Intelligent Systems Lab (CS 1763)

Naive Bayesian Classifier

Name: Lakshit Agarwal
Reg No: 201700382
Section: B

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

Ans.

1. IRIS DATASET

```
from sklearn.datasets import load_iris
dataset = load_iris()
X = dataset.data
y = dataset.target
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)
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
y_pred = gnb.predict(X_test)
from sklearn.metrics import accuracy_score
print("Gaussian Naive Bayes accuracy:", accuracy_score(y_test,
y_pred)*100)
```

Output :-

```
Gaussian Naive Bayes accuracy: 95.0
```

2. DIGITS DATASET

```
from sklearn.datasets import load_digits
dataset = load_digits()
X = dataset.data
y = dataset.target
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)
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
y_pred = gnb.predict(X_test)
from sklearn.metrics import accuracy_score
print("Gaussian Naive Bayes accuracy:", accuracy_score(y_test,
y_pred)*100)
```

Output :-

```
Gaussian Naive Bayes accuracy: 83.03198887343532
```

3. WINE DATASET

```
from sklearn.datasets import load_wine
dataset = load_wine()
X = dataset.data
y = dataset.target
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)
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
y_pred = gnb.predict(X_test)
from sklearn.metrics import accuracy_score
print("Gaussian Naive Bayes accuracy:", accuracy_score(y_test,
y_pred)*100)
```

Output :-

```
Gaussian Naive Bayes accuracy: 98.61111111111111
```

4. BREAST CANCER DATASET

```
from sklearn.datasets import load_breast_cancer
dataset = load_breast_cancer()
X = dataset.data
y = dataset.target
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)
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
y_pred = gnb.predict(X_test)
from sklearn.metrics import accuracy_score
print("Gaussian Naive Bayes accuracy:", accuracy_score(y_test,
y_pred)*100)
```

Output :-

```
Gaussian Naive Bayes accuracy: 94.2982456140351
```