**Roll No: 120A2036**

**Batch:A2**

**Date of performance: 29/3/23**

**EXPERIMENT NO. 8**

**TO PERFORM BASIC BINARY CLASSIFICATION OF ANY DATA OR PATTERN USING NAÏVE BAYES CLASSIFIER.**

**EXPERIMENT NO. 8: Classification**

**AIM: -** To perform basic binary classification of any data or pattern using Naive Bayes Classifier.

**OBJECTIVES:**

1. To use and understand dataset.
2. To extract classify data.
3. To calculate accuracy of the classifier.

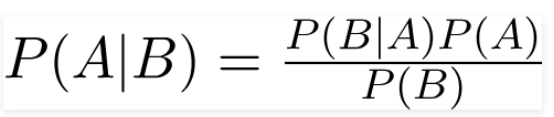
**EQUIPMENTS/SOFTWARE:** Python

**THEORY: -**

A classifier is a machine learning model that is used to discriminate different objects based on certain features.

A Naive Bayes classifier is a probabilistic machine learning model that’s used for classification task. The crux of the classifier is based on the Bayes theorem.

Bayes Theorem:



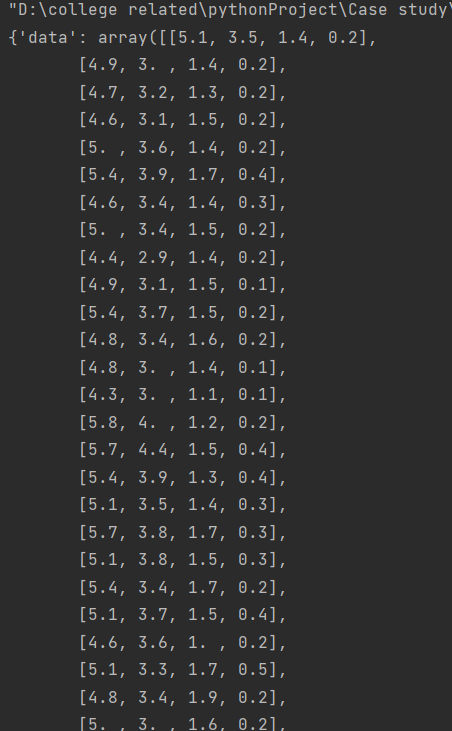
Using Bayes theorem, we can find the probability of A happening, given that B has occurred. Here, B is the evidence and A is the hypothesis. The assumption made here is that the predictors/features are independent.

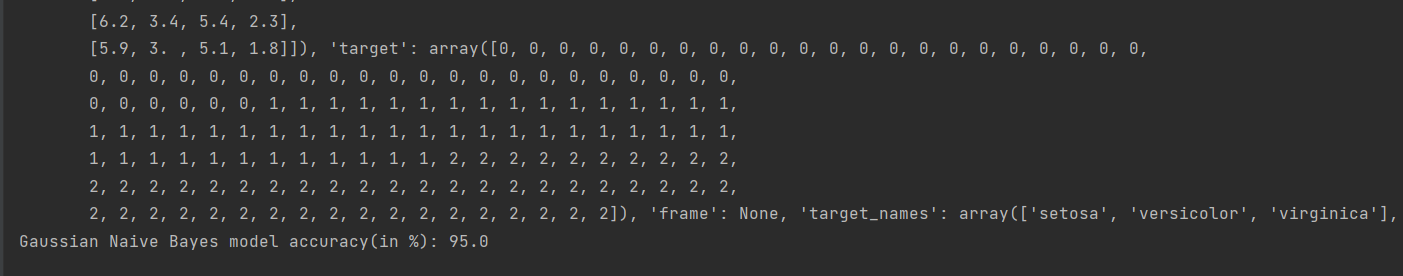
Naive Bayes algorithms are mostly used in sentiment analysis, spam filtering, recommendation systems etc. They are fast and easy to implement but their biggest disadvantage is that the requirement of predictors to be independent.

**Code:-**

import numpy as np  
import pandas as pd  
# load the iris dataset  
  
from sklearn.datasets import load\_iris  
iris = load\_iris()  
print(iris)  
  
# store the feature matrix (X) and response vector (y)  
X = iris.data  
y = iris.target  
  
# splitting X and y into training and testing sets  
from sklearn.model\_selection import train\_test\_split  
X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.4, random\_state=1)  
  
# training the model on training set  
from sklearn.naive\_bayes import GaussianNB  
gnb = GaussianNB()  
gnb.fit(X\_train, y\_train)  
  
# making predictions on the testing set  
y\_pred = gnb.predict(X\_test)  
  
# comparing actual response values (y\_test) with predicted response values (y\_pred)  
from sklearn import metrics  
print("Gaussian Naive Bayes model accuracy(in %):", metrics.accuracy\_score(y\_test, y\_pred)\*100)

**Output:-**

****

****

**CONCLUSION** :- Basic binary classification of iris data or pattern using Naive Bayes Classifier was performed successfully. Students were able to understand dataset and extract the classify data and were able to calculate accuracy. They were also able to understand the concept of training and testing of datasets.