

Paper1:

Khan, I., et al., An artificial intelligence approach to monitor student performance and devise preventive measures. Smart Learning Environments, 2021. 8(1): p. 1-18.

The study examines the importance of activating machine learning algorithms in monitoring student performance to help lecturers identify students at risk of poor results. Several multi-algorithm prediction models were developed.

Data sources:

Data sources were collected from the University of Amman over three semesters, covering 151 cases and ten predictive features of student performance .

Algorithms used:

- k-Nearest Neighbors (K-NN)
- Multilayer Perceptron (ANN)
- Naïve Bayes
- Decision Tree (RepTree)

Strengths:

- The study applied the system in practice and enabled lecturers to use it and make decisions.
- It used real data from the university and actual field experience from a single class after the first evaluation.
- It used multiple algorithms and compared them, which gave realistic results and better accuracy.

Weaknesses:

- The application was carried out on a simple sample of 151 cases in one course and one institution, which reduces the generalizability of the study to other institutions or courses.
- Less focus on behavioral or interactive characteristics, as the study focused on grades, attendance, and GPA.

Paper2:

Deo, R.C., et al., Modern artificial intelligence model development for undergraduate student performance prediction: An investigation on engineering mathematics courses. IEEE Access, 2020.

This study relied on the Extreme Learning Machine (ELM) intelligent system to analyze the performance of engineering students at an Australian university and used data based on actual attendance and remote attendance for a period of six years .

Algorithms used:

Use of student data over six years with two types of attendance patterns: in-person and remote .

The study used the ELM model and compared it with the Random Forest and Volterra models.

The ELM model proved its effectiveness and high predictive power .

Data sources:

Six years of student performance data from the University of Southern Queensland, Australia

Included several in-person and distance learning courses

Strengths:

- The data used is real and covers a long period of time
- The ELM model outperformed the RF and Volterra models
- Coverage of two types of education: in-person and distance learning

Weaknesses:

- The study focused more on prediction than on follow-up and the actual and practical impact after the predictions.
- Limited generalizability, as it was only applied to one course and one university .
- The study did not focus on analyzing behavior, but only on grades.
- There is a need to assess the impact of predictions on intervention decisions or educational improvements .

Peper3:

Zeineddine, H., U. Braendle, and A. Farah, Enhancing prediction of student success: Automated machine learning approach. Computers & Electrical Engineering, 2021. 89: p. 106903

The study proposes the use of machine learning to predict student success and improve the accuracy of results. It relied on available data and aims to automatically identify models and algorithms to identify students at risk of failure early on for intervention and improvement of their academic performance.

Algorithms used:

The study used several models :

- SVM
- K NN
- Random Forest
- Gradient Boosting
- AdaBoost

The study proposed testing multiple models using AutoML and automatically selecting the optimal model based on data accuracy.

Data sources

The study used university student data comprising 461 cases for a specific period between 2016 and 2021

Strengths:

- Use of early data, providing an opportunity for improvement for students at risk of failure
- Use of more than one model provides flexibility and better accuracy for predictions and determining the optimal model
- Use of real data from an educational institution

Weaknesses:

- The study lacks details on how the AutoML model is used.
- Reliance on data from a single institution and specific students, which affects the generalizability of the results.
- Focus on prediction rather than how it is applied to students and measuring its impact.
- Behavioral and interactive features were not included .

Paper 4:

Hashim, A.S., W.A. Awadh, and A.K. Hamoud. Student performance prediction model based on supervised machine learning algorithms. in IOP conference series: materials science and engineering. 2020. IOP Publishing.

The study aims to develop a predictive model using machine learning algorithms to predict student performance and enable institutions to identify those who may fail academically or suffer from weakness so that administrators or teachers can intervene and provide them with the necessary support.

Algorithms used:

- Decision Tree
- Naïve Bayes
- Logistic Regression
- Support Vector Machine (SVM)
- K Nearest Neighbour (KNN)
- Sequential Minimal Optimisation (SMO)
- Neural Network / Multi layer Perceptron

Data sources:

Data on students at the Faculty of Computer Science and Information Technology, University of Basra, between 2018 and 2019, with 499 cases

Strengths:

- The study used a variety of algorithms and did not rely on a single model
- Use of real data.
- Weaknesses:
- The sample was from a single university, which reduces the ability to generalize the results.
- The impact was not measured after the actual application of the prediction on students.

Paper 5

Lau, E.T., L. Sun, and Q. Yang, Modelling, prediction and classification of student academic performance using artificial neural networks. SN Applied Sciences, 2019. 1(9):

The study aims to solve the problem of the inadequacy of traditional statistical methods for predicting student performance by collecting traditional statistical analysis and then building an artificial neural network (ANN) model for prediction and academic classification.

Algorithms used :

- ANN neural networks

Data sources:

The study included 1,000 cases from a Chinese university and three university programs over four years and used 11 variables

Strengths:

Combining traditional statistical analysis with neural network analysis allows for better understanding before model construction.

Use of a good sample size of about 1,000 students from a real university.

Good preparation of the model and clear reporting of performance indicators, providing credibility and transparency in the evaluation.

Weaknesses:

The study was conducted at a single university, which reduces its generalizability .
The study did not address behavioral data.

Paper6:

The study seeks to use recurrent neural networks (RNNs) to predict final academic performance and identify those at risk, enabling teachers to intervene early to improve student levels.

Algorithms used:

Recurrent Neural Network (RNN)

Artificial Neural Network (ANN)

Deep Neural Network (DNN)

Data sources:

Data from Kaggle was used .

Strengths:

- The use of RNN to process sequential data is a suitable option for predicting academic performance.
- Comparison between several RNN models versus ANN and DNN models showed that the RNN model outperformed the other models .
- Use of real data rather than synthetic data, with multiple advantages .

Weaknesses:

- The sample and dataset are limited to one source and do not indicate whether they are diverse or extensive .
- The study did not address how teachers and educational institutions can use prediction in practice and intervene directly after the results are available.

Paper 7 :

Pallathadka, H., et al., Classification and prediction of student performance data using various

machine learning algorithms. Materials today: proceedings, 2023. 80: p. 3782-3785.

The study explores how machine learning algorithms can be used to classify student performance in universities early on to reduce failure rates.

Algorithms used:

- Naïve Bayes
- ID3
- C4.5
- SVM

Data sources:

Data from the UCI repository was used, such as the Student Performance Data Set.

which includes many features from previous schools and institutions.

Strengths:

- Use of multiple algorithms to determine the best one for predictions.
- Use of known data from the UCI platform.

Weaknesses:

- The study was applied to only one institution, which affects the generalization of the results.
- The link between prediction and decision-making is limited.

Paper8 :

Sukhbaatar, O., T. Usagawa, and L. Choimaa, An artificial neural network based early prediction of

failure-prone students in blended learning course. International Journal of Emerging Technologies

in Learning (iJET), 2019. 14(19): p. 77-92.

The study seeks to develop an early prediction system to identify students at risk of academic failure and combines face-to-face and distance learning.

Algorithms used:

- Use of an artificial neural network (ANN)

Data sources:

Data was collected from Kumamoto University for a compulsory course for second-year students, numbering 1,110 over six semesters from 2012 to 2017

Strengths:

- The study is based on real data and two learning styles.
- Extraction of LMS activity features that accurately measure student interaction.
- The ANN model used is suitable for non-linear data and enables early prediction of students at risk of failure.

Weaknesses:

- Limited generalizability to a single course at a single university
- Focus on LMS features related to tests and the digital section, but less coverage of behavioral data
- No clear actual intervention beyond prediction

Connections:

- The studies relied on real student data.
- The studies were conducted within a specific scope within a single educational institution.
- The researchers used more than one algorithm in the studies and selected the best one.
- The models showed remarkable accuracy in the results with a negligible error rate.

Research gap:

- The need to expand the scope of the study so as not to affect the generalizability of the results.
- Adding the largest possible segment of students to contribute to increasing the efficiency and accuracy of the model.
- Adding multiple features, different educational materials, and varying educational levels.
- Need to focus on understanding the model's recommendations to help teachers make decisions
- Need to study the practical application of prediction to guide real educational interventions, determine the effectiveness of the model, and measure its impact on students

Research background:

Artificial intelligence models play an important role in education because they facilitate, improve, and develop the educational process.

Older AI methods have proven their efficiency in data processing and prediction accuracy, but they have shown some weaknesses in dealing with large amounts of data. This has led to the emergence of some modern deep learning models that improve the educational experience for teachers and students in large institutions. Previous studies have shown that researchers have used some of the old models such as:

- Naïve Bayes
- Support Vector Machine – SVM
- K Nearest Neighbour – KNN
- Random Forest
- Decision Tree
- Logistic Regression
- Sequential Minimal Optimisation- SMO

There are also studies that have used deep methods, including:

- k-Nearest Neighbours(K-NN)
- Multilayer Perceptron (ANN)
- Recurrent Neural Network (RNN)
- Deep Neural Network (DNN)

From these methods, it is clear that we can use the RNN model to predict student grades early on and identify weaknesses. Data will be collected from several schools in Saudi Arabia over a period of two years, and noise, loss, and unbalanced distribution will be processed to ensure the effectiveness of the model and the accuracy and comprehensiveness of the predictions to help teachers identify students who are prone to failure and intervene to address their weaknesses. Then comes the role of measuring the impact on students following the intervention after the prediction and determining the extent to which the study's objective has been achieved.

