**Machine Learning Assignment-01**

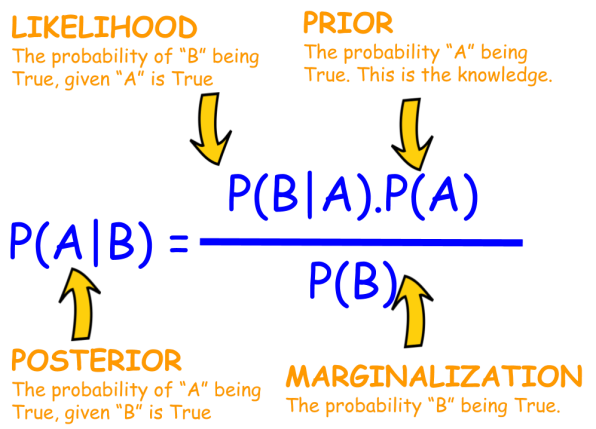
**Group No.36**

CLASSIFICATION OF CLASSES OF THE FLOWER SPECIES BASED ON THEIR FEATURES:-

Bayes Theorem**:**

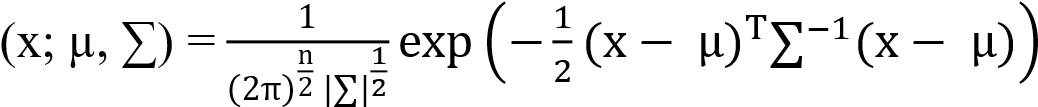
Bayes' theorem describes the probability of an event, based on prior knowledge of conditions that might be related to the event. It gives us the conditional probability.

Given a Hypothesis (**A)** and evidence (**B)**, Bayes Theorem tells that the relationship between the probability of the hypothesis before the evidence, **P(A)**,and the probability of the hypothesis after getting the evidence, **P(A|B)**, is :



**Bayes Classifier:**

A **Bayesian classifier** is based on the idea that the role of a class is to predict the values of features for members of that class.

To Calculate Likelihood we used Multivariate Gaussian Distribution and the formula is: p

Where ∑ means nothing but covariance matrix

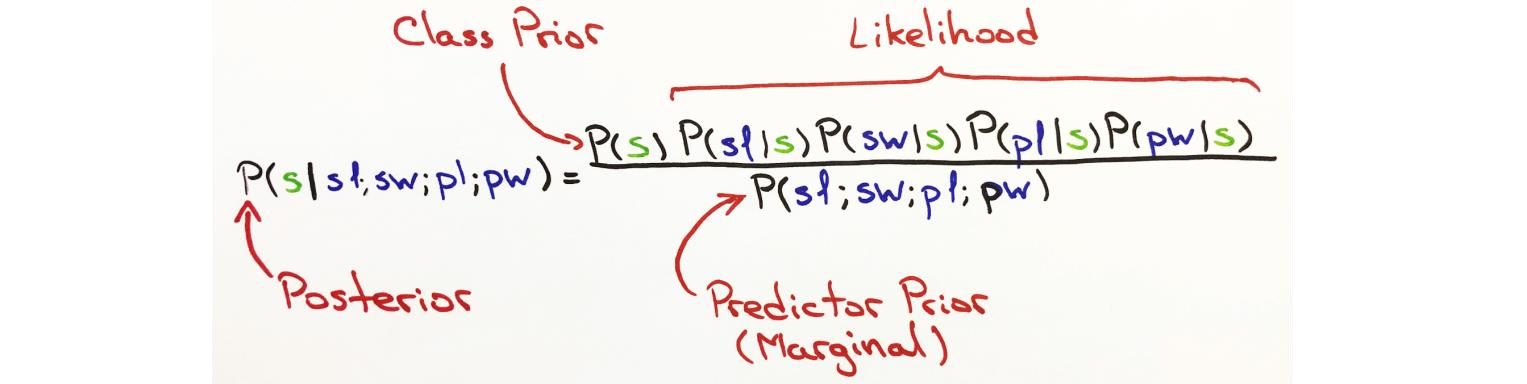
Form this we calculate the posterior probabilities and predict the class label that is nothing but the species of the flower it belongs to

**Naive Bayes Classifier:**

Naive Bayes is simple and the powerful algorithms for **classification** based on Bayes.

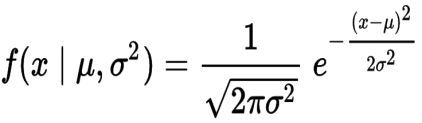
This model makes some strong assumptions. All of the features are assumed to be independent when calculating the likelihood; hence "Naive". Likelihood is calculated using the Gaussian Distribution (Normal Distribution) and all of the features are assumed to be normally distribtuted; hence "Gauss".

Overlooking Gauss NB's **strong assumptions**, the classifier is very fast and accurate. Gauss NB does not require a lot of data to be accurate and is highly scalable.



The Normal Distribution will help determine the Normal Probability for each new feature given the class. The product of all Normal Probabilities will result in the likelihood of the classoccurring given the new features. In other words, the Normal Distribution will calculate the Normal Probability value for each new feature. The product of all Normal Probabilities will be the likelihood.

**Normal pdf formula:**



**Implementation:**

**No inbuilt libraries such as scikit-learn or any other is used . we constructed every function and implemented both bayes and naïve bayes from scratch**

We implemented algorithm by broking them into 6 parts:

* Step 1: grouping the data set based on Class.
* Step 2: Summarize of this Dataset.

Step 3 Gaussian Probability Density Function that is nothing but calculating the likelihood

* Step 4 posterior probabilities
* Step5: checking the predicting labels with actual data
* Step6: accuracy

**Difference:**

The difference between Bayes and Naive Bayes Classifier is that the features are independent in naive bayes that is nothing but Cov(a,b) = 0 where a,b are the different features.

**Observations From Bayes Classifier:**

For this iris data set we got the accuracy 100 percent .but if there are many features, it is unlikely to find records that match exactly the same features.

**Observations From Naive Bayes Classifier:**

For this iris data set we got the accuracy 100 percent .but Naive Bayes classifiers compute the probabilities of each feature occurring in a class independently. So it could classify easily for larger data

**Result:**

1)Naive Bayes classifier is easy to implement than that of bayesclassifier

2)The Naive Bayes Classifier works better for larger data where as bayes classifier is used for smaller data.

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*The End\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***