WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

AI-Based Student Performance Prediction and Academic Recommendation System Using Deep Learning

P. Jai Ruby, A. Paulin

Associate Professor, Department of Computer Application and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli.
Student, Reg. No: 25081205300112021, Department of Computer Application and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli.

ABSTRACT:

Predicting student academic performance at an early stage is important for identifying students who may require additional academic support. Traditional evaluation methods mainly depend on examination results and manual assessment, which may not consider the combined influence of academic, behavioural, demographic, and life style-related factors. This paper presents an AI-based student performance prediction and academic recommendation system using deep learning to address this need.

The proposed system analyses 19 student-related input features, including study hours, attendance, previous scores, parental involvement, access to learning resources, motivation level, sleep hours, and other relevant factors. The Student Performance Factors dataset is preprocessed by removing duplicate records, handling missing values, encoding categorical features using label encoding, and applying feature scaling. An Artificial Neural Network (ANN)-based deep learning model is developed to predict the student's examination score. The model is evaluated using Mean Absolute Error (MAE) and the coefficient of determination (R^2).

The experimental results recorded for the developed model show an MAE of 0.97 and an R^2 score of 0.70 on the testing dataset. Based on the predicted examination score, the system generates an academic grade and provides recommendations to support student improvement. The application is developed using Python Tkinter, which provides a graphical user interface for entering student details and viewing prediction results. SQLite database integration is used to maintain prediction history for future reference.

The proposed system combines performance prediction, grade generation, academic recommendations, and result management within a single application. It demonstrates the potential of deep learning as a supportive tool for educational performance analysis and academic decision-making.

Keywords: Artificial Intelligence, Deep Learning, Student Performance Prediction, Artificial Neural Network, Educational Data Mining, Academic Recommendation.

I. INTRODUCTION

Artificial Intelligence (AI) has become an important technology in the field of education. The increasing availability of student academic and behavioural data provides opportunities to analyse learning patterns and predict academic performance. Predicting student performance at an early stage helps educational institutions identify students who may face academic difficulties and provide suitable support to improve their learning outcomes [1], [2].

Educational Data Mining (EDM) is an emerging research area that uses data analysis techniques to discover useful patterns from educational datasets. It helps identify the factors that influence students' academic achievement, such as attendance, study hours, previous examination scores, parental involvement, learning resources, and student behaviour [2], [3]. By analysing these factors, educational institutions can gain useful insights into students' learning progress and academic performance.

Machine Learning techniques have been widely used for predicting student academic performance. Algorithms such as decision trees, random forests, and support vector machines can analyse student-related data and estimate exam results. Previous studies have demonstrated the usefulness of these techniques in educational performance prediction [1], [4]. However, prediction alone may not provide sufficient support for students who require academic improvement.

To address this limitation, this research proposes an **AI-based student performance prediction and academic recommendation system using deep learning**. The proposed system analyses 19 student-related features, including academic, behavioural, and demographic information, to predict exam scores. The dataset undergoes preprocessing, categorical encoding, and feature scaling before being used for model training and prediction.

The system also generates grades and provides personalised academic recommendations based on the predicted exam scores. Students who are identified as being at academic risk can receive suitable suggestions to improve their performance. A Python-based Tkinter graphical user interface allows users to enter student details and view prediction results. In addition, an SQLite database is used to maintain prediction history for future reference.

The main objective of this research is to develop an intelligent system that combines student performance prediction with academic recommendations. The proposed approach aims to support early identification of students who may need additional academic assistance and help educational institutions make data-driven decisions to improve student learning outcomes.

II. LITERATURE REVIEW

Recent studies have applied artificial intelligence, machine learning, and educational data mining to analyze student data and predict academic performance. These approaches help identify academic patterns and factors associated with students' examination results.

Yağcı [1] examined machine learning algorithms such as Random Forest, Support Vector Machine, Logistic Regression, Naïve Bayes, and K-Nearest Neighbour using data from 1,854 university students.

Roslan and Chen [2] reviewed educational data mining research published between 2015 and 2021, while Chaka [3] summarized reviews on student performance prediction methods, algorithms, and tools.

Alalawi et al. [4] reviewed the use of machine learning for student performance prediction and discussed research trends and practical educational interventions.

Balcioğlu and Artar [5] studied the prediction of students' academic performance using machine learning. Adane et al. [6] analyzed the performance of machine learning algorithms for academic performance prediction. Meghji et al.

[7] explored the use of educational data mining to predict student academic performance, and Nayak et al. [8] focused on a machine learning-based classification model for mining educational data.

Sengupta [9] investigated the identification of a minimal set of features for predicting student performance through educational data mining. Wang et al. [10] presented ProbSAP, a system for student academic performance prediction. Overall, these studies establish the value of data-driven prediction.

The present work builds on this research by combining score prediction with grade generation, academic recommendations, at-risk analysis, a user-friendly interface, and prediction-history management in one application.

III. PROPOSED METHODOLOGY

The proposed system follows an integrated artificial intelligence and deep learning framework for student performance prediction and academic recommendation. The overall workflow consists of dataset collection, data preprocessing, feature encoding, feature scaling, model development, score prediction, grade generation, academic recommendation, at-risk analysis, and application-based result management.

A. Dataset and Preprocessing

The Student Performance Factors dataset is used for developing the proposed system. The dataset contains academic, behavioural, demographic, and lifestyle-related attributes that may influence students' examination performance. The input features include hours studied, attendance, parental involvement, access to resources, extracurricular activities, sleep hours, previous scores, motivation level, internet access, tutoring sessions, family income, teacher quality, school type, peer influence, physical activity, learning disabilities, parental education level, distance from home, and gender.

The target variable used for prediction is Exam_Score. Before model training, duplicate records are removed from the dataset. Missing values are handled using forward filling to prepare the data for further processing. The input features are separated from the target variable before model development.

B. Feature Encoding and Scaling

The dataset contains both numerical and categorical attributes. Label encoding is applied to convert categorical values into numerical representations that can be processed by the prediction model. The encoders are saved and reused during prediction to maintain consistency in data transformation.

Feature scaling is applied to the input features to bring numerical values into a comparable range. This preprocessing step helps prepare the dataset for model training and prediction. The trained scaler is also reused when processing new student details.

C. Machine Learning Model Comparison

Before developing the final deep learning model, traditional machine learning regression algorithms are evaluated for performance comparison. The algorithms considered are Decision Tree Regression, Random Forest Regression, and Support Vector Regression.

These models are trained using the prepared dataset and evaluated using Mean Absolute Error (MAE) and R^2 Score. The evaluation results are used as baseline measurements for comparing prediction performance.

D. ANN-Based Student Performance Prediction

An Artificial Neural Network (ANN) is used as the deep learning approach for predicting student examination scores. The processed student-related features are provided as input to the model, and Exam_Score is used as the target output.

The dataset is divided into training and testing sets to develop and evaluate the model. The ANN learns relationships between student-related attributes and examination performance during training. After training, the model is used to predict examination scores for new student records.

E. Academic Recommendation and At-Risk Analysis

After predicting the examination score, the system generates a corresponding academic grade based on predefined score ranges. It also provides academic recommendations according to the predicted performance.

Students with lower predicted scores can be identified as potentially at-risk students. The system provides suitable suggestions, such as improving study hours, maintaining attendance, completing assignments, and following a regular study schedule. These recommendations are intended to support academic improvement.

F. Class-Level Analysis and Application

A Python-based graphical user interface is developed using the Tkinter library. It allows users to enter student details and obtain predicted examination scores, grades, and academic recommendations.

The application also supports class-level prediction and result analysis. Users can view student prediction results and use available options for searching, filtering, sorting, and visualising the results. Prediction history is maintained using an SQLite database, and class-level prediction results can be exported in CSV format.

G. Performance Evaluation Metrics

The performance of the prediction models is evaluated using Mean Absolute Error (MAE) and the R^2 score.

Mean Absolute Error measures the average absolute difference between actual and predicted examination scores. A lower MAE indicates that the predicted values are closer to the actual values.

The R^2 score measures how well the model explains the variation in the target variable. A higher R^2 score generally indicates a better fit of the model to the test data.

These evaluation metrics are used to analyse and compare the prediction performance of the developed models.

Proposed System Workflow

IV. RESULTS AND DISCUSSION

The proposed system was evaluated to analyse its ability to predict students' examination performance using academic, behavioural, and demographic features. Different regression models were considered during experimentation, and their performance was compared using Mean Absolute Error (MAE) and R^2 Score.

A. Comparison of Machine Learning Models

Decision Tree Regression, Random Forest Regression, and Support Vector Regression were evaluated using the prepared student performance dataset. The experimental results are presented in Table.

PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS

Model	**MAE**	**R^2 Score**
Decision Tree Regression	1.69	0.28
Random Forest Regression	1.14	0.65
Support Vector Regression	1.22	0.65

The results show that Random Forest Regression achieved the lowest MAE of 1.14 among the three traditional machine learning models. Random Forest Regression and Support Vector Regression achieved the same R^2 score of 0.65. Decision tree regression obtained an MAE of 1.69 and an R^2 score of 0.28.

This comparison provides a baseline for evaluating student examination score prediction using traditional machine learning techniques.

B. ANN-Based Prediction and Application

The proposed system uses an Artificial Neural Network (ANN) as its deep learning approach for predicting student examination scores. The model analyses the processed input features and generates a predicted examination score.

Based on the experimental results recorded for the ANN model, it achieved a Mean Absolute Error (MAE) of 0.97 and an R^2 Score of 0.70. These results indicate the model's prediction performance on the testing dataset.

After prediction, the application generates a grade based on the predicted score and provides academic recommendations. This helps identify students who may require additional academic support.

C. Class-Level Analysis and Result Management

The developed Python Tkinter application provides an interface for entering student information and viewing prediction results. It also supports class-level prediction and analysis.

The application provides options for viewing student results, searching, filtering, sorting, and visualising prediction information. SQLite database integration is used to maintain prediction history, while CSV export allows users to save class-level prediction results for further analysis.

Overall, the results demonstrate how the proposed system combines student performance prediction, grade generation, academic recommendations, and result management within a single application.

CONCLUSION

This paper presented an AI-based student performance prediction and academic recommendation system using deep learning. The proposed system analyses 19 student-related input features, including academic, behavioural, and demographic factors, to predict examination performance. The dataset is preprocessed through duplicate removal, missing-value handling, categorical encoding, and feature scaling before training the deep learning model.

The developed model achieved a Mean Absolute Error (MAE) of 0.97 and an R^2 Score of 0.70 on the testing dataset. Based on the predicted examination score, the system generates grades and provides academic recommendations to support students in improving their performance. A Python Tkinter-based graphical user interface enables users to interact with the system, while SQLite database integration maintains prediction history for future reference.

The proposed system demonstrates the potential of deep learning in student performance prediction and academic support. In the future, the system can be enhanced using larger datasets and Explainable AI techniques to improve prediction reliability and provide more detailed explanations for individual predictions.

In the future, the system can be enhanced using larger datasets and explainable AI techniques to improve prediction reliability and provide more detailed explanations for individual predictions.

REFERENCES

[1] M. Yağcı, “Educational data mining: Prediction of students’ academic performance using machine learning algorithms,” Smart Learning Environments, vol. 9, article 11, 2022. doi: 10.1186/s40561-022-00192-z.

[2] M. H. Bin Roslan and C. J. Chen, “Educational data mining for student performance prediction: A systematic literature review (2015–2021),” International Journal of Emerging Technologies in Learning, vol. 17, no. 5, pp. 147–179, 2022. doi: 10.3991/ijet.v17i05.27685.

[3] C. Chaka, “Educational data mining, student academic performance prediction, prediction methods, algorithms and tools: An overview of reviews,” Journal of E-Learning and Knowledge Society, vol. 18, no. 2, pp. 58–69, 2022. doi: 10.20368/1971-8829/1135578.

[4] K. Alalawi, R. Athauda, and R. Chiong, “Contextualising the current state of research on the use of machine learning for student performance prediction: A systematic literature review,” Engineering Reports, vol. 5, no. 12, e12699, 2023. doi: 10.1002/eng2.12699.

[5] Y. S. Balcioğlu and M. Artar, “Predicting academic performance of students with machine learning,” Information Development, vol. 41, no. 3, 2025. doi: 10.1177/02666669231213023.

[6] M. D. Adane, J. K. Deku, and E. K. Asare, “Performance analysis of machine learning algorithms in prediction of student academic performance,” Journal of Advances in Mathematics and Computer Science, vol. 38, no. 5, pp. 74–86, 2023. doi: 10.9734/jamcs/2023/v38i51762.

[7] A. F. Meghji, F. B. Shaikh, S. A. Wadho, S. Bhatti, and R. K. Ayyasamy, “Using educational data mining to predict student academic performance,” VFAST Transactions on Software Engineering, vol. 11, no. 2, pp. 43–49, 2023. doi: 10.21015/vtse.v11i2.1475.

[8] P. Nayak, S. K. Vaheed, S. Gupta, and N. Mohan, “Predicting students’ academic performance by mining the educational data through machine learning-based classification model,” Education and Information Technologies, vol. 28, no. 11, pp. 14611–14637, 2023.

[9] S. Sengupta, “Towards finding a minimal set of features for predicting students’ performance using educational data mining,” International Journal of Modern Education and Computer Science, vol. 15, no. 3, pp. 44–54, 2023. doi: 10.5815/ijmecs.2023.03.04.

[10] X. Wang, Y. Zhao, C. Li, and P. Ren, “ProbSAP: A comprehensive and high-performance system for student academic performance prediction,” Pattern Recognition, vol. 137, article 109309, 2023. doi: 10.1016/j.patcog.2023.109309.