# Naive Bayes and its applications

Name : Hazem Mamdouh Ali
Sec : 2

## What is Naive Bayes ?

Naive Bayes is a simple but surprisingly powerful probabilistic machine learning algorithm used for predictive modeling and classification tasks. Some typical applications of Naive Bayes are spam filtering, sentiment prediction, classification of documents, etc. It is a popular algorithm mainly because it can be easily written in code and predictions can be made real quick which in turn increases the scalability of the solution. The Naive Bayes algorithm is traditionally considered the algorithm of choice for practical-based applications mostly in cases where instantaneous responses are required for user requests.

# What is Naive Bayes theorem ?

Bayes' Theorem helps you examine the probability of an event based on the prior knowledge of any event that has correspondence to the former event. Its uses are mainly found in probability theory and statistics. The term naive is used in the sense that the features given to the model are not dependent on each other. In simple terms, if you change the value of one feature in the algorithm, it will not directly influence or change the value of the other features.

Consider, for example the probability that the price of a house is high can be calculated better if we have some prior information, like the facilities around it, compared to another assessment made without the knowledge of the location of the house.

# Equation used in Naive Bayes theorem ?

$$P(A|B) = [P(B|A)P(A)]/[P(B)]$$

The equation above shows the basic representation of Bayes' theorem where A and B are two events and:

**P(A|B):** The conditional probability that event A occurs, given that B has occurred. This is termed as the posterior probability.

**P(A) and P(B):** The probability of A and B without any correspondence with each other.

**P(B|A):** The conditional probability of the occurrence of event B, given that A has occurred.

# What is a Naive Bayes Classifier in Machine Learning?

A classifier is a machine learning model which is used to classify different objects based on certain behavior. Naive Bayes classifiers in machine learning are a family of simple probabilistic machine learning models that are based on Bayes' Theorem. In simple words, it is a classification technique with an assumption of independence among predictors.

The Naive Bayes classifier reduces the complexity of the Bayesian classifier by making an assumption of conditional dependence over the training dataset.

Consider you are given variables X, Y, and Z. X will be conditionally independent of Y given Z if and only if the probability distribution of X is independent of the value of Y given Z. This is the assumption of conditional dependence.

In other words, you can also say that X and Y are conditionally independent given Z if and only if, the knowledge of the occurrence of X provides no information on the likelihood of the occurrence of Y and vice versa, given that Z occurs. This assumption is the reason behind the term naive in Naive Bayes.

**The likelihood can be written considering n different attributes as:**

$$P(X_1 \ldots X_n | Y) = \prod_{i=1}^{n} P(X_i | Y)$$

In the mathematical expression, X represents the attributes, Y represents the response variable. So, P(X|Y) becomes equal to the product of the probability distribution of each attribute given Y.

## Maximizing a Posteriori:

If you want to find the posterior probability of P(Y|X) for multiple values of Y, you need to calculate the expression for all the different values of Y.

Let us assume a new instance variable X_NEW. You need to calculate the probability that Y will take any value given the observed attributes of X_NEW and given the distributions P(Y) and P(X|Y) which are estimated from the training dataset.

In order to predict the response variable depending on the different values obtained for P(Y|X), you need to consider a probable value or the maximum of the values. Hence, this method is known as maximizing a posteriori.

## Maximizing Likelihood

You can simplify the Naive Bayes algorithm if you assume that the response variable is uniformly distributed which means that it is equally likely to get any response. The advantage of this assumption is that the a priori or the P(Y) becomes a constant value.

Since the a priori and the evidence become independent from the response variable, they can be removed from the equation. So, maximizing the posteriori becomes maximizing the likelihood problem. You can solve similar machine learning problems and apply Bayes theorem in data science with python.

## What are the types of Naive Bayes classifier?

**Multinomial Naive Bayes** — These types of classifiers are usually used for the problems of document classification.  It checks whether the document belongs to a particular category like sports or technology or political etc and then classifies them accordingly. The predictors used for classification in this technique are the frequency of words present in the document.

**Complement Naive Bayes** — This is basically an adaptation of the multinomial naive bayes that is particularly suited for imbalanced datasets.

**Bernoulli Naive Bayes** — This classifier is also analogous to multinomial naive bayes but instead of words, the predictors are Boolean values. The parameters used to predict the class variable accepts only yes or no values, for example, if a word occurs in the text or not.

**Out-of-Core Naive Bayes** — This classifier is used to handle cases of large scale classification problems for which the complete training dataset might not fit in the memory.

**Gaussian Naive Bayes** — In a Gaussian Naive Bayes, the predictors take a continuous value assuming that it has been sampled from a Gaussian Distribution. It is also called a Normal Distribution.

# What are the advantages of the Naive Bayes?

## Pros of Naive Bayes —

- It is easy and fast to predict the class of the training data set.
- It performs well in multiclass prediction.
- It performs better as compared to other models like logistic regression while assuming the independent variables.
- It requires less training data.
- It performs better in the case of categorical input variables as compared to numerical variables.

# What are the disadvantages of the Naive Bayes?

## Cons of Naive Bayes —

- The model is not able to make a prediction in situations where the categorical variable has a category that was not observed in the training data set and assigns a 0 (zero) probability to it. This is known as the ‘Zero Frequency’. You can solve this using the Laplace estimation.
- Since Naive Bayes is considered to be a bad estimator, the probability outputs are not taken seriously.
- Naive Bayes works on the principle of assumption of independent predictors, but it is practically impossible to get a set of predictors that are completely independent.

## What are the applications of the Naive Bayes?

- **Real-time prediction** — It is a fast and eager machine learning classifier, so it is used for making predictions in real-time.
- **Multi-class prediction** — It can predict the probability of multiple classes of the target variable.
- **Classification/ Spam Filtering / Sentiment Analysis** — They are mostly used in text classification problems because of its multi-class problems and the independence rule. They are used for identifying spam emails and also to identify negative and positive customer sentiments on social platforms.

- **Recommendation Systems** — A Recommendation system is built by combining Naive Bayes classifier and Collaborating Filtering. It filters unseen information and predicts whether the user would like a given resource or not using machine learning and data mining techniques.

# Thanks