

Big Data Analytics



NITHISH KUMAR G.
21BCS036.
BSC COMPUTER SCIENCE.

AGENDA

- ❑ What is Bayes Theorem ?
- ❑ What is NAIVE BAYES CLASSIFIER ?
- ❑ Types of Naive Bayes Algorithm.

BAYES THEOREM

BAYES THEOREM

1. In probability theory and statistics, Bayes' theorem (alternatively Bayes' law or Bayes' rule) describes the probability of an event, based on prior knowledge of conditions that might be related to the event.
2. For example, if cancer is related to age, then, using Bayes' theorem, a person's age can be used to more accurately to assess the probability that they have cancer, compared to the assessment of the probability of cancer made without knowledge of the person's age.

Formula for Baye's Theorem :

$P(H / E) = P(E/H) * P(H) / P(E)$ where ,

$P(H)$ is the probability of hypothesis H being true. This is known as the prior probability.

$P(E)$ is the probability of the evidence (regardless of the hypothesis).

$P(E / H)$ is the probability of the evidence given that hypothesis is true.

$P(H / E)$ is the probability of the hypothesis given that the evidence is there.

NAIVE BAYES CLASSIFIER

- Naive Bayes is a kind of classifier which uses the Bayes Theorem.
- It predicts membership probabilities for each class such as the probability that given record or data point belongs to a particular class.
- The class with the highest probability is considered as the most likely class. This is also known as Maximum A Posteriori (MAP).
- $\text{MAP}(H)$
- $= \max(P(H|E))$
- $= \max((P(E|H)*P(H))/P(E))$
- $= \max(P(E|H)*P(H))$ $P(E)$ is evidence probability, and it is used to normalize the result. It remains same so, removing it won't affect.

TYPES OF NAVIE BAYES ALGORITHM

- ❖ **Gaussian:** It is used in classification and it assumes that features follow a normal distribution.
- ❖ **Bernoulli:** The binomial model is useful if your feature vectors are binary (i.e. zeros and ones). One application would be text classification with 'bag of words' model where the 1s & 0s are "word occurs in the document" and "word does not occur in the document" respectively.
- ❖ **Multinomial** Naive Bayes is preferred to use on data that is multinomially distributed.
- ❖ **For example,** let's say, we have a text classification problem. Here we can consider bernoulli trials which is one step further and instead of "word occurring in the document", we have "count how often word occurs in the document", you can think of it as "number of times outcome number x_i is observed over the n trials".