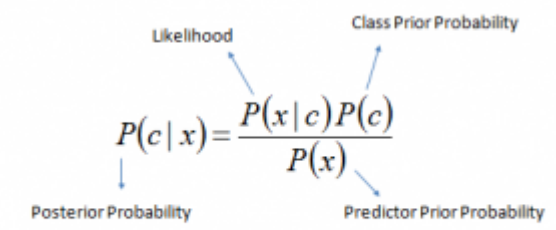# What is Naive Bayes algorithm?

It is a classification technique based on <u>Bayes' Theorem</u> with an assumption of independence among predictor simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelate the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, about 3 inches in diameter. Even if these features depend on each other or upon the existence of the features, all of these properties independently contribute to the probability that this fruit is an apple and that is it is known as 'Naive'.

Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, N Bayes is known to outperform even highly sophisticated classification methods.

Bayes theorem provides a way of calculating posterior probability P(c|x) from P(c), P(x) and P(x|c). Look a equation below:

$$P(c\,|\,x) = \frac{P(x\,|\,c)\,P(c)}{P(x)}$$

- Likelihood
- Class Prior Probability
- Posterior Probability
- Predictor Prior Probability

$$P(c\,|\,X) = P(x_1\,|\,c) \times P(x_2\,|\,c) \times \cdots \times P(x_n\,|\,c) \times P(c)$$

Above,

- *P(c|x)* is the posterior probability of *class* (c, *target*) given *predictor*(x, *attributes*).
- *P(c)* is the prior probability of *class*.
- *P(x|c)* is the likelihood which is the probability of *predictor* given *class*.
- *P(x)* is the prior probability of *predictor*.

How Naive Bayes algorithm works?

Let's understand it using an example. Below I have a training data set of weather and corresponding target var 'Play' (suggesting possibilities of playing). Now, we need to classify whether players will play or not base weather condition. Let's follow the below steps to perform it.

Step 1: Convert the data set into a frequency table

Step 2: Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of pla is 0.64.

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Frequency Table

Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table

Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with t highest posterior probability is the outcome of prediction.

Problem: Players will play if weather is sunny. Is this statement is correct?

We can solve it using above discussed method of posterior probability.

$P(Yes | Sunny) = P(Sunny | Yes) * P(Yes) / P (Sunny)$

Here we have P (Sunny |Yes) = 3/9 = 0.33, P(Sunny) = 5/14 = 0.36, P(Yes)= 9/14 = 0.64

Now, P (Yes | Sunny) = 0.33 * 0.64 / 0.36 = 0.60, which has higher probability.

Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification and with problems having multiple classes.

What are the Pros and Cons of Naive Bayes?

Pros:

- It is easy and fast to predict class of test data set. It also perform well in multi class prediction
- When assumption of independence holds, a Naive Bayes classifier performs better compar other models like logistic regression and you need less training data.
- It perform well in case of categorical input variables compared to numerical variable(s). For nume variable, normal distribution is assumed (bell curve, which is a strong assumption).

Cons:

- If categorical variable has a category (in test data set), which was not observed in training data then model will assign a 0 (zero) probability and will be unable to make a prediction. This is often know "Zero Frequency". To solve this, we can use the smoothing technique. One of the simplest smoo techniques is called Laplace estimation.
- On the other side naive Bayes is also known as a bad estimator, so the probability ou from `predict_proba` are not to be taken too seriously.

- Another limitation of Naive Bayes is the assumption of independent predictors. In real life, it is al impossible that we get a set of predictors which are completely independent.

4 Applications of Naive Bayes Algorithms

- **Real time Prediction:** Naive Bayes is an eager learning classifier and it is sure fast. Thus, it could be for making predictions in real time.
- **Multi class Prediction:** This algorithm is also well known for multi class prediction feature. Here we predict the probability of multiple classes of target variable.
- **Text classification/ Spam Filtering/ Sentiment Analysis:** Naive Bayes classifiers mostly used in classification (due to better result in multi class problems and independence rule) have higher success as compared to other algorithms. As a result, it is widely used in Spam filtering (identify spam e-mail) Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments)
- **Recommendation System:** Naive Bayes Classifier and <u>Collaborative Filtering</u> together build Recommendation System that uses machine learning and data mining techniques to filter unseen informa and predict whether a user would like a given resource or not