Naïve Bayes Classifier

Naïve Bayes Classifier Algorithm

•Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.

•It is mainly used in text classification that includes a high-dimensional training dataset.

•Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.

•It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

•Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis and classifying articles

•The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:

•Naïve: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other.

•Bayes: It is called Bayes because it depends on the principle of Bayes' Theorem.

Bayes' Theorem:

•Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.

•The formula for Bayes' theorem is given as:

|  |  |
| --- | --- |
| •Naïve Bayes Classifier Algorithm |  |

Where,

•P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

•P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

•P(A) is Prior Probability: Probability of hypothesis before observing the evidence.

•P(B) is Marginal Probability: Probability of Evidence.

STEPS

• find the P(A)- prior probability

•Convert the given dataset into frequency tables.

•Generate Likelihood table by finding the probabilities of given features. (Conditional probabilities)

•Now, use Bayes theorem to calculate the posterior probability.

Advantages of Naïve Bayes Classifier  
• Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.• It can be used for Binary as well as Multi-class Classifications.

• It performs well in Multi-class predictions as compared to the other Algorithms.• It is the most popular choice for text classification problems.

•Disadvantages of Naïve Bayes Classifier:  
• Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

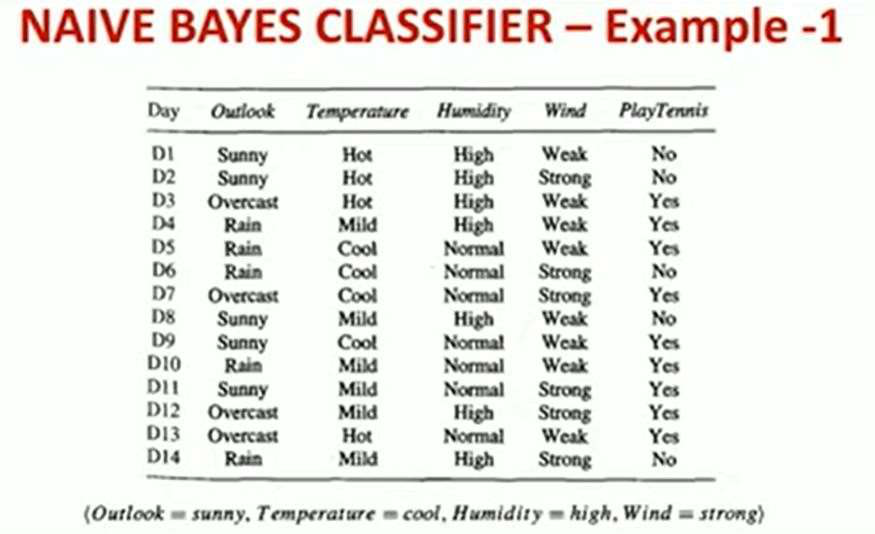
Applications of Naïve Bayes Classifier:

•It is used for Credit Scoring.

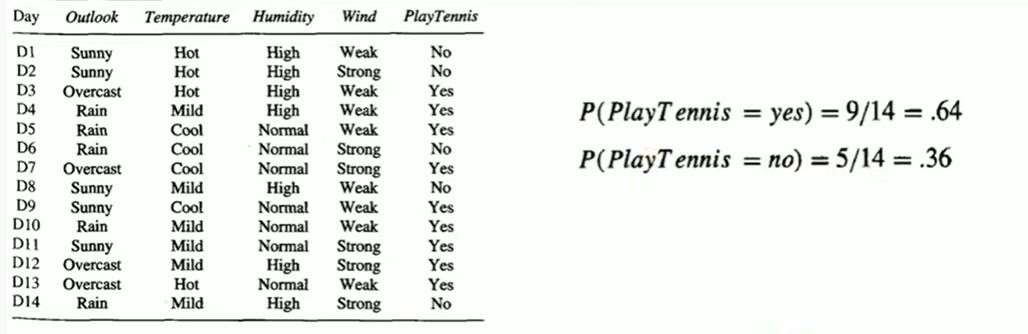
•It is used in medical data classification.

•It can be used in real-time predictions because Naïve Bayes Classifier is an eager learner.

•It is used in Text classification such as Spam filtering and Sentiment analysis.



Calculate prior probability(P)-> Y & N



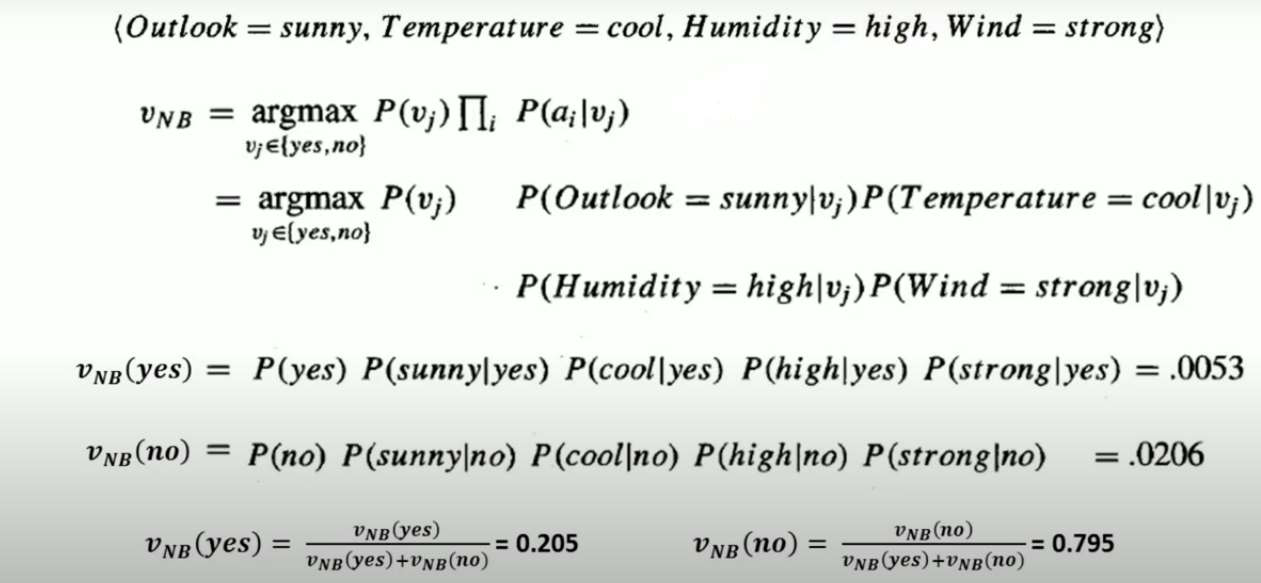
Calculate the conditional probability->each attribute

|  |  |
| --- | --- |
|  |  |

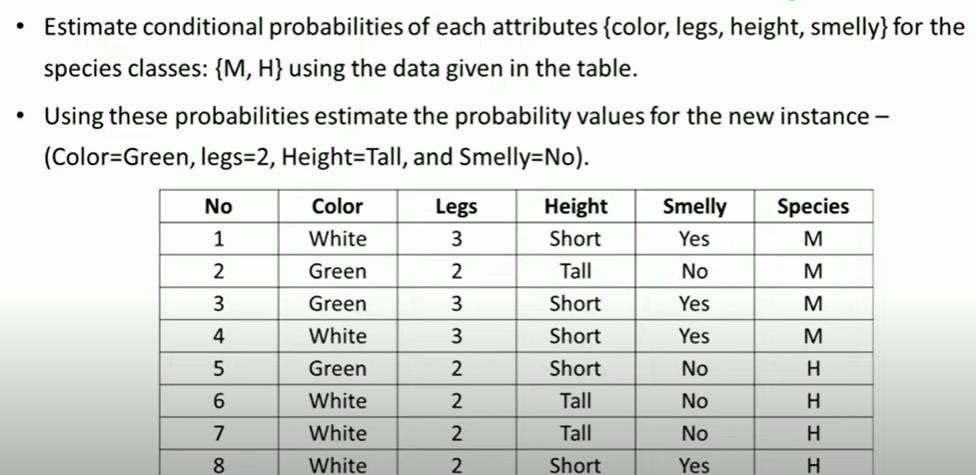
Calculate the conditional probability->each attribute

|  |  |
| --- | --- |
|  |  |

Classify the new example ->Y or N



Example 2



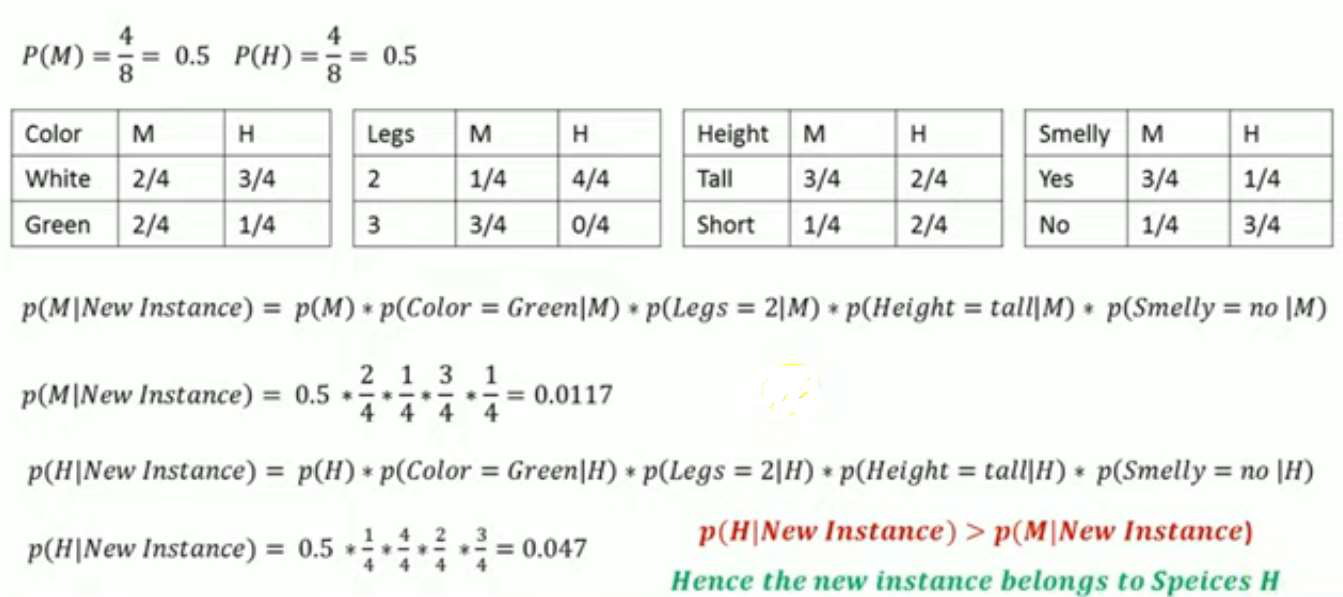
Calculate Prior probability (M) & P(H)

|  |  |
| --- | --- |
|  |  |

Calculate for all attributes(conditional probabilities) ->possible outcome, possible values

|  |  |
| --- | --- |
|  |  |

Calculate the probability of new instance->classify whether M or H( Using prior & Conditional)



**Text Classification**

• **Text Classification** is an example of supervised **machine learning** task since a labelled dataset containing **text** documents and their labels is used for train a **classifier.**

• The **Naive Bayes classifier** is a simple **classifier** that classifies based on probabilities of events. It is the applied commonly to **text classification**. **With** the training set, we can train a **Naive Bayes classifier** which we can **use** to automatically categorize a new sentence.

• TYPES

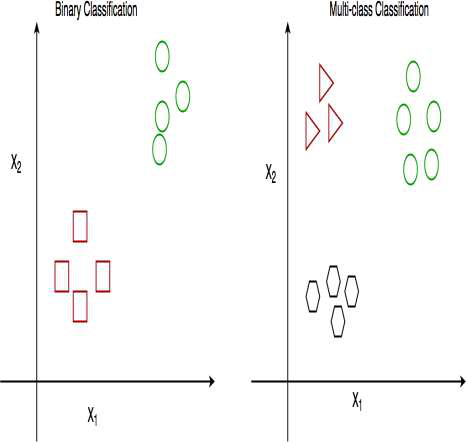
|  |  |  |
| --- | --- | --- |
| • **Binary Classification** : When we have to categorize given data into 2 | | distinct classes. |
| Example – On the basis of given health conditions of a | person, we have to determine | |

whether the person has a certain disease or not.

• **Multiclass Classification** : The number of classes is more than 2. For Example – On the

basis of data about different species of flowers, we have to determine which specie does our

observation belong to.

Example Fig : Binary and Multiclass Classification. Here x1 and x2 are our variables upon which the class is predicted.   
How does classification works?

• Which means there are two possible outcomes:

• Suppose we have to predict whether a given patient has a certain disease or not, on the basis of 3 variables, called features.

• The patient has the said disease. Basically a result labelled “Yes” or “True”.

• The patient is disease free. A result labelled “No” or “False”.

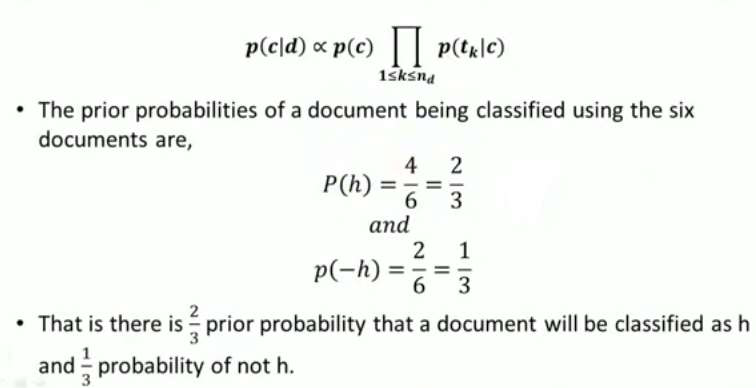
• This is a binary classification problem.

• We have a set of observations called training data set, which comprises of sample data with actual classification results. We train a model, called Classifier on this data set, and use that model to predict whether a certain patient will have the disease or not.

EXAMPLE 3



Priori Probability



Conditional Probability

|  |  |
| --- | --- |
|  |  |

