

NAME- AVROJIT DUTTA

ROLL NO- 25/AIML-A6/NOV-8047

BATCH- AIML with Python Program | InternsElite

November/December 2025

## Student Grade Prediction Using Artificial Neural Network

### Colab Link:

<https://colab.research.google.com/drive/1UI05tkZhC37QRBIINvk7PYKDg5Pu9Qkg?usp=sharing>

Academic performance is influenced by several factors such as study habits, attendance, and past academic history. Identifying students who may struggle academically at an early stage can help educators provide timely support and help students take corrective measures. With the increasing availability of educational data, machine learning techniques can be used to analyse these factors and predict student outcomes effectively.

Artificial Neural Networks (ANNs) are powerful machine learning models capable of learning complex and non-linear relationships from data. Unlike traditional statistical models, ANNs can capture hidden patterns between multiple features and the target outcome. This makes them suitable for predicting student performance, where academic success is often influenced by a combination of interconnected factors.

This project focuses on building a Student Grade Prediction model using an Artificial Neural Network to classify students as either pass or fail based on academic attributes.

### Problem Statement

The objective of this project is to predict student academic performance using an Artificial Neural Network model.

The problem is formulated as a **binary classification task**, where students are classified into two categories:

- Pass
- Fail

The dataset includes academic attributes such as study time, number of absences, and previous failures. The final grade (G3) is used to generate the output label, while the remaining selected attributes serve as input features.

The problem can be stated as:

*Given academic attributes of students, design an Artificial Neural Network that predicts whether a student will pass or fail the final examination.*

### Dataset Description

The dataset used in this project is the **Student Performance Dataset (Mathematics)** obtained from the **UCI Machine Learning Repository**. The data was collected through school records and student questionnaires.

The dataset contains information related to students' academic behaviour and past performance. For this project, relevant numerical features such as study time, absences, and previous failures were selected. The final grade (G3) was converted into a binary target variable representing pass or fail.

## **Methodology**

The following steps were followed to implement the prediction model:

### **1. Data Loading and Exploration**

The dataset was loaded and examined to understand its structure and attributes.

### **2. Target Variable Creation**

The final grade (G3) was transformed into a binary output variable, where students scoring above a defined threshold were labeled as pass and others as fail.

### **3. Feature Selection and Scaling**

Important academic features were selected and standardized to ensure efficient training of the neural network.

### **4. ANN Model Construction**

A feedforward Artificial Neural Network was built using TensorFlow and Keras. The model consists of one input layer, two hidden layers with ReLU activation functions, and one output layer with a sigmoid activation function.

### **5. Model Training and Validation**

The model was trained using the Adam optimizer and binary cross-entropy loss function. Training and validation data were used to monitor performance.

### **6. Evaluation**

The trained model was evaluated using accuracy metrics and a confusion matrix to assess classification performance.

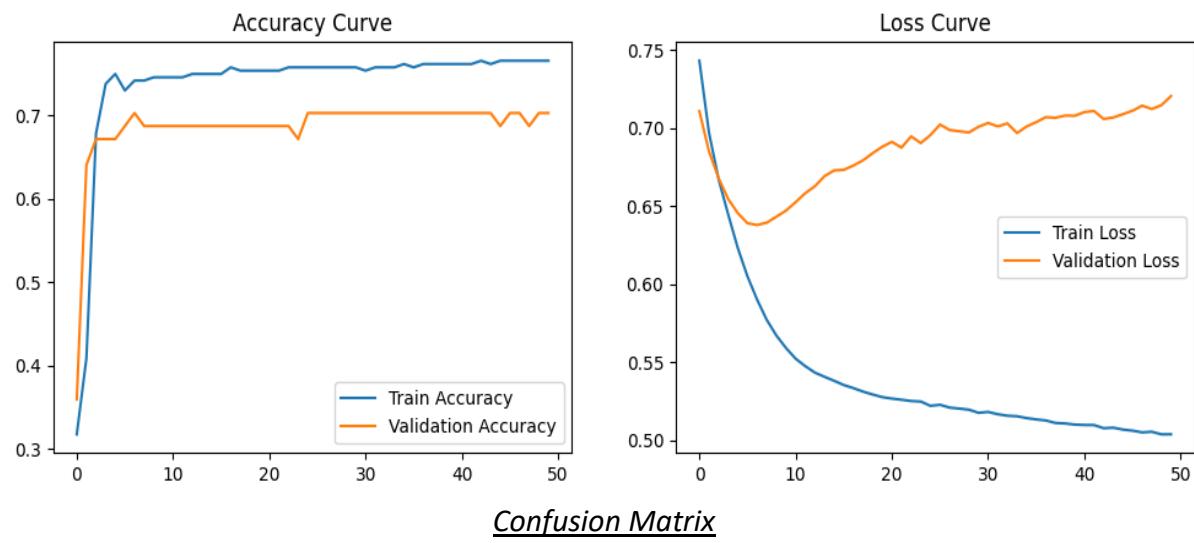
## **Results and Discussion**

The Artificial Neural Network demonstrated effective learning during training, with stable accuracy observed on both training and validation datasets. The model was able to correctly classify a significant number of students into pass and fail categories.

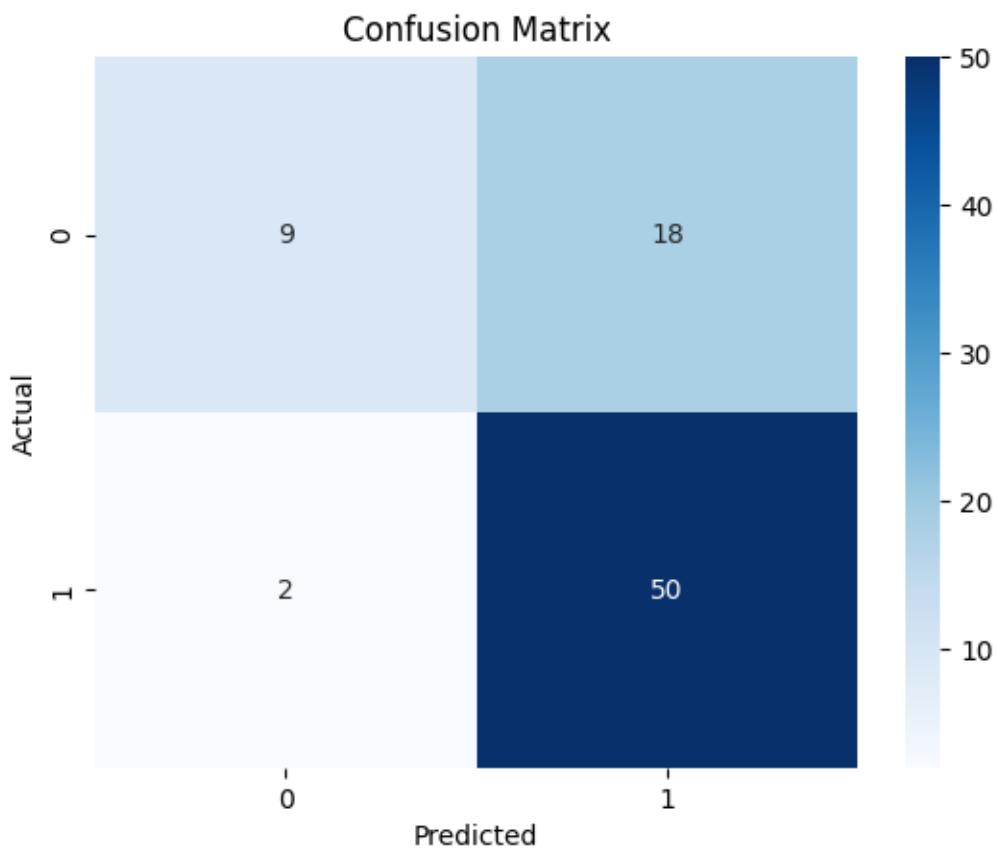
The results indicate that academic factors such as previous failures and attendance play an important role in determining final academic outcomes. Students with fewer absences and no history of failures showed a higher likelihood of passing the final examination.

The confusion matrix further confirmed that the ANN model performed well in distinguishing between the two classes, making it a reliable approach for student performance prediction.

## Accuracy Curve & Loss Curve



Confusion Matrix



## Conclusion

In this project, an Artificial Neural Network was successfully implemented to predict student academic performance using academic attributes. Despite using a limited number of features, the model achieved satisfactory prediction accuracy and demonstrated the effectiveness of neural networks in educational data analysis.

This model can assist educators in identifying students who may require additional academic support and enable students to understand their performance trends in advance. Future

improvements may include incorporating additional behavioural and socio-economic features and deploying the model as a web-based application for real-time usage.

### **Reference**

P. Cortez and A. Silva, *Using Data Mining to Predict Secondary School Student Performance*, Proceedings of the 5th International Conference on Future Business Technology (FUBUTEC 2008), Porto, Portugal, April 2008.