NUML University, Multan, Pakistan*
[*2]*Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan*
[3]*Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan*
[4]*Department of Computer Science, Emerson University, Multan, Pakistan*
[5]*Muhammad Nawaz Sharif University of Engineering and Technology, Pakistan*
[6]*Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan*

[*1]sundus.israr@numl.edu.pk, [*2]sajid.maqbool@nfciet.edu.pk, [3]aoaisrar@bzu.edu.pk, [4]nmuqadas587@gmail.com, [5]abdul.basit@mnsuet.edu.pk, [6]aimanalibatool71@gmail.com

[*2]ORCID: 0000-0001-5583-3550

DOI: https://doi.org/10.5281/zenodo.20156230

Keywords
Educational Data Mining (EDM), Student Performance Prediction, Machine Learning, Deep Learning, Classification Algorithms, Principal Component Analysis (PCA)

Article History
Received: 16 March 2026
Accepted: 25 April 2026
Published: 13 May 2026

Corresponding Author: *
Muhammad Sajid Maqbool

Abstract

In recent years, data mining techniques have gained significant attention in educational institutions for improving the quality of education and enhancing academic decision-making processes. Accurate prediction of student academic performance plays a vital role in identifying students at risk of poor achievement and supports the development of effective educational strategies. Numerous studies have focused on predicting student performance at the higher education level, as academic success in earlier semesters strongly influences students' future learning progress and retention. In semester-based educational systems, many students experience academic difficulties or fail to achieve satisfactory grades during the initial stages of higher education. Therefore, the early prediction of student performance is essential for improving student retention and academic outcomes. Educational Data Mining (EDM) provides techniques for extracting meaningful information, hidden patterns, and valuable knowledge from large volumes of educational data. These extracted insights can be utilized to predict students' future academic success and support timely interventions. The primary objective of this research is to evaluate student performance using multiple classification techniques and identify the model that achieves the highest predictive accuracy. The educational dataset used in this study is obtained from a Kaggle repository. The proposed methodology consists of several stages. Initially, the dataset undergoes preprocessing, including the removal of duplicate records and handling of missing values through appropriate data imputation techniques. Subsequently, three classification algorithms are implemented using the Weka data mining tool. These algorithms include Deep Learning-based Neural Networks (NN) and traditional machine learning techniques such as Random Forest (RF), Support

Vector Machine (SVM). To enhance feature quality and reduce dimensionality, Principal Component Analysis (PCA) is applied for optimized feature extraction. Furthermore, the performance of all classification models is evaluated using a training–testing split validation available in the Weka environment. The models are assessed using standard performance evaluation metrics, including Training Accuracy, Testing Accuracy, Precision, Recall, and F1-Score. Experimental results indicate that the Neural Network and Random Forest classifiers outperform the SVM model in terms of predictive accuracy and overall classification performance.

1 INTRODUCTION

Educational Data Mining (EDM) has emerged as an important interdisciplinary research field that combines education, data analysis, machine learning, and data mining techniques to extract meaningful knowledge from educational datasets. The increasing availability of academic data and the rapid growth of digital educational platforms have created a strong need for intelligent systems capable of analyzing large volumes of educational information. Educational institutions continuously generate massive amounts of data through learning management systems, online educational platforms, classroom activities, examinations, seminars, workshops, and student assessments. However, raw educational data has limited value unless it is effectively processed and analyzed to support academic decision-making and improve the quality of education. Data mining techniques provide efficient methods for discovering hidden patterns, extracting useful information, and identifying meaningful relationships from large educational databases. These extracted insights can support institutions in monitoring academic progress, improving teaching methodologies, and predicting student performance. Educational systems are considered one of the most critical components for the social, economic, scientific, and technological development of any nation. Therefore, maintaining a well-structured and effective educational system is essential for sustainable national development. Developed countries have implemented advanced educational systems and evaluation frameworks to improve academic quality and student outcomes. In modern educational environments, learning is no longer restricted to traditional classroom teaching; instead, it extends to online learning systems, intelligent tutoring systems, web-based education platforms, project-based learning, seminars, and workshops. Despite these advancements, the success of such systems largely depends on accurate evaluation and assessment mechanisms.

Student academic performance prediction has become a significant research area because it enables educational institutions to identify students who are at risk of poor academic achievement or failure. Early prediction of student performance allows educators to take timely interventions and provide appropriate academic support to improve learning outcomes and student retention. Educational Data Mining facilitates this process by analyzing historical educational data and generating predictive models capable of forecasting student success or failure. Unlike traditional database analysis methods that only provide descriptive information, EDM techniques can answer predictive and analytical questions such as estimating a student's future academic performance, identifying learning difficulties, and determining the factors influencing academic achievement. The prediction of student performance is closely related to the effectiveness of the teaching and learning process. Learning outcomes represent the knowledge, cognitive skills, attitudes, and behavioral changes acquired by students during the educational process. These outcomes can be evaluated through assessments, assignments, examinations, and classroom activities. By analyzing these indicators, institutions can better understand students' learning behavior and improve instructional strategies. Consequently, the accurate prediction of student performance can contribute significantly to improving educational quality and institutional effectiveness.

In recent years, machine learning and deep learning approaches have been widely applied in educational data mining for student performance prediction. These approaches enable automated analysis of educational datasets and improve prediction accuracy compared to traditional statistical methods. However, challenges such as insufficient datasets, irrelevant features, imbalanced data, and low prediction accuracy still affect the effectiveness of existing prediction systems. Therefore, selecting appropriate classification techniques and optimized feature sets is essential for developing reliable predictive models. The primary objective of this research is to develop an intelligent framework for predicting student academic performance using machine learning and deep learning techniques. The study aims to organize and preprocess educational datasets collected from academic institutions and analyze student performance using multiple academic and demographic features. Various classification algorithms, including Deep Learning and Machine Learning models, are implemented and evaluated to determine the most effective predictive approach. Furthermore, the proposed framework is designed to assist educational institutions in accurately forecasting student academic outcomes and supporting data-driven educational decision-making processes.

2 Literature Review

Numerous studies in the field of Educational Data Mining (EDM) have focused on developing predictive models to evaluate and forecast student academic performance. These studies primarily aim to identify students at academic risk, analyze factors affecting learning outcomes, and improve educational decision-making processes through predictive analytics [1], [8]. Existing research has concentrated on identifying influential academic attributes and designing efficient classification models capable of accurately predicting future student performance. Several significant contributions in this domain are discussed below. V. Ganesh et al. proposed a Hybrid Educational Data Mining (HEDM) model to evaluate overall student academic performance and improve educational standards. The proposed framework integrated the J48 decision tree classifier with the Naïve Bayes classification approach to analyze academic factors and categorize student performance effectively. The hybridization of these techniques enhanced classification efficiency and improved prediction accuracy by combining probabilistic and rule-based learning mechanisms. G. Jayanthi and V. Ramesh developed a student performance prediction model using multiple classification algorithms, including Fully Connected Neural Networks, Naïve Bayes (NB), Sequential Minimal Optimization (SMO), J48, and REPTree classifiers [2]. Their model was designed to identify weak and unmotivated students at an early stage, enabling educators to implement corrective measures and improve student learning outcomes. Experimental findings demonstrated that classification-based predictive systems can significantly support academic monitoring and student performance enhancement. Altabrawee et al. [6] introduced a predictive framework for student performance evaluation using four machine learning techniques: Artificial Neural Networks (ANN), Naïve Bayes (NB), Decision Trees (DT), and Linear Regression (LR). The study focused on predicting students' overall academic performance in specific subjects. Furthermore, the research analyzed the influence of internet usage and social networking activities on academic achievement, highlighting the relationship between digital learning behavior and educational performance. Educational Data Mining techniques have increasingly contributed to the modernization and expansion of educational systems. Chitti M. et al. [5] investigated several educational factors associated with EDM, including student performance, dropout prediction, cognitive abilities, instructor effectiveness, course delivery, and educational content development. The study emphasized the use of data analytics tools for extracting meaningful patterns from educational datasets. However, the authors also noted that many existing predictive models remain overly complex and difficult to interpret, limiting their practical implementation in educational environments. Kim B. et al. [6] addressed the challenge of predicting student success in online learning environments using deep learning techniques. The researchers proposed a framework named *GritNet*, which utilizes Bidirectional Long Short-Term Memory (BLSTM) networks for student performance prediction. The proposed deep learning model demonstrated superior performance compared to traditional logistic regression-based methods, particularly during the early weeks of online courses when accurate predictions are most critical for timely educational interventions.

Recent advancements in Federated Learning (FL) have also influenced educational data mining research. Dayan I. et al. [7] demonstrated the effectiveness of FL in enabling collaborative model training while preserving data privacy. Federated Learning is a distributed machine learning paradigm in which multiple clients collaboratively train a shared global model without exchanging raw data. Instead, only model parameters are transmitted to a centralized aggregation server, ensuring data privacy and security. FL has gained significant attention in domains such as healthcare, intelligent systems, wireless communication, and educational analytics due to its ability to support collaborative learning across heterogeneous datasets while maintaining user confidentiality. Several studies have highlighted the advantages of FL for educational prediction systems. Hu K. et al. [9] provided a comprehensive overview of Federated Learning, including its architecture, development process, classification, and comparison with traditional distributed learning methods. The study also discussed key challenges associated with FL, such as communication overhead, system heterogeneity, unreliable model aggregation, and privacy preservation. Additionally, the authors explored various federated deep learning algorithms and emphasized the future potential of FL-based intelligent educational systems. Federated Learning has demonstrated promising performance in predictive modeling tasks due to its ability to combine local models into a unified global model while preserving data privacy. Existing FL-based predictive frameworks have shown improved reliability, efficiency, and accuracy compared to conventional machine learning approaches. In these systems, multiple localized datasets are used to develop independent machine learning models, which are subsequently aggregated into a global predictive model. This collaborative learning mechanism enables institutions to benefit from distributed knowledge without sharing sensitive educational data. Bayer et al. [13] combined social network analysis and data mining techniques to identify potentially successful and unsuccessful high school students during the early stages of their academic careers. The study enhanced predictive accuracy by incorporating social interaction patterns into the educational prediction framework. Similarly, Zahyah A. et al. [20] proposed a framework that utilized admission-time data to predict students' first-semester academic outcomes. The framework classified students into high-risk and high-performance categories, enabling institutions to implement early academic interventions. Brijesh K. B. and Pal S. [22] proposed a predictive model that considered multiple student-related characteristics to forecast academic performance. Their research emphasized the importance of identifying both advanced and slow learners using data mining techniques. The study highlighted the role of predictive analytics in supporting personalized educational strategies and improving academic outcomes. M. Pandey and S. Taruna [23] introduced a hybrid student performance prediction framework that combined Aggregating One-Dependence Estimators (AODE) with the K-Nearest Neighbor (KNN) algorithm. The proposed hybrid classifier utilized probabilistic rule-based learning for predicting engineering students' academic performance. The model was evaluated on multiple student datasets and compared with classifiers such as KSTAR, OneR, Naïve Bayes, and NBTree. Experimental results demonstrated improved predictive performance and classification accuracy compared to individual classifiers.

Recent studies have also explored Federated Learning frameworks for prediction tasks beyond education. In [24], a Federated Learning-based earthquake prediction framework was proposed to address issues related to multidimensional data processing, communication latency, and data privacy. The framework utilized localized machine learning models trained on distributed seismic datasets, which were aggregated using the FedQuake algorithm to generate a global predictive model. The proposed system achieved a prediction accuracy of 88.87% using seismic data collected from the Western Himalayas region. This study demonstrated the effectiveness of Federated Learning in handling distributed datasets and developing highly accurate predictive systems while maintaining data privacy and security. Overall, the literature indicates that machine learning, deep learning, and federated learning approaches have significantly enhanced the capability of educational systems to predict student academic performance. However, challenges such as data heterogeneity, model interpretability, privacy preservation, and prediction accuracy remain important research issues. Therefore, further research is required to develop robust, interpretable, and privacy-preserving predictive frameworks capable of improving educational decision-making and student success prediction.

Table 1: Summary of Related Research Studies

Authors	Accuracy
V. Ganesh et al. (2020)	90%
Altabrawee et al. (2019)	91%
M. Pandey & S. Taruna (2016)	94%
Brijesh K. B. & Pal S (2011)	85%
Zahyah A et al. (2016)	83%
Abu, A. (2016)	87%
Sangho S. (2016)	90%
G. Jayanthi & V. Ramesh (2015)	92%
V. Gharat et al. (2020)	89%
Shetty I. D. et al. (2019)	91%
Al Mazidi et al. (2018)	90%

During the last few years, a lot of work has been done on Educational Data Mining, to predict the student's education performance and progress, Educational Dropout ratio prediction, Student placement prediction, and Student final result prediction. There are many advantages of these predictions, and very helpful for establishing as well as implementing new rules and regulations in educational institutions, adopting new teaching methodology, improving placement records, improving curriculum in educational institutions, and using educational data mining techniques to gather different information about the student that affects the student's performance. Students' performance depends upon many internal and external factors. The internal factors that are related to the student's basic information age, gender leaving area marks data and student's basic information age, gender, leaving area marks.

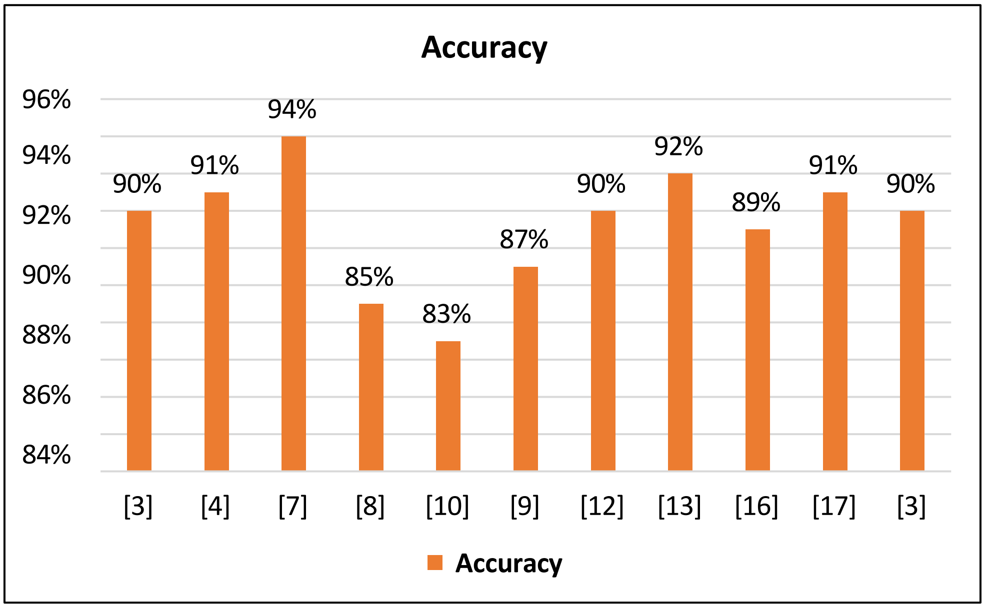

Figure 1 Comparison of Different Studies

3 Proposed Methodology and Implementation

The proposed approach for predicting student academic performance is illustrated in Figure 2. The methodology is designed as a systematic pipeline that integrates data preprocessing, feature optimization, model development, and performance evaluation to ensure accurate and reliable predictions. Initially, a publicly available student performance dataset is obtained from the Kaggle repository (https://www.kaggle.com/datasets). The collected dataset undergoes a comprehensive preprocessing stage, which includes data cleaning, handling missing values, and encoding categorical (string) attributes into a numerical format suitable for machine learning models. This step ensures data consistency and improves the quality of input features. After preprocessing, the dataset is partitioned into training and testing sets, and a cross-validation split is performed using the WEKA data mining tool. This division enables robust model evaluation and helps in assessing the generalization capability of the developed models. In the next stage, Principal Component Analysis (PCA) is applied to reduce and select features. PCA identifies the most significant and informative features from the dataset while eliminating redundant and less relevant attributes. This dimensionality reduction enhances model efficiency and improves predictive performance. Subsequently, multiple machine learning and deep learning models are developed and trained using WEKA software. These models include a Neural Network (NN) as a deep learning approach, along with traditional machine learning classifiers such as Random Forest (RF) and Support Vector Machine (SVM), among others. Each model is trained and tested on the processed dataset to evaluate its predictive capability. Finally, the performance of all implemented models is systematically evaluated and compared using standard evaluation metrics, including training accuracy, testing accuracy, precision, recall, and F1-score. The model demonstrating the highest performance across these metrics is identified as the best-performing approach for student academic performance prediction. This comparative analysis provides a clear understanding of the effectiveness of each algorithm and supports the selection of an optimal predictive model.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

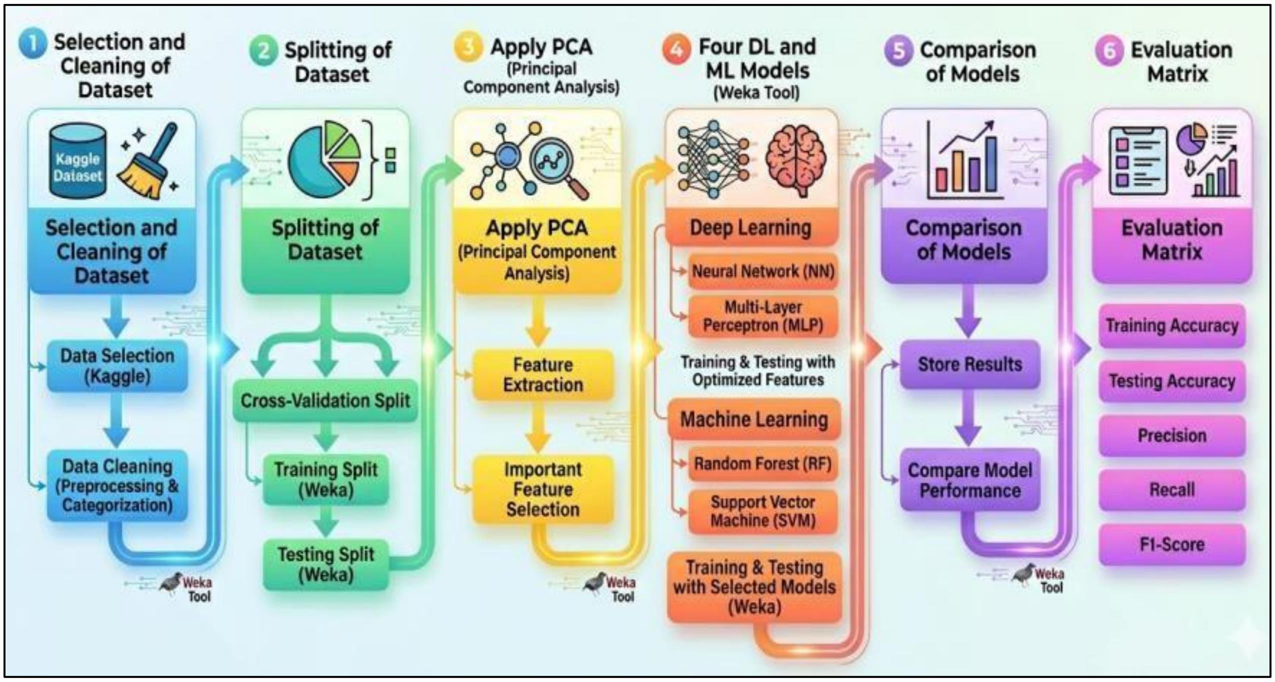

Figure 2 Proposed Methodology

3.1 Description and Selection of the Dataset

The dataset is downloaded from the Kaggle dataset repository, which contains 581 rows and 20 features. The features of the dataset are given as (AGE, GENDER, HS_TYPE, SCHOLARSHIP, FATHER_EDU, MOTHER_EDU, SIBLINGS, KIDS, MOTHER_JOB, FATHER_JOB, PREV_STUDY, PREV_EXAM, NOTES, LISTENS, LIKES_DISCUSS, CLASSROOM, CUML_GPA, EXP_GPA, COURSE ID, GRADE). In these 20 features, some features are binary, some features are numeric, and some features contain string values. Table 2 shows the type and values of each feature. There are 9 binary-type features, 9 categorical-type features, and 2 numeric features. After applying PCA, the number of features is reduced due to the unimportance of their dependency in the accuracy and performance of machine learning models. The last column, Grade, is the label of the dataset that predicts the students in four categories (Good, Low, Average, and Drop).

Table 2 Dataset Features Description

Feature	Type	Values
AGE	Numeric	1 to 10
GENDER	Binary	Male and Female
HS_TYPE	Binary	Yes and No
SCHOLARSHIP	Binary	Yes and No
FATHER_EDU	Categorical	High, Low, Average
MOTHER_EDU	Categorical	High, Low, Average
SIBLINGS	Categorical	5, Less than 5, greater than 5
KIDS	Binary	Yes and No
MOTHER_JOB	Binary	Yes and No
FATHER_JOB	Binary	Yes and No

PREV_STUDY	Categorical	High, Low, Average
PREV_EXAM	Categorical	High, Low, Average
NOTES	Binary	Yes and No
LISTENS	Binary	Yes and No
LIKES_DISCUSS	Binary	Yes and No
CLASSROOM	Categorical	High, Low, Average
CUML_GPA	Categorical	High, Low, Average
EXP_GPA	Categorical	High, Low, Average
COURSE ID	Numeric	Different values
GRADE	Categorical	Low, Good, Average and Drop

The cleaning of the dataset is done by removing the noisy data and fill the blank cells of the dataset with appropriate values. Labeling of String features are done in this step to accurately process the dataset on the machine and deep learning models.

3.2 Splitting of the dataset

After selecting the dataset, the second step is splitting the dataset into training and testing. Weka provides many ways of splitting the dataset, but we use two ways to split our dataset according to our needs. Two ways are given below:

- Training Testing Split
- 10-K fold Split

3.2.1 Training and Testing Split

Training and testing split options are given in Weka for the evaluation of model performance. Training and testing split refers to the process of dividing a dataset into two separate subsets: a training set and a testing set. We use a 75 percent dataset for training of models and the remaining 25 percent dataset for testing of the selected deep learning and machine learning models.

Table 3 Distribution of the dataset

Parameter	Percentage	Count
Training Dataset	75	435 records
Testing Dataset	25	146 records
Total Dataset	100	581 records

A total of 581 records of dataset 435 records are used for the training of models, and the remaining 146 records are used for the testing of models.

4 Experiment and Results

Many ML models are applied to the dataset to evaluate the students' academic performance. Many Evaluation metrics are used in our purpose methodology, such as the accuracy of Models, the precision of models on each class, the recall of models on each class, and the F-1 score for each class.

4.1.1 Accuracy

The accuracy of each ML Model is collected. Accuracy is the degree to which a set of measurements is accurate in relation to its actual value.

4.1.2 Precision

The consistency of many determinations of the same quantity is known as precision. The results are extremely reproducible because there is a reduced difference in values, the better the precision.

4.1.3 Recall

The recall is the capacity of a model to locate all pertinent instances in a data source. The recall is calculated as the sum of the number of true positives minus the number of false negatives.

4.1.4 F-1 score.

The harmonic mean of recall and precision is used to calculate the F1 score.

4.2 Results of NN

We evaluate a neural network using Weka on a training-testing split dataset and a 10-K Folding dataset, and the results of both splits are given in Table 4.

Table 4 Accuracy of NN

Splitting	Accuracy
Training Testing Split	92.41
10-K Folding Split	93.79

Table 4 shows that the accuracy of the training testing split is 92.41 percent, and the accuracy of 10-K Folding is 93.79 percent. The results show that the neural network model is well on 10-K Folding split.

4.2.1 Confusion Matrix of NN

The confusion matrix of NN for both the training and testing split, and 10-K Folding is calculated. The confusion matrix of the training-testing split results is given in Table 5.

Table 5 CM of NN using Training Testing Split

Class	F1-Score	Precision	Recall
Good	95.8	95.8	95.8
Average	93.3	95.5	91.3
Low	89.4	91.3	87.5
Drop	87.7	83.4	92.6
Average	92.5	92.4	92.4

Table 5 shows that the F1-score of NN is achieved by class Good, Average, Low, and Drop as95.8 percent, 93.3 percent, 89.4 percent, and 87.7 percent. Precision of NN is achieved by class Good, Average, Low, and Drop as 95.8 percent, 95.5 percent, 91.3 percent, and 83.4 percent. Recall of NN is achieved by classes Good, Average, Low, and Drop as 95.8 percent, 91.3 percent, 87.5 percent, and 92.6 percent.

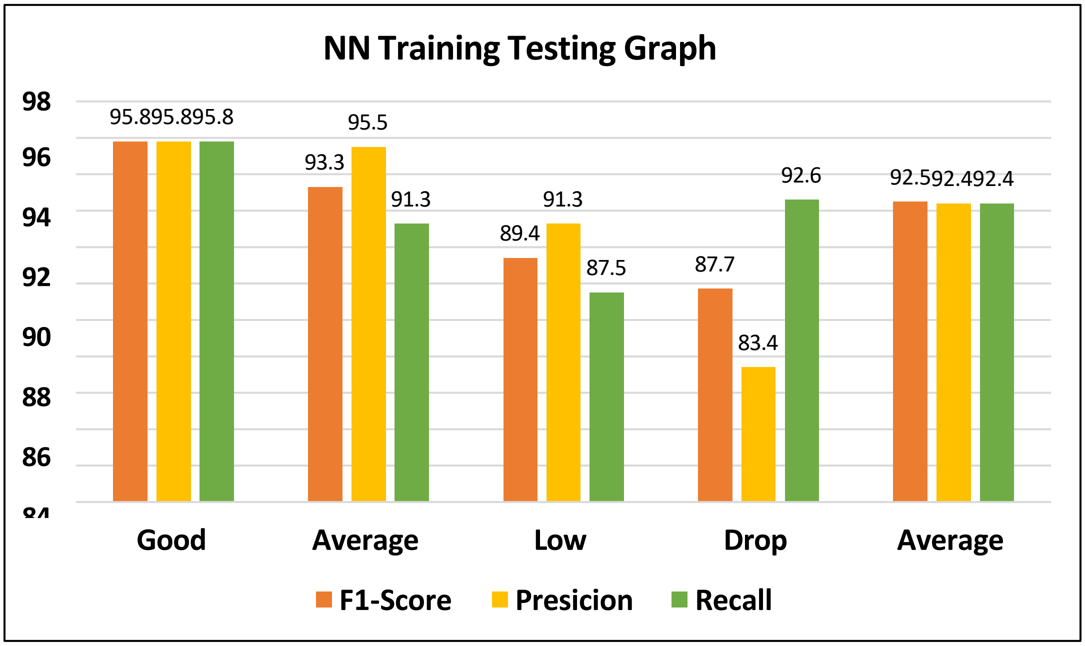

Figure 3 Result Graph of NN for Training Testing Split

The graph in Figure 3 shows the CM results of NN using a training-testing split. The graph shows that Maximum precision is achieved by class Good, maximum Recall is achieved by class Good, and maximum f1-Score is achieved by class Average.

4.3 Result of RF

We evaluate the Random Forest model using Weka on a training-testing split dataset and a 10-K Folding dataset, and the results of both splits are given in Table 6. Table 6 shows that the RF accuracy of the training testing split is 91.3 percent, and the accuracy of 10-K Folding is 90.34 percent. The results show that the RF model performs well on the training-testing split.

Table 6 RF Accuracy

Splitting	Accuracy
Training Testing Split	91.3
10-K Folding Split	90.34

4.3.1 Confusion Matrix of RF

The confusion matrix of RF of both training and testing splits and 10-K Folding is calculated. The confusion matrix of the training-testing split is given in Table 7.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

Table 7 CM of RF using Training Testing Split

Class	F1-Score	Precision	Recall
Good	93.8	93.8	93.3
Average	85.1	93.3	87
Low	82.6	86.4	79.2
Drop	100	100	100
Average	90.3	90.4	90.3

Table 7 shows that the F1-score of RF is achieved by class Good, Average, Low and Drop as 93.8 percent, 85.1 percent, 82.6 percent, and 100 percent. Precision of RF is achieved by class Good, Average, Low, and Drop as 93.8 percent, 93.3 percent, 86.4 percent and 100 percent. Recall of RF is achieved by classes Good, Average, Low, and Drop as 93.3 percent, 87 percent, 79.2 percent, and 100 percent.

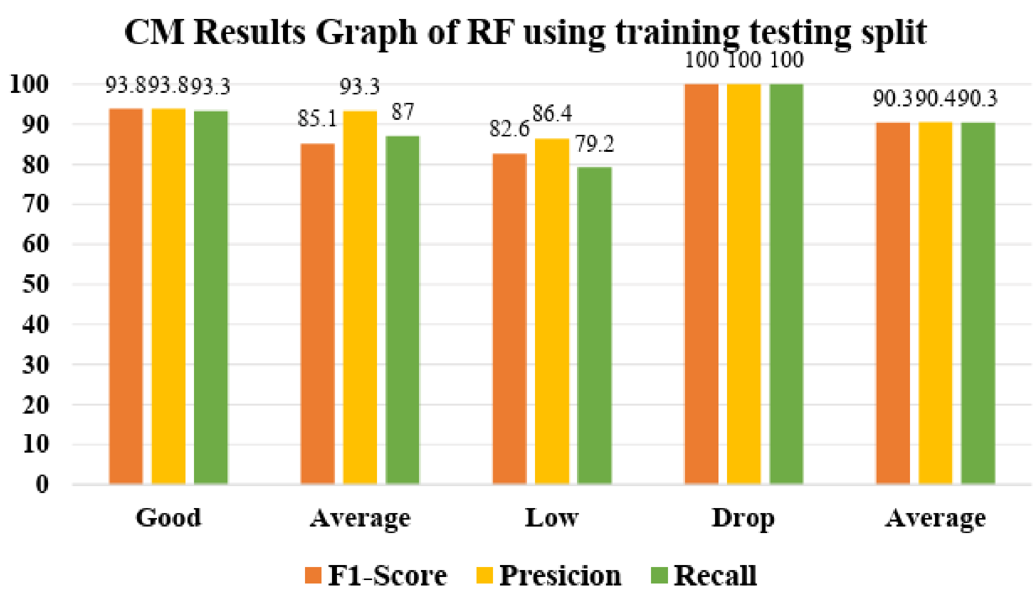

Table 8 CM Results of RF using Training Testing Split

The graph in Figure 8 shows the CM results of RF using a training-testing split. The graph shows that Maximum precision is achieved by class Good, maximum Recall is achieved by class Good, and maximum f1-Score is achieved by class.
Average.

4.4 Results of SVM

We evaluate the SVM model using Weka on the training-testing split dataset, and the 10-K Folding dataset, and the results of both splits are given in Table 9.
Table 9 shows that the accuracy of the training testing split is 43.44 percent, and the accuracy of 10-K Folding is 50.17 percent. The results show that the SVM model is well on 10-K Folding split.

Table 9: Accuracy of SVM

Splitting	Accuracy
Training Testing Split	43.44
10-K Folding Split	50.17

4.4.1 Confusion Matrix of SVM

The confusion matrix of the SVM of both training and testing splits, and 10-K Folding is calculated. The confusion matrix of the training-testing split results is given in Table 10.

Table 10 CM of SVM using Training Testing Split

Class	F1-Score	Precision	Recall
Good	16.4	38.5	10.04
Average	49.6	37.4	73.9
Low	11.8	20	8.3
Drop	75.9	71	81.5
Average	37.2	41.1	43.4

Table 10 shows that the F1-score of SVM is achieved by class Good, Average, Low, and Drop as 16.4 percent, 49.6 percent, 11.8 percent, and 75.9 percent. Precision of SVM is achieved by classes Good, Average, Low, and Drop as 38.5 percent, 37.4 percent, 20 percent, and 71 percent. Recall of SVM is achieved by class Good, Average, Low, and Drop as 10.04 percent, 73.9 percent, 8.3 percent, and 81.5 percent.

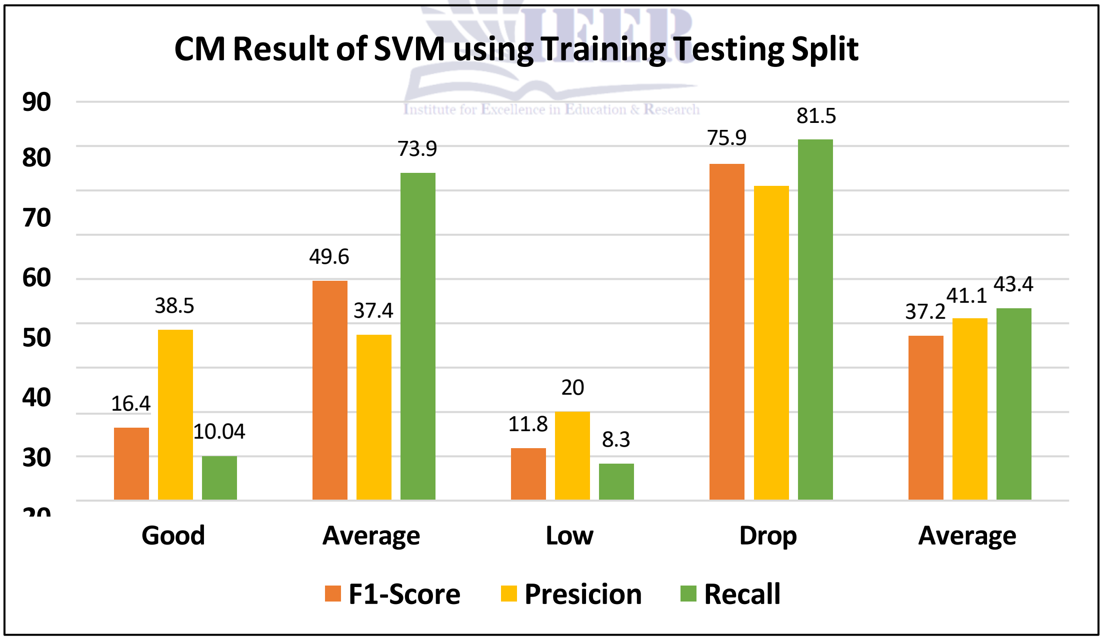

Figure 4 CM of SVM using Training Testing Split

The graph in Figure 4 shows the CM results of SVM using a training-testing split. The graph shows that Maximum precision is achieved by class Drop, maximum Recall is achieved by class Drop and maximum f1-Score is achieved by class Drop.

Table 11: Accuracy Comparison of Models

Models	Training Testing Accuracy	10-K Folding Accuracy
NN	92.41	93.79
RF	91.3	90.34
SVM	43.44	50.17

The graph in Figure 5 shows that maximum accuracy on the training testing split is achieved by the NN model with an accuracy of 92.4 percent, and minimum accuracy is achieved by the SVM model with an accuracy of 43 percent.

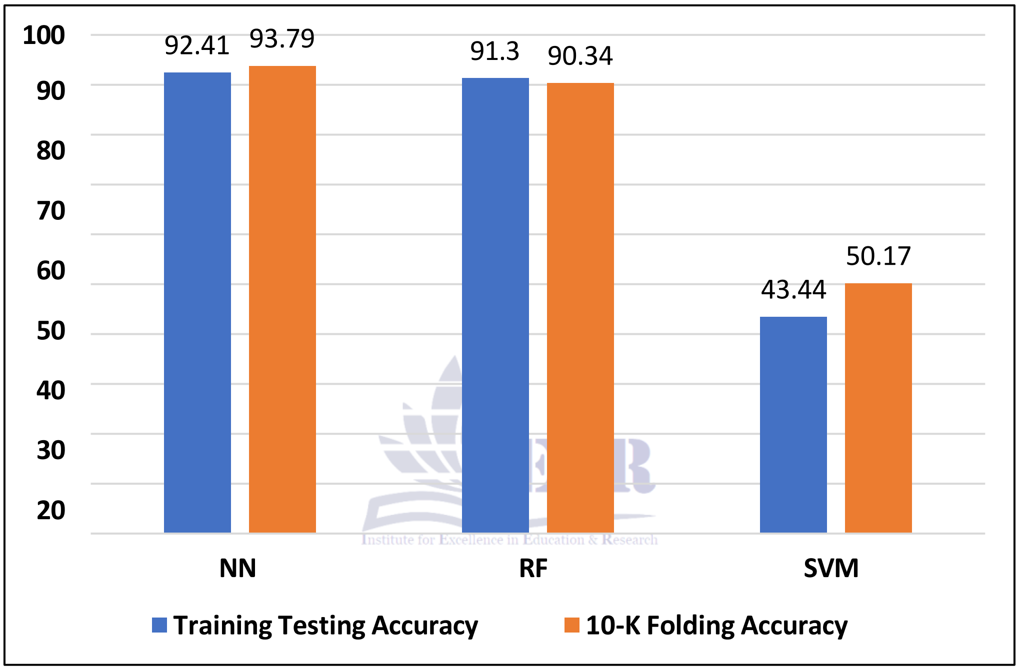

Figure 5 Comparison Graph of Models

4.5 Conclusion

Educational institutes have been increasingly interested in data mining over the past few years as a means of enhancing the quality of education. Accurately predicting student performance in class was crucial to improving the value of academic education. Studies primarily focused on predicting student performance at the higher education level since understanding academic progress was critical for further education. Failure or poor grades during higher education caused students to drop out, making timely performance prediction critical for retention. Educational data mining (EDM) is the process of extracting. valuable data and patterns from accumulated academic data to predict student success.

Our research aimed to evaluate student performance using multiple classification strategies, using a Kaggle repository database as the dataset source. We removed duplicate records and filled in any empty fields with appropriate information. We applied six different classifiers, including deep learning methods like Neural Network and data mining methods like Random Forest, to the dataset using the performance measure tool Weka. We also applied Principal Component Analysis (PCA) to optimize feature extraction. We evaluated all models using the Training and Testing split and 10-K Fold options

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

available in Weka. Finally, we calculated the Training Accuracy, Testing Accuracy, Precision, Recall, and F1-Score for each model and compared the results. Our findings showed that Neural Network and Random Forest produced the best results compared to other models. In the future, this work can be expanded by adding more records to the dataset, and reinforcement or federated learning can be used to improve the accuracy of the model. Other datasets can be used to classify the Grades and fine-tune the classes.

REFERENCES

Al-Mahmoud, H., & Al-Razvan, M. (2015). Arabic Text Mining: A Systematic Review of the Published Literature. 2015 International Conference on Cloud Computing, ICCC 2015, July. https://doi.org/10.1109/CLOUDCOMP.2015.7149632.

Altabrawee, H., Ali, O. A. J., & Ajmi, S. Q. (2019). Predicting Students' Performance Using Machine Learning Techniques. JOURNAL OF UNIVERSITY OF BABYLON for Pure and Applied Sciences, 27(1), 194–205. https://doi.org/10.29196/jubpas.v27i1.2108.

Yacef, K., Zaïane, O., Hershkovitz, A., & Yudelson, M. (2012). EDM 2012 5 th International Conference on Educational Data Mining.

Al Mazidi, A., & Abusham, E. (2018). Study of general education diploma students' performance and prediction in the Sultanate of Oman, based on data mining approaches. International Journal of Engineering Business Management, 10, 1–11.https://doi.org/10.1177/1847979018807020.

Estrada-Danell, R. I., Zamarripa-Franco, R. A., Zúñiga-Garay, P. G., & Martínez-Trejo, I. (2016). Aportaciones desde la minería de datos al proceso de captación de matrícula en instituciones de Educación Superior particulares. Revista Electrónica Educare, 20(3), 1. https://doi.org/10.15359/ree.20-3.11.

Imran, M., Latif, S., Mehmood, D., & Shah, M. S. (2019). Student academic performance prediction using supervised learning techniques. International Journal of Emerging Technologies in Learning, 14(14), 92–104. https://doi.org/10.3991/ijet.v14i14.10310

J. Bayer, H. Bydzovska, J. Geryk, T. Obsivac, and L. Popelinsky, "Predicting drop-out from social behaviour of students," in International Educational Data Mining Society. ERIC, 2012.

Al Mazidi, A., & Abusham, E. (2018). Study of general education diploma students' performance and prediction in the Sultanate of Oman, based on data mining approaches. International Journal of Engineering Business Management, 10, 1–11.https://doi.org/10.1177/1847979018807020.

Estrada-Danell, R. I., Zamarripa-Franco, R. A., Zúñiga-Garay, P. G., & Martínez-Trejo, I. (2016). Aportaciones desde la minería de datos al proceso de captación de matrícula en instituciones de Educación Superior particulares. Revista Electrónica Educare, 20(3), 1. https://doi.org/10.15359/ree.20-3.11.

Imran, M., Latif, S., Mehmood, D., & Shah, M. S. (2019). Student academic performance prediction using supervised learning techniques. International Journal of Emerging Technologies in Learning, 14(14), 92–104. https://doi.org/10.3991/ijet.v14i14.10310

J. Bayer, H. Bydzovska, J. Geryk, T. Obsivac, and L. Popelinsky, "Predicting drop-out from social behaviour of students," in International Educational Data Mining Society. ERIC, 2012.

Yassin, N. A., M Helali, R. G., & Mohamad, S. B. (2017). Predicting Student Academic Performance in KSA using Data Mining Techniques. Jour 300nal of Information Technology & Software Engineering, 07(05). https://doi.org/10.4172/2165-7866.1000213.

G. Jayanthi & Dr.V. Ramesh. (2015). Design of an Academic Performance Prediction System Using Multi-Layer Perceptron. 1(1), 9–15.

L.Pandeeswari, K. Rajeswari (2014). Student Academic Performance Using Data Mining Techniques, International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 3, Issue. 10, October 2014, pg.726 – 731.

Bacher-Hicks, A., Chin, M., Kane, T., & Staiger, D. (2017). An Evaluation of Bias in Three Measures of Teacher Quality: Value-Added, Classroom Observations, and Student Surveys. National Bureau of Economic Research.https://doi.org/10.3386/w23478 .

Al-Mahmoud, H., & Al-Razgan, M. (2015). Arabic Text Mining: A Systematic Review of the Published Literature. 2015 International Conference on Cloud Computing, ICCC 2015, July. https://doi.org/10.1109/CLOUDCOMP. 2015.7149632.

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 the prediction of student performance using a 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 EfficientNet. 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 an efficient deep fake detection technique on a video dataset using a 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 is 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.

Mahnoor Zaman, Nosheen Fatima, Muhammad Sajid Maqool, Dr. Naeem Aslam, Rubaina Nazeer, & Hira Saleem. (2026). Ingredient: 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 Multi-Temporal 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 EfficientNet. VFAST Transactions on Software Engineering, 12(3), 147-165.

Aslam, N., Maqbool, M. S., Nadeem, M., & Saleem, H. (2026). DEEPFAKESHIELD: ENHANCED VIDEO AUTHENTICITY DETECTION VIA CONVOLUTIONAL VISION TRANSFORMER. Spectrum of Engineering Sciences, 4(3), 1650-1665.

Aslam, M., Maqbool, M. S., Aoun, M., Aslam, N., Razzaq, A. M., & Ali, S. (2026). HIGH-PERFORMANCE AND EFFICIENT BRAIN TUMOR SEGMENTATION FOR ENHANCED CLINICAL ANALYSIS. Spectrum of Engineering Sciences, 4(3), 195-210.

Maqbool, M. S., Zahra, S. R., Ismail, S., Nadeem, M., Fatima, N., & Ahmad, J. (2026). A CNN-BASED FRAMEWORK FOR EFFICIENT DETECTION OF EYE DISEASE IN FUNDUS IMAGES. Spectrum of Engineering Sciences, 4(4), 1157-1169.

Farwa Zainab, Farwa Nazim, Muhammad Kashaf, Naeem Aslam, & Muhammad Sajid Maqbool. (2026). PREDICTIVE ANALYTICS FOR CUSTOMER CHURN IN SUBSCRIPTION-BASED BUSINESSES USING MACHINE LEARNING. Spectrum of Engineering Sciences, 4(4), 596–618. Retrieved from https://thesesjournal.com/index.php/1/article/view/2460.

Meiraj Aslam, Mohammad Sajid Maqbool, Muhammad Aoun, Naeem Aslam, Abdul Manan Razzaq, Abdul Manan Razzaq, & Salman Ali. (2026). HIGH-PERFORMANCE AND EFFICIENT BRAIN TUMOR SEGMENTATION FOR ENHANCED CLINICAL ANALYSIS. Spectrum of Engineering Sciences, 4(3), 195–210. Retrieved from https://thesesjournal.com/index.php/1/article/view/2169.

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.

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 from.https://policyrj.com/1/article/view/1770.

Rabia Hassan, Muhammad Sajid Maqbool, Dr. Naeem Asalam, Ariba Afzal, Hira Saleem, & Muqadas Nadeem. (2026). AN IN-DEPTH STUDY ON STUDENTS' PERFORMANCE EVALUATION USING MULTIPLE MACHINE LEARNING CLASSIFIERS AND DATA ANALYTICS APPROACHES. Policy Research Journal, 4(4), 86–99. Retrieved from.https://policyrj.com/1/article/view/1771.

Rabia Ikhlaq, Muhammad Sajid Maqbool, Hira Saleem, Dr. Naeem Aslam, Zeeshan Manzoor, & Muqadas Nadeem. (2026). DEEP CONVOLUTIONAL NEURAL NETWORKS FOR AUTOMATED BREAST CANCER DIAGNOSIS. Policy Research Journal, 4(4), 170–182. Retrieved from.https://policyrj.com/1/article/view/1795.

Izhar, Dr. Naeem Aslam, Muhammad Sajid Maqbool, Muqadas Nadeem, & Hira Saleem. (2026). DEEPFAKESHIELD: ENHANCED VIDEO AUTHENTICITY DETECTION VIA CONVOLUTIONAL VISION TRANSFORMER. Spectrum of Engineering Sciences, 4(3), 1650–1665. Retrieved from https://thesesjournal.com/index.php/1/article/view/2360

Altabrawee, H., Ali, O. A. J., & Ajmi, S. Q. (2019). Predicting Students' Performance Using Machine Learning Techniques. JOURNAL OF UNIVERSITY OF BABYLON for Pure and Applied Sciences, 27(1), 194–205. https://doi.org/10.29196/jubpas.v27i1.2108.

Karthikeyan, V. G., Thangaraj, P., & Karthik, S. (2020). Towards developing a hybrid educational data mining model (HEDM) for efficient and accurate student performance evaluation. Soft Computing, 24(24), 18477–18487. https://doi.org/10.1007/s00500.020.05075-4.

Al Mazidi, A., & Abusham, E. (2018). Study of general education diploma students' performance and prediction in the Sultanate of Oman, based on data mining approaches. International Journal of Engineering Business Management, 10, 1–11. https://doi.org/10.1177/1847979018807020.

Estrada-Danell, R. I., Zamarripa-Franco, R. A., Zúñiga-Garay, P. G., & Martínez-Trejo, I. (2016). Aportaciones desde la minería de datos al proceso de captación de matrícula en instituciones de Educación Superior particulares. Revista Electrónica Educare, 20(3), 1. https://doi.org/10.15359/ree.20-3.11.

Tanveer, K., Aslam, N., Saleem, H., Maqbool, M. S., Akhter, M., & Rashid, M. H. (2026). INTEGRATING BLOCKCHAIN AND MACHINE LEARNING FOR ROBUST IOT DATA INTEGRITY IN CRITICAL INFRASTRUCTURES. Spectrum of Engineering Sciences, 4(3), 1875-1886.

Imran, M., Latif, S., Mehmood, D., & Shah, M. S. (2019). Student academic performance prediction using supervised learning techniques. International Journal of Emerging Technologies in Learning, 14(14), 92–104. https://doi.org/10.3991/ijet.v14i14.10310.