04. Write a program to implement the naive Bayesian classifier for a sample training dataset stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.

```
import pandas as pd
from sklearn.model_selection import train_test_split
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_classnames = ['diabetes']

x = df[feature_col_names].values
y = df[predicted_class_names].values

print(df.head)
xtrain, xtest, ytrain, ytest = train_test_split
(x, y, test_size=0.33)

print('The total number of training data:',
      ytrain.shape)
print('The 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("Accuracy of the classifier is :", metrics.
      accuracy_score(ytest, predicted))
print("The value of precision", metrics.precision_
      score(ytest, predicted))
print("The value of Recall", metrics.recall_score
      (ytest, predicted))
print("predicted value :", predictTestData)
```

Output:

```
confusion matrix
[[139  20]
 [74  51]]

Accuracy of the classifier is : 0.748031
The value of precision: 0.7183098591549296
The value of Recall : 0.536842
predicted value : [1]
```