Logistic Regression

- **Purpose**: Predict the probability of an input belonging to a certain **category** (e.g., "yes/no," "spam/not spam").
- **Type of Problem**: Used for **binary classification** (2 categories), but can be extended to multi-class.
- **Key Idea**: Instead of predicting a continuous value (like linear regression), it predicts a probability between **0 and 1**.

Real-Life Example

Imagine you want to predict whether a student will **pass (1) or fail (0)** an exam based on their study hours.

- **Linear Regression Problem**: Predicts marks (e.g., 75, 90), which can be >100 or <0.
- **Logistic Regression Problem**: Predicts **probability of passing**, e.g., 0.8 (80% chance of passing).

Sigmoid Function: The "S" Curve

Logistic Regression uses the **sigmoid function** to squeeze the output between 0 and 1.

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

Where:

- $z = b_0 + b_1 \cdot \text{StudyHours}$ (just like linear regression).
- If $\sigma(z) \geq 0.5$, predict class **1** (pass).
- If $\sigma(z) < 0.5$, predict class **0** (fail).

Visualization:

- The curve looks like an "S" (close to 0 for low study hours, close to 1 for high study hours).

How Does It Learn? (Intuition)

- **Goal**: Adjust weights (b_0, b_1) so that:
  - If the true label is **1**, predicted probability $\sigma(z)$ should be close to **1**.
  - If the true label is **0**, predicted probability $\sigma(z)$ should be close to **0**.
- **Loss Function (Cross-Entropy)**: Penalizes wrong predictions more if they're very confident (e.g., predicting 0.9 when the true label is 0).

---

5. Assumptions of Logistic Regression

1. The dependent variable is **binary** (e.g., yes/no).
2. No high multicollinearity among features.
3. Large sample size helps (small samples may lead to overfitting).