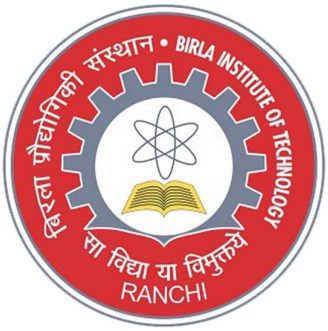
Birla Institute of Technology,Mesra,

Patna Campus



ML-Assignment

Name-Shubham Sourabh

Roll-Btech/15044/18

Sec-CSE 6th

#Assignment-5

Objective: P. Write a program to implement the naive Bayesian classifier for a sample training dataset stored as .CSV file. Compute the accuracy of the classifier, considering few test datasets.

Solution. Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.

Code:

```
# load the iris dataset

from sklearn.datasets import load_iris

iris = load_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: