**GDA**

In this algorithm, we analyse the samples and assume that they are normally distributed or in gaussian distribution. We learn P(x|y) first and then during prediction use Baiyes theorem to find P(y|x).

In our case we had to classes, so we will initially learn

and

Where,

‘x’ is a sample row vector with ‘n’ features.

‘µi’ is the mean of the class ‘i’.

‘’ is the covariance matrix.

‘’ is the probability of an event.

Joint Likelihood here is the product of all probabilities of individual samples, i.e,

Maximizing this likelihood, we get,

Basically, it is the average of the samples belonging to the class ‘i’.

‘’ can also be chosen as a hyperparameter. Tuning it will affect the performance of the algorithm.

‘m’ is the number of training samples.

(x(i), y(i)) is the i th training sample

During prediction we will return