

Summary Report

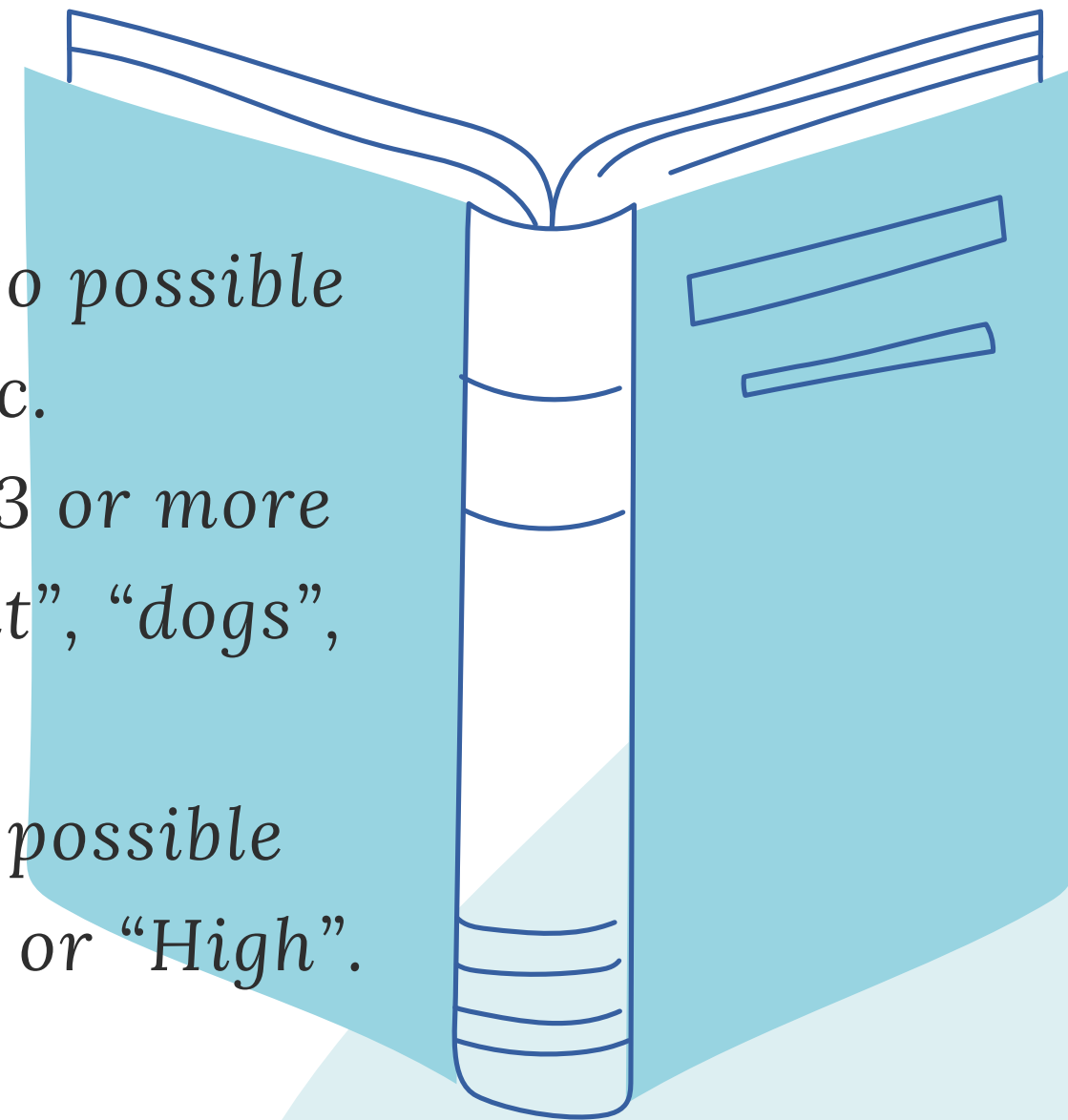
By Bassant_Tamer

Logistic regression :

is a supervised machine learning algorithm used for classification tasks where the goal is to predict the probability that an instance belongs to a given class or not.

Types of Logistic Regression :

1. **Binomial:** In binomial Logistic regression, there can be only two possible types of the dependent variables, such as 0 or 1, Pass or Fail, etc.
2. **Multinomial:** In multinomial Logistic regression, there can be 3 or more possible unordered types of the dependent variable, such as “cat”, “dogs”, or “sheep”
3. **Ordinal:** In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as “low”, “Medium”, or “High”.



Confusion Matrix

A table used to evaluate the performance of a classification model. It consists of:

- True Positives (TP): Correctly predicted positive cases.
- True Negatives (TN): Correctly predicted negative cases.
- False Positives (FP): Incorrectly predicted positive cases (Type I Error).
- False Negatives (FN): Incorrectly predicted negative cases (Type II Error).

Accuracy, Precision, Recall, F1 Score

- **Accuracy:**

Measures how many predictions were correct overall.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

- **Precision:**

Measures how many of the predicted positives were actually correct.

$$\text{Precision} = (\text{TP}) / (\text{TP} + \text{FP})$$

- **Recall (Sensitivity or True Positive Rate):**

Measures how many actual positives were correctly predicted.

$$\text{Recall} = (\text{TP}) / (\text{TP} + \text{FN})$$

- **F1 Score:**

Harmonic mean of precision and recall, providing a balance between them.

$$\text{F1 Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$



Performance Metrics for Classification

- If false positives are costly → Focus on Precision.
- If false negatives are costly → Focus on Recall.
- For a balanced evaluation → Use F1 Score.

AUC (Area Under the Curve):

Measures how well the model separates classes. The closer AUC is to 1, the better the model.