NAIVE BAYES

-MUKESH KUMAR

What is Naive Bayes?

- Naive Bayes is a supervised machine learning algorithm used for classification tasks.
- It is based on Bayes' Theorem, which calculates the probability of an event occurring given the probability of another event that has already occurred.
- Naive Bayes makes the assumption that features are independent of each other, which simplifies the calculations

Key Concepts:

- Bayes' Theorem:
- Naive Assumption:
- Maximum A Posteriori (MAP)

Bayes' Theorem

- Bayes' Theorem is a fundamental concept in probability theory and statistics that describes how to update the probability of a hypothesis based on new evidence.

- It provides a way to calculate the conditional probability of an event, given prior knowledge of conditions that might be related to the event.

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Where:

- $P(A|B)$ is the probability of event A occurring given B is true (posterior probability).
- $P(B|A)$ is the probability of event B given that A is true (likelihood).
- $P(A)$ is the prior probability of A (before considering evidence B).
- $P(B)$ is the probability of B.

HOW DOES NAIVE BAYES WORK?

- Naive Bayes uses Bayes' Theorem to calculate the probability of a data point belonging to each class.

- The class with the highest probability is assigned to the data point. The formula used is:

Steps to Use Naive Bayes:

- **Calculate Prior Probabilities**: The probability of each class occurring in the dataset.
- **Calculate Likelihood**: For each feature, calculate the likelihood of observing a particular feature value given the class.
- **Predict Class**: Use Bayes' theorem to compute the posterior probability for each class and choose the one with the highest probability.

Example

- Notebook : Naïve_bayes_example

Advantages of Naive Bayes

- Easy to implement and computationally efficient
- Effective with a large number of features
- Performs well with limited training data
- Handles categorical and numerical data well

Disadvantages of Naive Bayes

- Assumes features are independent, which is often not true in real-world data
- Can be influenced by irrelevant attributes
- May assign zero probability to unseen events, leading to poor generalization

Applications of Naive Bayes

- Spam email filtering
- Text classification (sentiment analysis, document categorization, topic classification)
- Medical diagnosis
- Credit scoring
- Weather prediction