Logistic Regression - Student Learning Guide

1. What is Logistic Regression?

Logistic Regression is a supervised machine learning algorithm used for binary classification problems.

It estimates the probability that a data point belongs to a certain class, using the logistic (sigmoid) function.

2. The Sigmoid Function

The Sigmoid Function:

$$S(x) = 1 / (1 + e^{(-x)})$$

It maps any real-valued number into a value between 0 and 1, representing probability.

3. Why Use Logistic Regression?

Why Use Logistic Regression:

- It's fast and interpretable
- Works well with linearly separable data
- Output probabilities are easy to understand
- Great baseline model for classification tasks

4. Decision Boundary

Decision Boundary:

Logistic regression draws a straight line (in 2D) or a hyperplane (in higher dimensions) to separate the two classes based on the best fit to the data.

5. Loss Function

Loss Function:

We use log-loss (binary cross-entropy) to penalize wrong predictions:

$$Loss = -[y*\log(p) + (1-y)*\log(1-p)]$$

where y is the true label, and p is the predicted probability.

6. Common Use Cases

Logistic Regression - Student Learning Guide

Common Use Cases:

- Disease diagnosis (e.g., ASD detection)
- Spam detection
- Customer churn prediction
- Credit risk classification