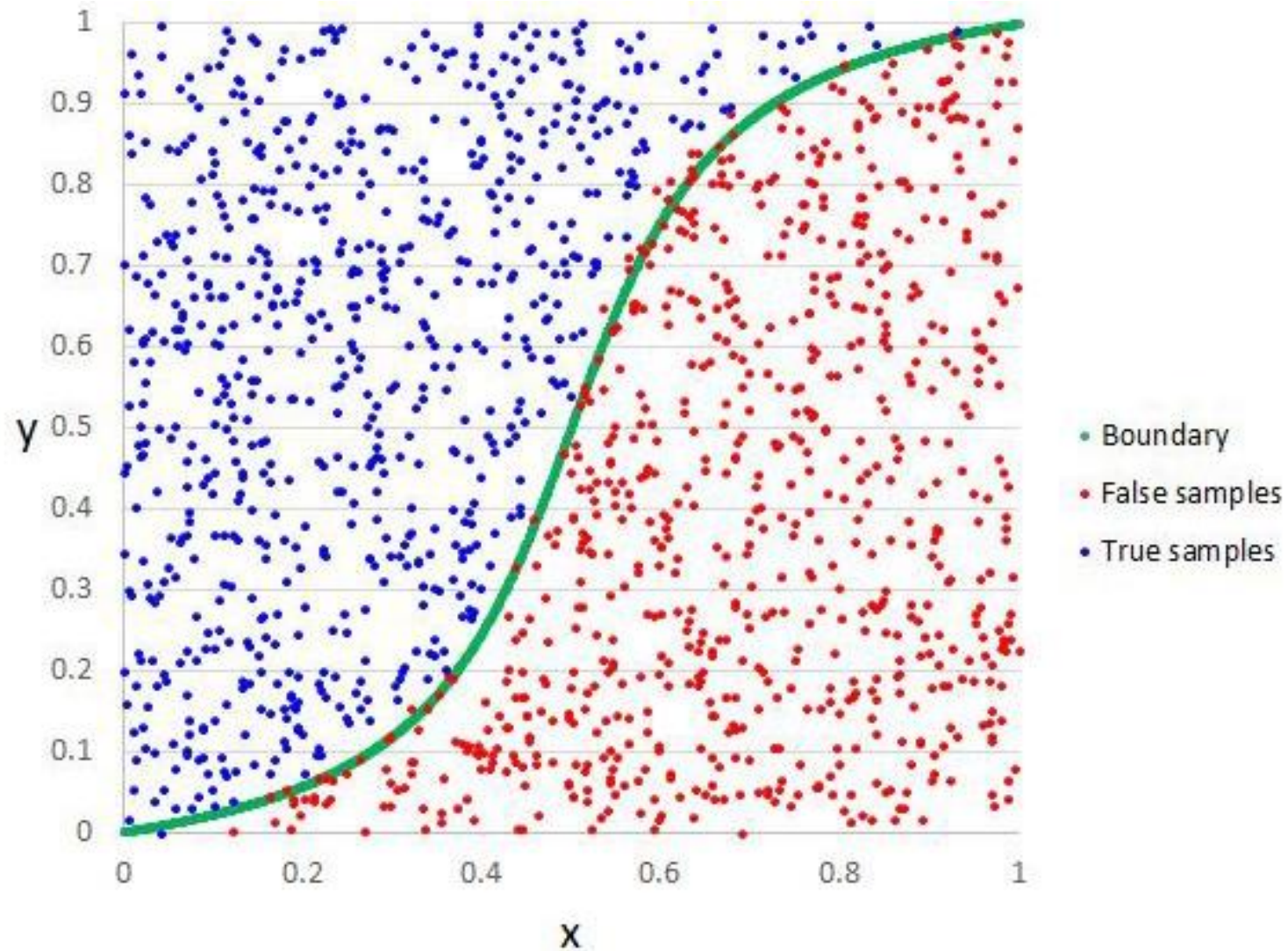


Logistic regression

Logistic regression

- Logistic Regression is used when the dependent variable(target) is categorical.
 - Email: Spam / Not Spam?
 - Online Transactions: Fraudulent (Yes / No)?
 - Tumor: Malignant / Benign ?
 - 0: “Negative Class” (e.g., benign tumor)
 - 1: “Positive Class” (e.g., malignant tumor)
- $$y \in \{0, 1\}$$
- The output from the hypothesis is the estimated probability.
 - This is used to infer how confident can predicted value be actual value when given an input X.

Logistic regression

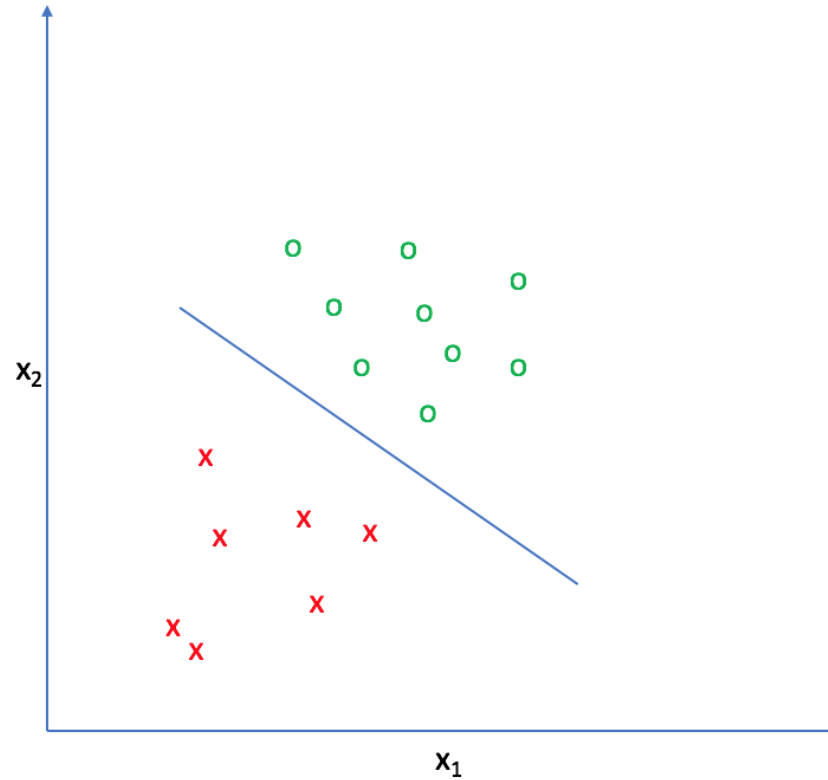


Types of Logistic Regression

- Binary Logistic Regression
 - The categorical response has only two 2 possible outcomes. Example: Spam or Not
- Multinomial Logistic Regression
 - Three or more categories. Example: Predicting which food is preferred more (Veg, Non-Veg, Vegan)

Decision Boundary

- To predict which class a data belongs, a threshold can be set.
- Based upon this threshold, the obtained estimated probability is classified into classes.

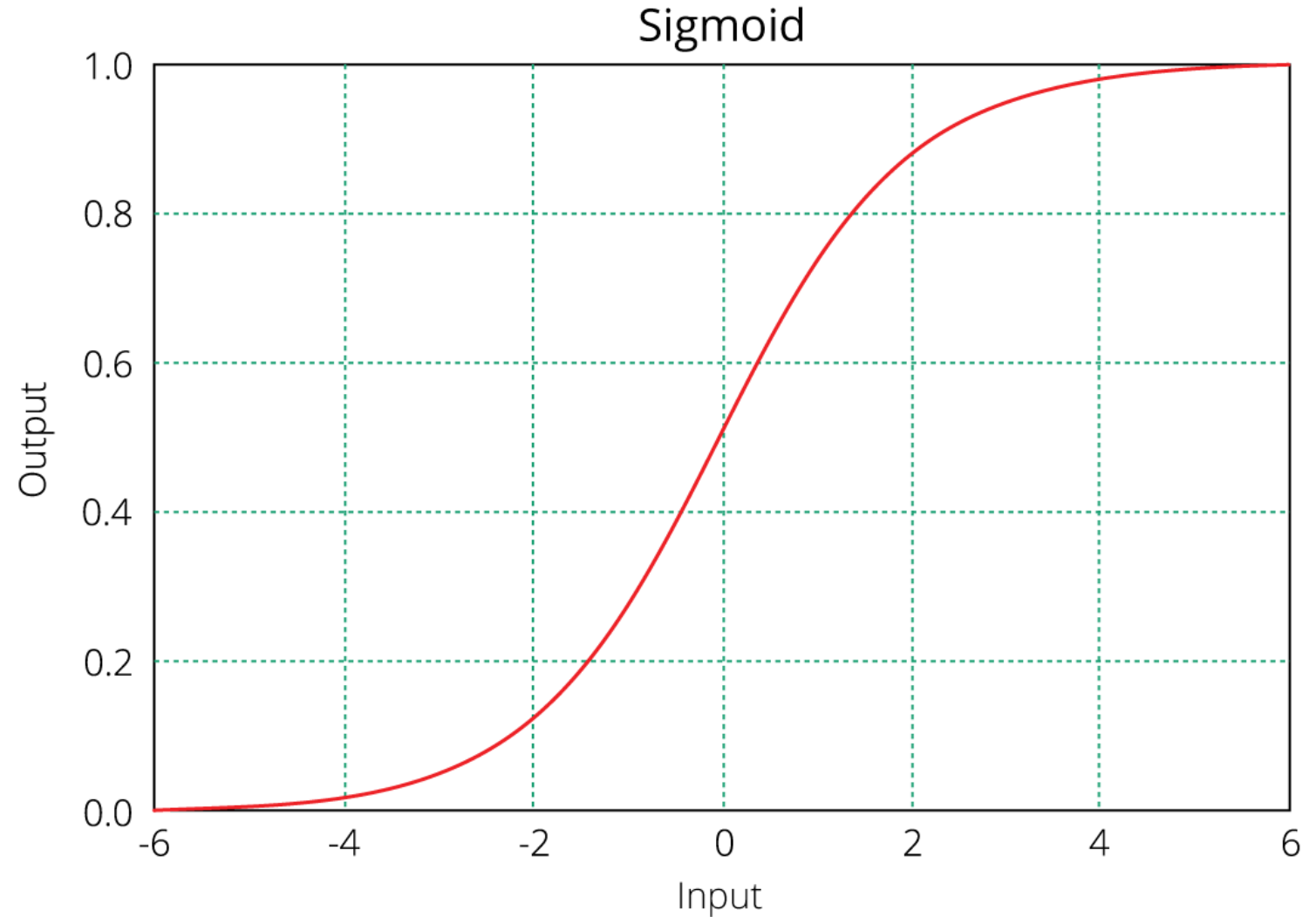


How does Logistic Regression work?

- Logistic Regression assumes that the dependent variable follows a binomial distribution:
 - Each trial can have only two outcomes; i.e., the response variable can have only two unique categories.
 - The probability of success (p) and failure (q) should be the same for each trial.
- For Linear Regression, we write the equation as:
 - $Y = \beta_0 + \beta_1 X + \epsilon$
- In Logistic Regression, we use the same equation but with some modifications made to Y.
- In Logistic Regression, we calculate probability of the class.

The Sigmoid Function

With the help of Sigmoid function, we transform the output from linear equation to the probability scale (0-1).



Confusion Matrix

- Confusion matrix is commonly used to evaluate classification models.
- It's quite confusing but make sure you understand it by heart.

	1 (Predicted)	0 (Predicted)
1 (Actual)	True Positive	False Negative
0 (Actual)	False Positive	True Negative

Thanks