Naïve Bayes algorithm theory

Varnika Mulay

Naïve Bayes is a probabilistic classifier in Machine Learning. It is based on the working of Bayes' theorem in statistics which is used for classifying data into different categories or classes. It is a simplistic algorithm that assumes that the features of a dataset are independent of each other. In a classification context, it calculates the posterior probability for each class given the input features and assigns the class with the highest posterior probability. It works well for classifying tasks like sentiment analysis, text classification, and different classification problems.

Bayes' Theorem

Bayes theorem helps to determine the probability of an event with random knowledge. It calculates the probability of occurring one event while the other one has already occurred.

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

Where,

- P(A|B) is the probability of the hypothesis "A being true given evidence B" which is also called the **posterior probability** of the hypothesis.
- P(B|A) is the probability of evidence B given that hypothesis A is true. This is also called the **likelihood**.
- P(A) is the probability of hypothesis A being true which is also called as **prior probability**.
- P(B) is the total probability of evidence B which is also called as **marginal probability**.

$$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

$$P(target|predictor) = \frac{P(predictor|target) \cdot P(target)}{P(predictor)}$$