import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.naive\_bayes import GaussianNB

from sklearn import metrics

df = pd.read\_csv('pima\_indian.csv')

feature\_col\_names = ['num\_preg','glucose\_conc','diastolic\_bp','thickness','insulin','bmi','diab\_pred','age']

predicted\_class\_names = ['diabetes']

x = df[feature\_col\_names].values

y = df[predicted\_class\_names].values

xtrain,xtest,ytrain,ytest = train\_test\_split(x,y,test\_size=0.33)

print('\n Total number of Training\_data :', ytrain.shape)

print('\n Total number of Test\_data :', ytest.shape)

clf = GaussianNB().fit(xtrain,ytrain.ravel())

predicted = clf.predict(xtest)

PredictTestData = clf.predict([[6,148,72,35,0,33.6,0.627,50]])

print("\n Confusion matrix")

print(metrics.confusion\_matrix(ytest,predicted))

print("\n Accuracy of the classifier is:",metrics.accuracy\_score(ytest,predicted))

print("\n The value of Precision:",metrics.precision\_score(ytest,predicted))

print("\n The value of Recall:",metrics.recall\_score(ytest,predicted))

print("Predicted value for individual Test Data:",PredictTestData)

OUTPUT:

Total number of Training\_data : (514, 1)

Total number of Test\_data : (254, 1)

Confusion matrix

[[137 25]

[ 36 56]]

Accuracy of the classifier is: 0.7598425196850394

The value of Precision: 0.691358024691358

The value of Recall: 0.6086956521739131

Predicted value for individual Test Data: [1]