**Logistic Regression**

Logistic regression is a [supervised machine learning](https://www.geeksforgeeks.org/supervised-unsupervised-learning/) algorithm mainly used for [classification](https://www.geeksforgeeks.org/getting-started-with-classification/) tasks where the goal is to predict the probability that an instance of belonging to a given class. It is used for classification algorithms its name is logistic regression. it’s referred to as regression because it takes the output of the [linear regression](https://www.geeksforgeeks.org/ml-linear-regression/)function as input and uses a sigmoid function to estimate the probability for the given class. The [difference between linear regression and logistic regression](https://www.geeksforgeeks.org/ml-linear-regression-vs-logistic-regression/) is that linear regression output is the continuous value that can be anything while logistic regression predicts the probability that an instance belongs to a given class or not.

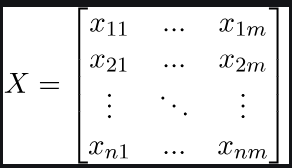
**Logistic Regression:**

It is used for predicting the categorical dependent variable using a given set of independent variables.

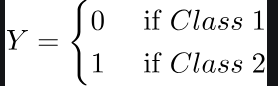
* Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value.
* It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.
* Logistic Regression is much similar to the Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems.
* In Logistic regression, instead of fitting a regression line, we fit an “S” shaped logistic function, which predicts two maximum values (0 or 1).
* The curve from the logistic function indicates the likelihood of something such as whether the cells are cancerous or not, a mouse is obese or not based on its weight, etc.
* Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.
* Logistic Regression can be used to classify the observations using different types of data and can easily determine the most effective variables used for the classification.

**Logistic Function (Sigmoid Function):**

* The sigmoid function is a mathematical function used to map the predicted values to probabilities.
* It maps any real value into another value within a range of 0 and 1. o The value of the logistic regression must be between 0 and 1, which cannot go beyond this limit, so it forms a curve like the “S” form.
* The S-form curve is called the Sigmoid function or the logistic function.
* In logistic regression, we use the concept of the threshold value, which defines the probability of either 0 or 1. Such as values above the threshold value tends to 1, and a value below the threshold values tends to 0.
* The logistic regression model transforms the [linear regression](https://www.geeksforgeeks.org/ml-linear-regression/) function continuous value output into categorical value output using a sigmoid function, which maps any real-valued set of independent variables input into a value between 0 and 1. This function is known as the logistic function.
* Let the independent input features be



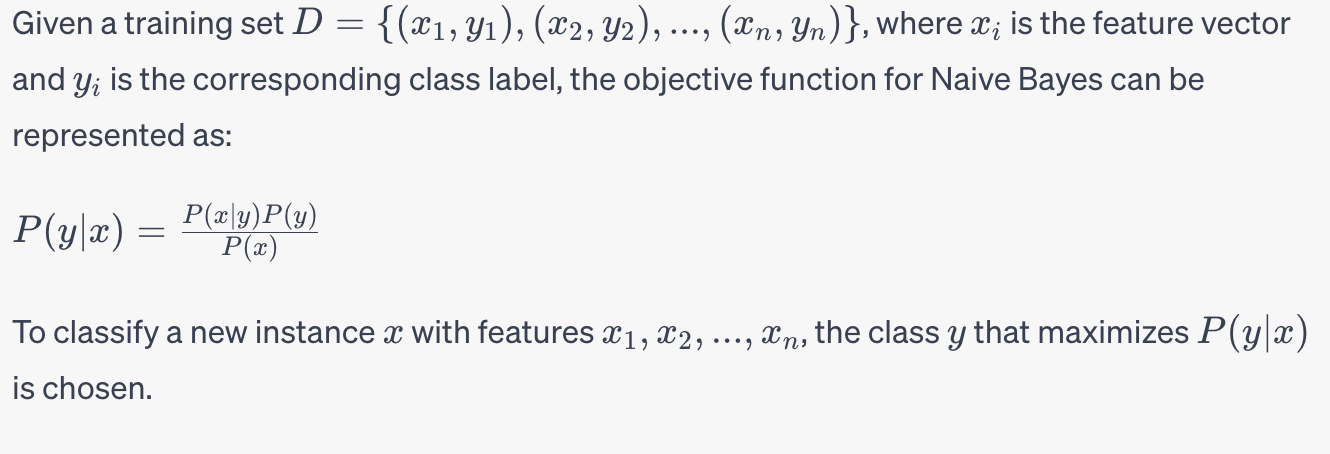
* The dependent variable is Y having only binary value i.e. 0 or 1.



* Then apply the multi-linear function to the input variables X
* Here  is the ith observation of X,  is the weights or Coefficient, and is the bias term also known as intercept. simply this can be represented as the dot product of weight and bias.
* Now we use the [sigmoid function](https://www.geeksforgeeks.org/derivative-of-the-sigmoid-function/) where the input will be z and we find the probability between 0 and 1. i.e predicted y.



**Naive Bayes classifier**

The Naive Bayes classifier can be understood in the context of an objective function. The objective function in this case is typically related to maximizing the posterior probability of the class given the features. This involves maximizing the likelihood of the training data under the assumption that the features are conditionally independent given the class. The objective function can be formulated as follows:

* ***Prior probability (P(y)):*** This represents the probability of each class occurring in the dataset. It can be estimated as the frequency of each class in the training set.
* ***Likelihood (P(x | y)):*** This represents the probability of the features given the class. In the case of the Naive Bayes assumption, it is assumed that the features are conditionally independent given the class. Therefore, the likelihood is calculated as the product of the individual feature probabilities.
* ***Evidence (P(x)):*** This is the probability of the features occurring, which acts as a normalization factor. In practice, this term is often ignored since it doesn't affect the comparison of different class probabilities.

The objective function involves the following components: To find the optimal parameters, the algorithm usually maximizes the likelihood of the training data. This can involve finding the maximum likelihood estimation or maximum a posteriori estimation, depending on the incorporation of prior knowledge. Understanding the objective function helps in grasping the underlying mathematical principles that drive the Naive Bayes classifier and its decision-making process.