***Practical 12***

**Aim:** *Implement Naive Bayes’ learning algorithm for the restaurant waiting problem.*

***Theory:***

*Naïve Bayes Classifier Algorithm*

* *Naïve Bayes algorithm is a supervised learning algorithm, which is based on****Bayes theorem****and used for solving classification problems.*
* *It is mainly used in text classification that includes a high-dimensional training dataset.*
* *Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.*
* *It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.*
* *Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.*

**Why is it called Naïve Bayes?**

*The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:*

* ***Naïve****: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other.*
* ***Bayes****: It is called Bayes because it depends on the principle of Bayes' Theorem.*

***Bayes' Theorem:***

* *Bayes' theorem is also known as****Bayes' Rule****or****Bayes' law****, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.*
* *The formula for Bayes' theorem is given as:*

Naïve Bayes Classifier Algorithm

***Where,***

***P(A|B) is Posterior probability****: Probability of hypothesis A on the observed event B.*

***P(B|A) is Likelihood probability****: Probability of the evidence given that the probability of a hypothesis is true.*

***P(A) is Prior Probability****: Probability of hypothesis before observing the evidence.*

***P(B) is Marginal Probability****: Probability of Evidence.*

*Working of Naïve Bayes' Classifier:*

*Working of Naïve Bayes' Classifier can be understood with the help of the below example:*

*Suppose we have a dataset of****weather conditions****and corresponding target variable "****Play****". So using this dataset we need to decide that whether we should play or not on a particular day according to the weather conditions. So to solve this problem, we need to follow the below steps:*

1. *Convert the given dataset into frequency tables.*
2. *Generate Likelihood table by finding the probabilities of given features.*
3. *Now, use Bayes theorem to calculate the posterior probability.*

***Problem****: If the weather is sunny, then the Player should play or not?*

***Solution****: To solve this, first consider the below dataset:*

|  |  |  |
| --- | --- | --- |
|  | ***Outlook*** | ***Play*** |
| ***0*** | *Rainy* | *Yes* |
| ***1*** | *Sunny* | *Yes* |
| ***2*** | *Overcast* | *Yes* |
| ***3*** | *Overcast* | *Yes* |
| ***4*** | *Sunny* | *No* |
| ***5*** | *Rainy* | *Yes* |
| ***6*** | *Sunny* | *Yes* |
| ***7*** | *Overcast* | *Yes* |
| ***8*** | *Rainy* | *No* |
| ***9*** | *Sunny* | *No* |
| ***10*** | *Sunny* | *Yes* |
| ***11*** | *Rainy* | *No* |
| ***12*** | *Overcast* | *Yes* |
| ***13*** | *Overcast* | *Yes* |

***Frequency table for the Weather Conditions:***

|  |  |  |
| --- | --- | --- |
| *Weather* | *Yes* | *No* |
| *Overcast* | *5* | *0* |
| *Rainy* | *2* | *2* |
| *Sunny* | *3* | *2* |
| *Total* | *10* | *5* |

***Likelihood table weather condition:***

|  |  |  |  |
| --- | --- | --- | --- |
| *Weather* | *No* | *Yes* |  |
| *Overcast* | *0* | *5* | *5/14= 0.35* |
| *Rainy* | *2* | *2* | *4/14=0.29* |
| *Sunny* | *2* | *3* | *5/14=0.35* |
| *All* | *4/14=0.29* | *10/14=0.71* |  |

***Applying Bayes'theorem:***

***P(Yes|Sunny)= P(Sunny|Yes)\*P(Yes)/P(Sunny)***

*P(Sunny|Yes)= 3/10= 0.3*

*P(Sunny)= 0.35*

*P(Yes)=0.71*

*So P(Yes|Sunny) = 0.3\*0.71/0.35=****0.60***

***P(No|Sunny)= P(Sunny|No)\*P(No)/P(Sunny)***

*P(Sunny|NO)= 2/4=0.5*

*P(No)= 0.29*

*P(Sunny)= 0.35*

*So P(No|Sunny)= 0.5\*0.29/0.35 =****0.41***

*So as we can see from the above calculation that****P(Yes|Sunny)>P(No|Sunny)***

***Hence on a Sunny day, Player can play the game.***

*Advantages of Naïve Bayes Classifier:*

* *Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.*
* *It can be used for Binary as well as Multi-class Classifications.*
* *It performs well in Multi-class predictions as compared to the other Algorithms.*
* *It is the most popular choice for****text classification problems****.*

*Disadvantages of Naïve Bayes Classifier:*

* *Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.*

**Applications of Naïve Bayes Classifier:**

* *It is used for****Credit Scoring****.*
* *It is used in****medical data classification****.*
* *It can be used in****real-time predictions****because Naïve Bayes Classifier is an eager learner.*
* *It is used in Text classification such as****Spam filtering****and****Sentiment analysis****.*

***Types of Naïve Bayes Model:***

*There are three types of Naive Bayes Model, which are given below:*

* ***Gaussian****: The Gaussian model assumes that features follow a normal distribution. This means if predictors take continuous values instead of discrete, then the model assumes that these values are sampled from the Gaussian distribution.*
* ***Multinomial****: The Multinomial Naïve Bayes classifier is used when the data is multinomial distributed. It is primarily used for document classification problems, it means a particular document belongs to which category such as Sports, Politics, education, etc.  
  The classifier uses the frequency of words for the predictors.*
* ***Bernoulli****: The Bernoulli classifier works similar to the Multinomial classifier, but the predictor variables are the independent Booleans variables. Such as if a particular word is present or not in a document. This model is also famous for document classification tasks.*

***Code:***

import numpy as nm

import matplotlib.pyplot as mtp

import pandas as pd

# Importing the dataset

dataset = pd.read\_csv('user\_data.csv')

x = dataset.iloc[:, [2, 3]].values

y = dataset.iloc[:, 4].values

# Splitting the dataset into the Training set and Test set

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

x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size = 0.25, random\_state = 0)

# Feature Scaling

from sklearn.preprocessing import StandardScaler

sc = StandardScaler()

x\_train = sc.fit\_transform(x\_train)

x\_test = sc.transform(x\_test)

from sklearn.naive\_bayes import GaussianNB

classifier = GaussianNB()

classifier.fit(x\_train, y\_train)

# Predicting the Test set results

y\_pred = classifier.predict(x\_test)

# Making the Confusion Matrix

from sklearn.metrics import confusion\_matrix

cm = confusion\_matrix(y\_test, y\_pred)

# Visualising the Training set results

from matplotlib.colors import ListedColormap

x\_set, y\_set = x\_train, y\_train

X1, X2 = nm.meshgrid(nm.arange(start = x\_set[:, 0].min() - 1, stop = x\_set[:, 0].max() + 1, step = 0.01),

                     nm.arange(start = x\_set[:, 1].min() - 1, stop = x\_set[:, 1].max() + 1, step = 0.01))

mtp.contourf(X1, X2, classifier.predict(nm.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),

             alpha = 0.75, cmap = ListedColormap(('purple', 'green')))

mtp.xlim(X1.min(), X1.max())

mtp.ylim(X2.min(), X2.max())

for i, j in enumerate(nm.unique(y\_set)):

    mtp.scatter(x\_set[y\_set == j, 0], x\_set[y\_set == j, 1],  c = ListedColormap(('purple', 'green'))(i), label = j)

mtp.title('Naive Bayes (Training set)')

mtp.xlabel('Age')

mtp.ylabel('Estimated Salary')

mtp.legend()

mtp.show()

# Visualising the Test set results

from matplotlib.colors import ListedColormap

x\_set, y\_set = x\_test, y\_test

X1, X2 = nm.meshgrid(nm.arange(start = x\_set[:, 0].min() - 1, stop = x\_set[:, 0].max() + 1, step = 0.01),

                     nm.arange(start = x\_set[:, 1].min() - 1, stop = x\_set[:, 1].max() + 1, step = 0.01))

mtp.contourf(X1, X2, classifier.predict(nm.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),

             alpha = 0.75, cmap = ListedColormap(('purple', 'green')))

mtp.xlim(X1.min(), X1.max())

mtp.ylim(X2.min(), X2.max())

for i, j in enumerate(nm.unique(y\_set)):

    mtp.scatter(x\_set[y\_set == j, 0], x\_set[y\_set == j, 1],

                c = ListedColormap(('purple', 'green'))(i), label = j)

mtp.title('Naive Bayes (test set)')

mtp.xlabel('Age')

mtp.ylabel('Estimated Salary')

mtp.legend()

mtp.show()

***user\_data.csv***

User ID,Gender,Age,Estimated Salary,Purchased

15622,Male,36,15000,1

15669,Female,39,21000,1

15678,Male,31,15000,1

15622,Female,28,46000,0

15641,Female,32,42000,1

15647,Male,35,38000,0

15658,Female,32,28000,1

15650,Female,38,8000,1

15652,Male,21,12000,1

15618,Female,45,17000,0

15682,Female,34,13000,1

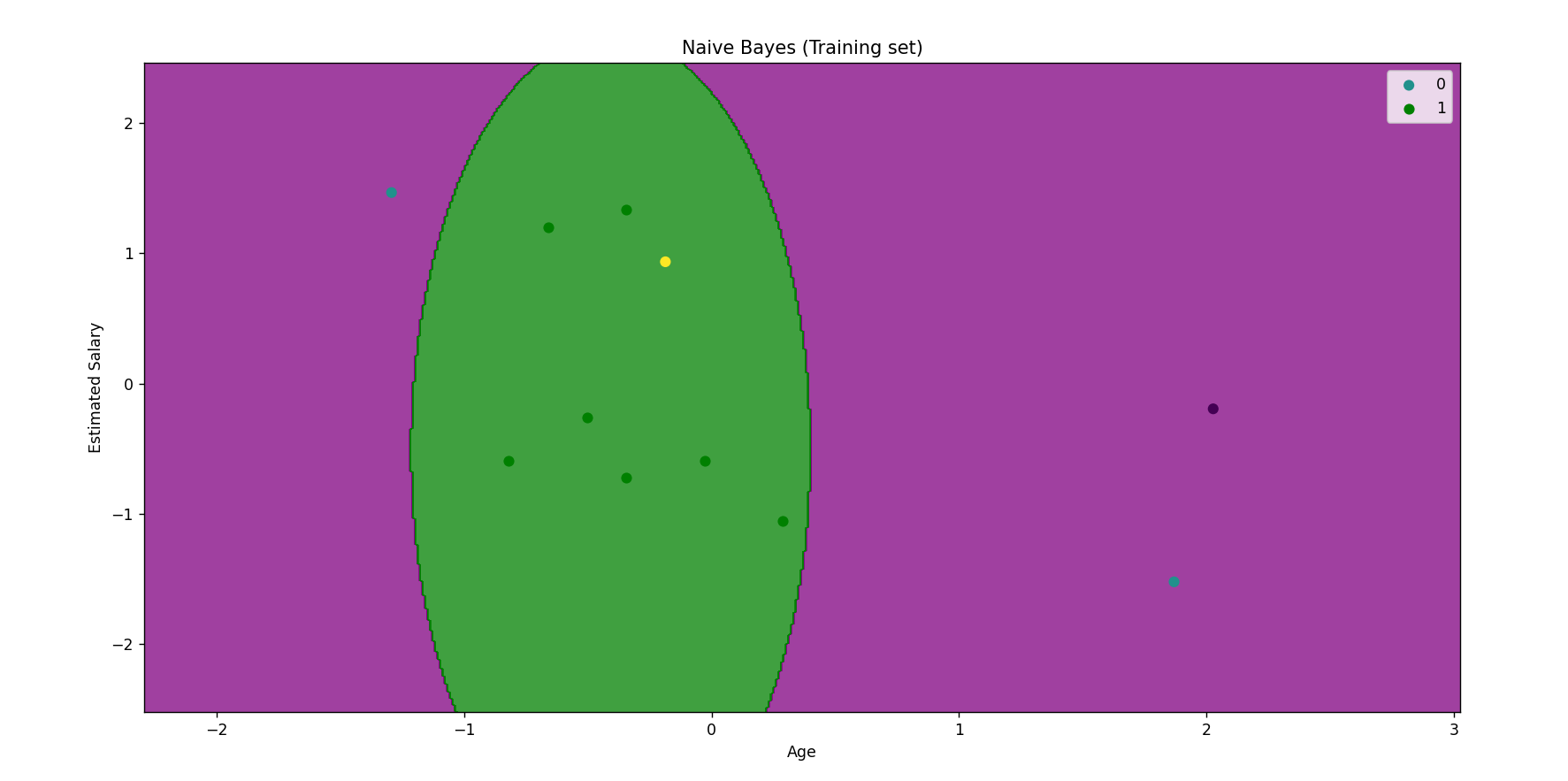
15696,Male,34,44000,1

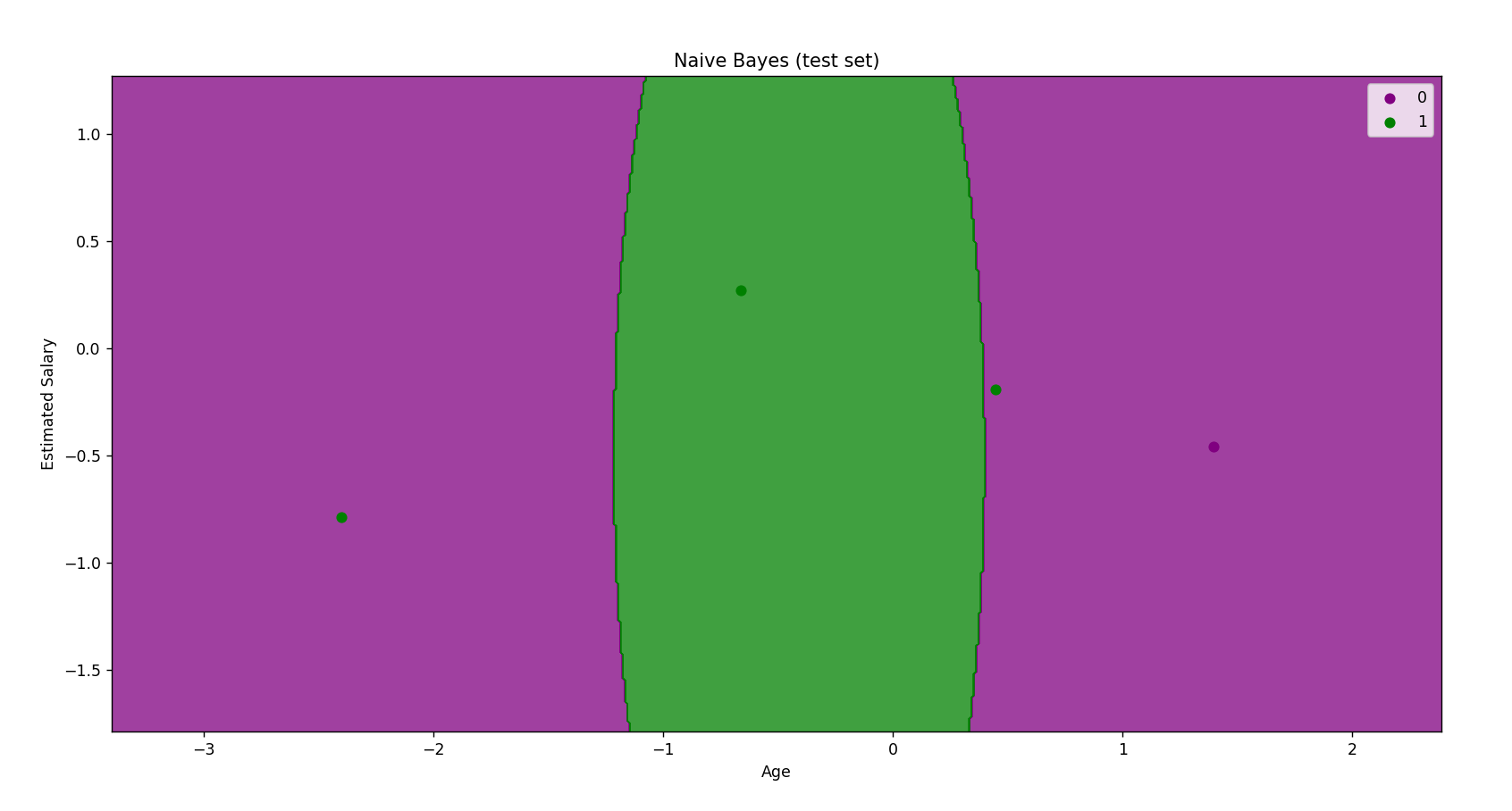
15682,Male,33,20000,1

15650,Female,49,21000,0

15609,Male,48,1000,0

***Output:***

******

******

***Conclusion:***

*Implemented Naive Bayes’ learning algorithm for the restaurant waiting problem.*