Experiment No. 3

Aim: To apply Logistic Regression for binary classification problems using machine learning, and assess model performance through appropriate evaluation metrics.

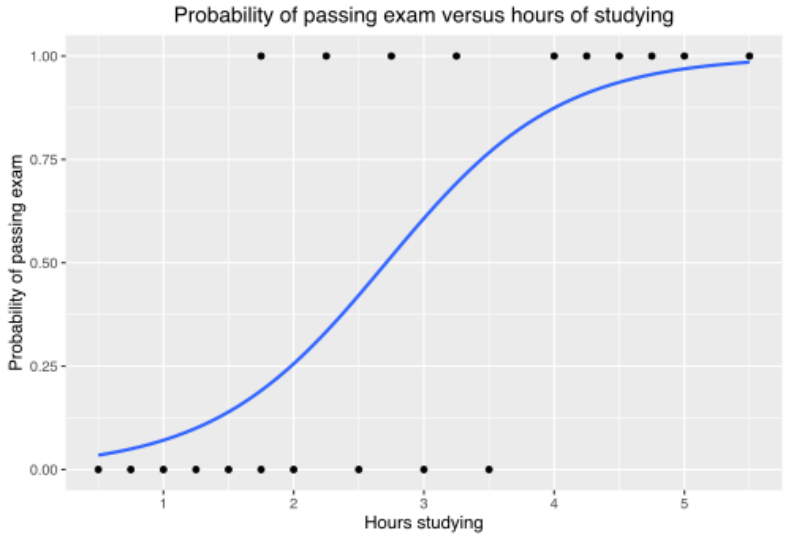
Platform used: Google Colab

Theory:

Logistic Regression

Logistic regression, like linear regression, is a type of linear model that examines the relationship between predictor variables (independent variables) and an output variable (the response, target or dependent variable). The key difference is that linear regression is used when the output is a continuous value—for example, predicting someone's credit score. Logistic regression is used when the outcome is categorical, such as whether a student passed or not.

In logistic regression, the model predicts the probability that a specific outcome occurs. For instance, given someone's hours of studying, we might predict the probability that the student passes or fails the exam. The output of the model is a value between 0 and 1. Based on a threshold, often 0.5, we classify the outcome as either "passed" or "failed." Instead of drawing a straight line through the data as we would in linear regression, logistic regression fits an S-shaped curve to map input values to a probability.



Logistic regression

Logistic regression function

Logistic regression is a statistical model that uses the logistic function, or logit function, in mathematics as the equation between x and y. The logit function maps y as a sigmoid function of x.

$$f(x) = \frac{1}{1 + e^{-x}}$$

Conclusion: Thus, we successfully applied the Logistic Regression for binary classification problems using machine learning, and assessed the model performance through appropriate evaluation metrics.