**Experiment No.: 8**

**Aim**

Naïve Bayes

**Questions**

Naïve Bayes With CSV

**Program and Output**

1. #upload,read,print csv file

import pandas as pd

import io

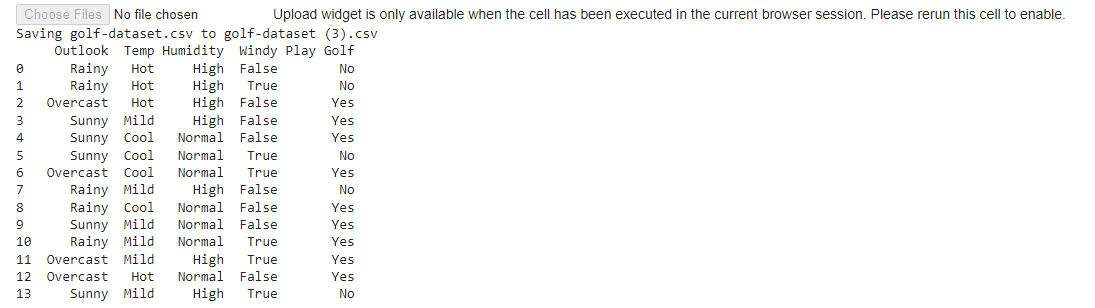
from google.colab import files

uploaded = files.upload()

df1=pd.read\_csv('/content/golf-dataset.csv')

print(df1)

**Output**



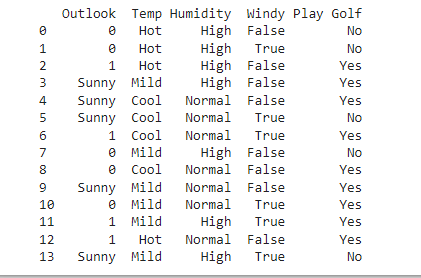
2. #replace attributes in csv

df1['Outlook'].replace(['Rainy'],[0], inplace=True)

df1['Outlook'].replace(['Overcast'],[1], inplace=True)

print(df1)

output



3.

import pandas as pd

import io

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

df1.head(10)

X = df1.iloc[:, [0,1,2, 3]].values

y = df1.iloc[:, -1].values

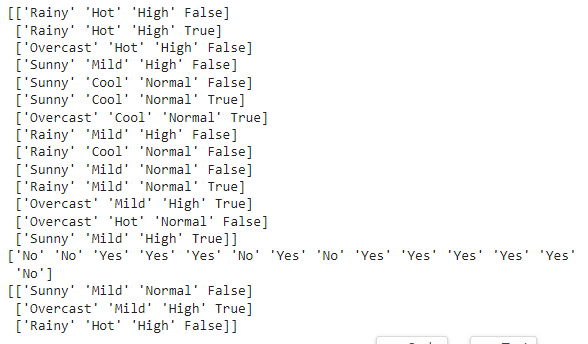
print(X)

print(y)

X\_train, X\_test, y\_train, y\_test = train\_test\_split( X, y, test\_size = 0.2, random\_state=42)

print(X\_test)

**Output**

****

4.

a. # confusion matrix

import pandas as pd

import numpy as np

from sklearn import preprocessing

data= pd.read\_csv('golf-dataset.csv')

label\_encoder = preprocessing.LabelEncoder()

data['Outlook']= label\_encoder.fit\_transform(data['Outlook'])

data['Temp']= label\_encoder.fit\_transform(data['Temp'])

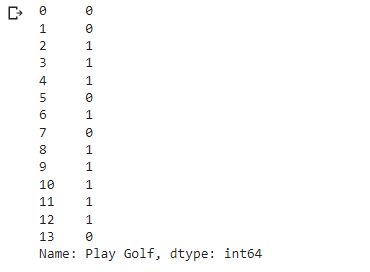
data['Humidity']= label\_encoder.fit\_transform(data['Humidity'])

data['Windy']= label\_encoder.fit\_transform(data['Windy'])

data['Play Golf']= label\_encoder.fit\_transform(data['Play Golf'])

data

**output**



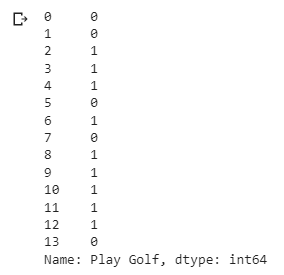
b.

X=data.drop('Play Golf', axis=1)

Y=data['Play Golf']

Y

**Output**



c.

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

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, Y, test\_size=0.1, random\_state=42)

from sklearn.naive\_bayes import GaussianNB

gnb = GaussianNB()

gnb.fit(X\_train, y\_train)

y\_pred = gnb.predict(X\_test)

from sklearn import metrics

print("Gaussian Naive Bayes model accuracy(in %):", metrics.accuracy\_score(y\_test, y\_pred)\*100)

**output**



d. from sklearn import metrics

confusion\_matrix = metrics.confusion\_matrix(Y, Y)

confusion\_matrix

**output**

