

Naïve Bayes

It is a probabilistic machine learning algorithm based on Bayes' Theorem. It assumes independence among predictors (features). In simple terms, the presence of a particular feature in a class is unrelated to the presence of any other feature. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

Bayes Theorem

Shows the relation b/w one conditional probability and its inverse.

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

$P(A | B)$ is referred to as posterior which means the probability of occurrence of A given B.

$P(B | A)$ is referred to as likelihood ratio which measures the probability of given event A of occurrence of B.

$P(A)$ is referred to as prior which represents the actual probability distribution of A.

$P(B)$ is Marginal Probability -- probability of Evidence.

Types of Naïve Bayes

The **Gaussian** model assumes that features follow a normal distribution. This means if predictors take continuous values instead of discrete, then the model assumes that these values are sampled from the Gaussian distribution.

The **Multinomial** Naïve Bayes classifier is used when the data is multinomial distributed. It is primarily used for document classification problems, it means a particular document belongs to which category such as Sports, Politics, education, etc.

The classifier uses the frequency of words for the predictors.

The **Bernoulli** classifier works similar to the Multinomial classifier, but the predictor variables are the independent Booleans variables. Such as if a particular word is present or not in a document. This model is also famous for document classification tasks.