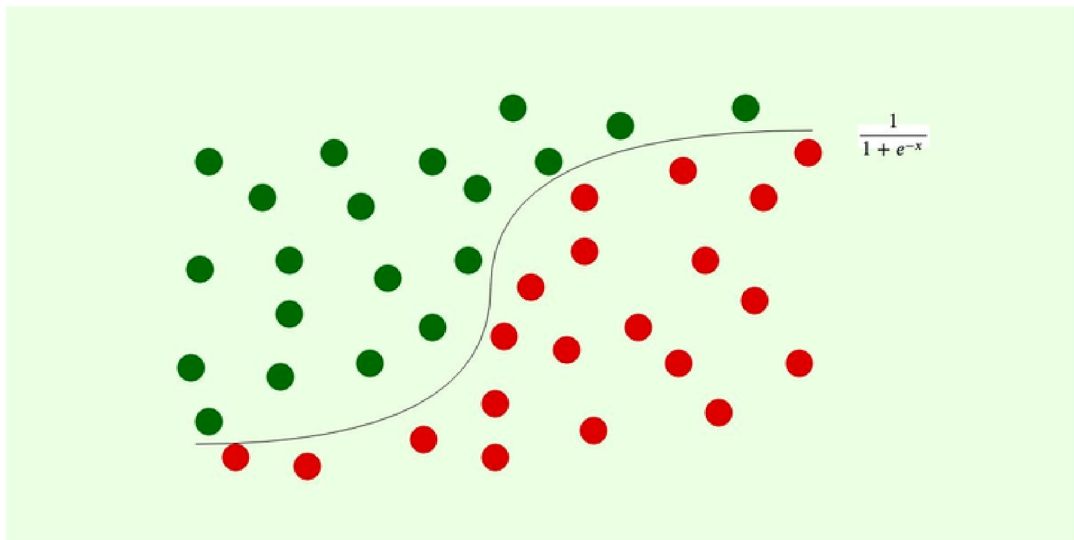


Day 12: Logistic Regression





What is Logistic Regression

- Logistic regression is the simplest classification algorithm. Though the name is regression, it is a classification algorithm.
- It uses sigmoid function to get the output between 0 and 1 as a probability.
- Based on the threshold selected the classification will be performed.
- It is a regression model that provides discrete output instead of continuous output. Thus helps in solving classification problem.
- It does not require linear relationship between input and output.



Types of Logistic Regression

- There are three main types of logistic regression, Binary, Multinomial, and Ordinal.
- Binary logistic regression is used to predict probability of a binary class.
- Multinomial logistic regression is used to predict the probability of three or more classes.
- Ordinal logistic regression is used to predict the probability of ordered class.



Sigmoid Function

- Sigmoid function maps any real value between 0 and 1. It is also called logistic function.
- It is defined as $1/(1+e^{(-x)})$, where x is the input value and e is the mathematical constant of 2.718.
- The graph of sigmoid function is S-shaped. It is a non-linear function.
- It is also used in neural networks as the activation function.
- It is an important function used in logistic regression.
- The linear regression method can't be used for classification problem because it uses linear function. Hence we need sigmoid function to introduce non-linearity.



Decision Boundary

- Decision boundary decides whether the data falls in class A or class B.
- It is decided based on the threshold chosen.
- For example, if probability ≥ 0.5 then the patient has heart disease.
- The threshold is chosen according to the goal of the problem.
- Example 1: In heart disease detection, we don't want any positive patient to be detected as negative. Hence in this case we choose low value as a threshold. This gives us more Positives.
- Example 2: In the prediction of probability whether a bank customer is interested in new credit card, we don't want non-interested customer to be predicted as interested, because this invests more energy in wrong customer. Hence in this case we choose high value as a threshold which will reduce unnecessary positives.