Name : Akshit Kothari

Section : BCA A1

Roll no : 07 (2221118)

PROBLEM STATEMENT (16):- Write a program to implement decision tree.

PROGRAM:-

```python
from sklearn.datasets import load_iris

import numpy as np

iris=load_iris()

x=iris.data

y=iris.target


from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=42)


from sklearn.tree import DecisionTreeClassifier

clf=DecisionTreeClassifier(criterion='gini',max_depth=3,random_state=42)

clf.fit(x_train,y_train)

y_pred=clf.predict(x_test)

from sklearn.metrics import accuracy_score

accu=accuracy_score(y_test,y_pred)

print("accuracy",accu*100)


from sklearn import tree

import matplotlib.pyplot as plt

plt.figure(figsize=(12,8))

tree.plot_tree(clf,filled=True,feature_names=iris.feature_names,class_names=iris.target_names)

plt.show()
```
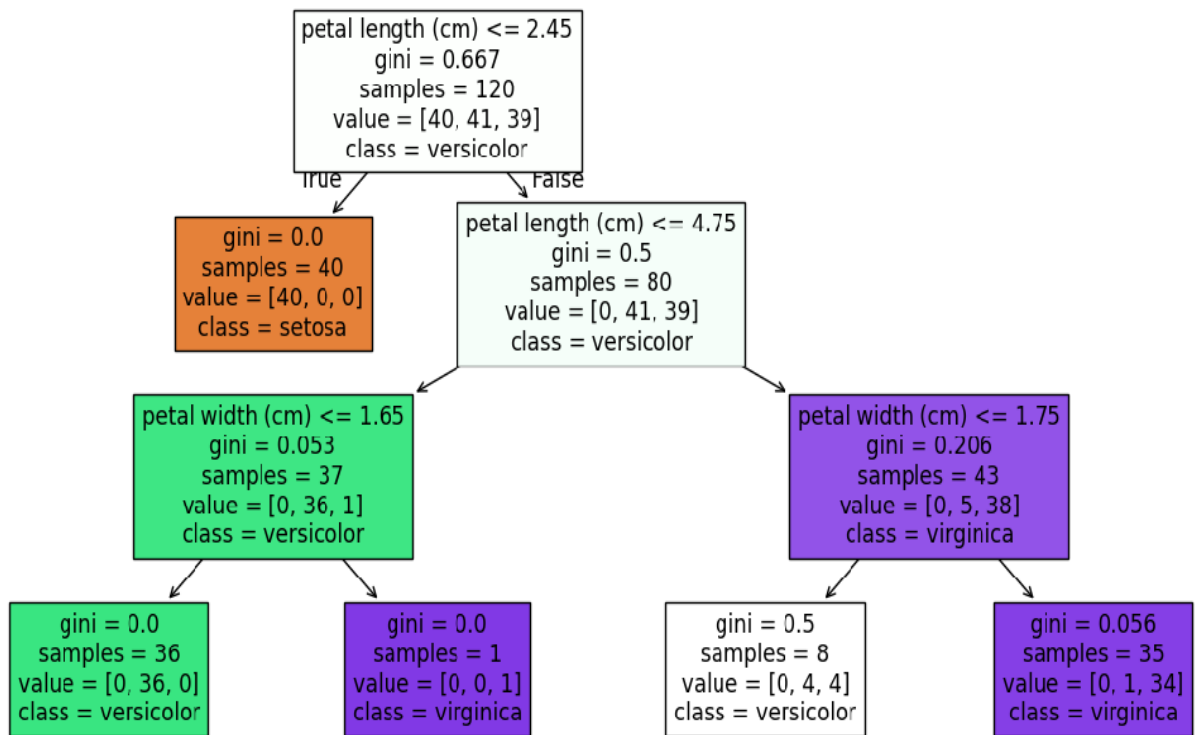
Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

Output:

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

PROBLEM STATEMENT (15):- Write a program to implement naïve bayes.

PROGRAM:-

from sklearn.datasets import load_iris

import numpy as np

iris=load_iris()

x=iris.data

y=iris.target

from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=42)

from sklearn.naive_bayes import GaussianNB

gb=GaussianNB()

gb.fit(x_train,y_train)

y_pred=gb.predict(x_test)

from sklearn.metrics import accuracy_score

accu=accuracy_score(y_test,y_pred)

print("prediction",y_pred)

print("accuracy",accu*100)

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

Output:

prediction [1 0 2 1 1 0 1 2 1 1 2 0 0 0 0 1 2 1 1 2 0 2 0 2 2 2 2 2 0 0]

accuracy 100.0

Name : Akshit Kothari

Section : BCA A1

Roll no : 07 (2221118)

PROBLEM STATEMENT (14):- Write a program to implement SVM.

PROGRAM:-

```
from sklearn.datasets import load_iris

iris=load_iris()

x=iris.data[:,:2]

y=iris.target


from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)


from sklearn.svm import SVC

svm=SVC(kernel='linear')

svm.fit(x_train,y_train)


y_pred=svm.predict(x_test)


from sklearn.metrics import accuracy_score

acu=accuracy_score(y_test,y_pred)

print(acu)

print(y_pred)


import numpy as np

x_min,x_max=x[:,0].min()-1,x[:,0].max()+1

y_min,y_max=x[:,1].min()-1,x[:,1].max()+1

xx,yy=np.meshgrid(np.arange(x_min,x_max,0.02),np.arange(y_min,y_max,0.02))

z=svm.predict(np.c_[xx.ravel(),yy.ravel()])
```
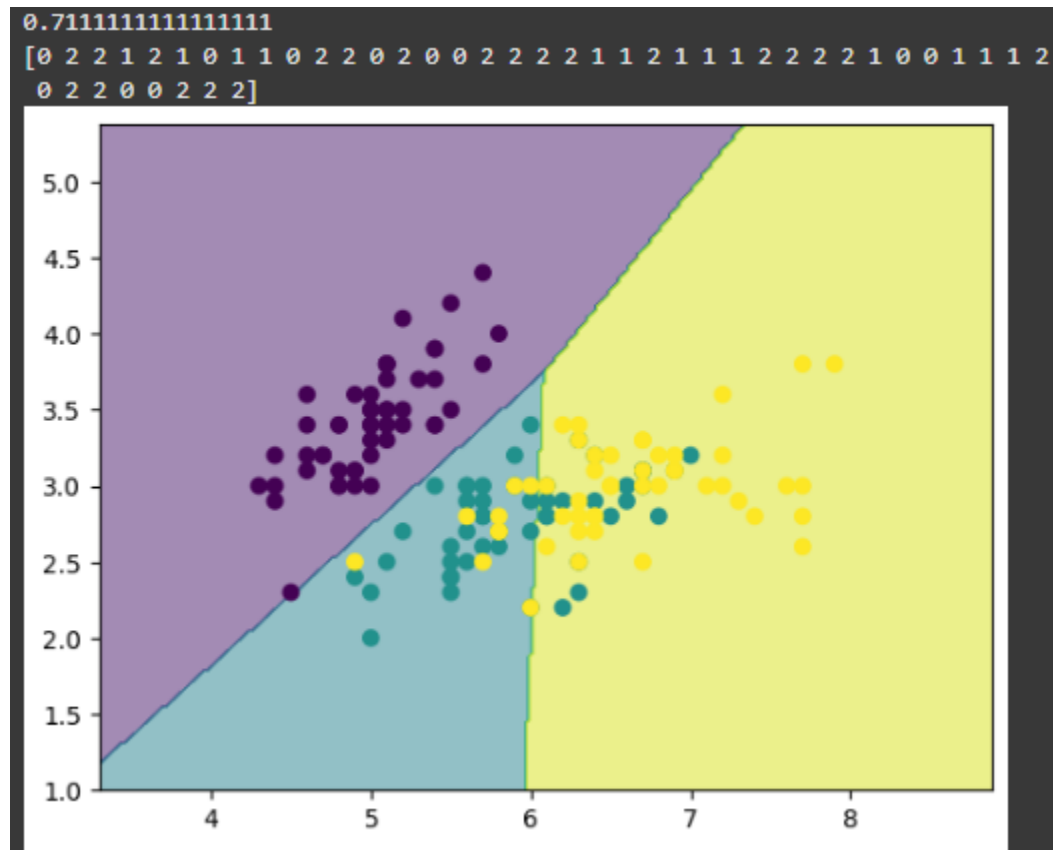
Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

```python
import matplotlib.pyplot as plt

plt.contourf(xx,yy,z.reshape(xx.shape),alpha=0.5)

plt.scatter(x[:,0],x[:,1],c=y,cmap='viridis')

plt.show()
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

Output:

```
0.71111111111111111
[0 2 2 1 2 1 0 1 1 0 2 2 0 2 0 0 2 2 2 2 1 1 2 1 1 1 2 2 2 2 1 0 0 1 1 1 2
 0 2 2 0 0 2 2 2]
```

Name : Akshit Kothari

Section : BCA A1

Roll no : 07 (2221118)

PROBLEM STATEMENT (17):- Write a program to implement PCA.

PROGRAM:-

```python
import numpy as np

import matplotlib.pyplot as plt

from sklearn.datasets import load_iris

iris=load_iris()

x=iris.data

y=iris.target


from sklearn.preprocessing import StandardScaler

x_scaled=StandardScaler().fit_transform(x)



from sklearn.decomposition import PCA

pca=PCA(n_components=2)

x_pca=pca.fit_transform(x_scaled)



plt.scatter(x_pca[:,0],x_pca[:,1],c=y,cmap='viridis')

plt.colorbar()

plt.show()
```
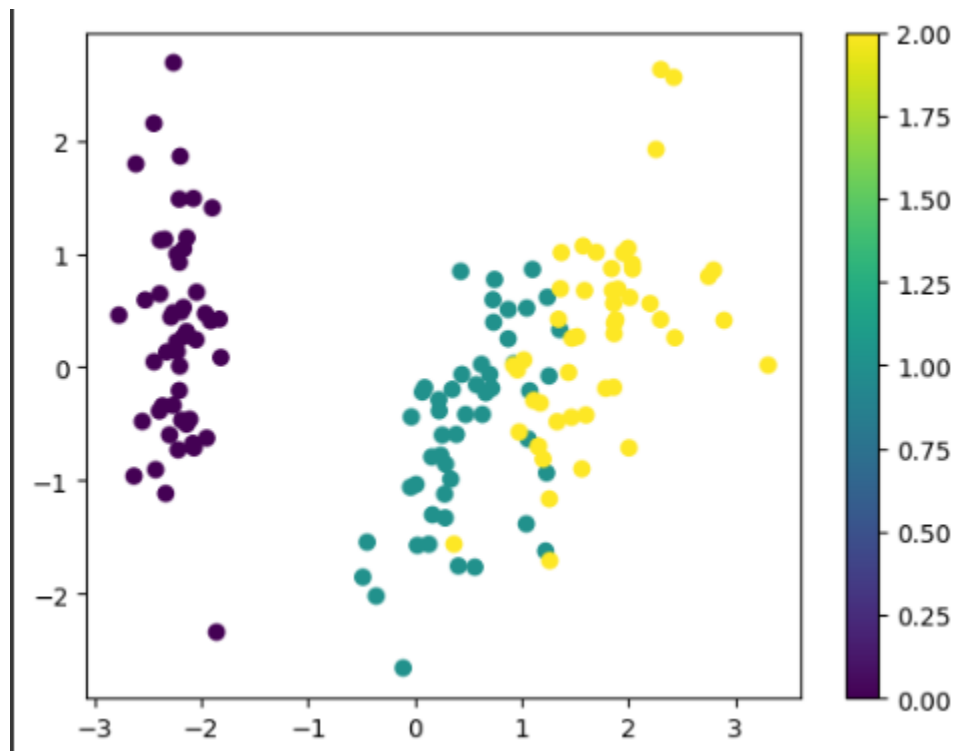
Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

Output:

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

PROBLEM STATEMENT (18):- Uploading an excel/csv data file (containing Student_ID, Student_Name, Gender, Sub1, Sub2, Sub3 with the marks of 30 students). Perform the following tasks:

I.Check for missing values, and replace them with suitable replacement.

PROGRAM:-

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

print(df.head())

df.fillna(80,inplace=True)

print(df.head())


```
= RESTART: C:\Users\Akshit Kothari\Desktop\newwwwwwwwwww.py
    Student_ID Student_Name Gender  Sub1  Sub2  Sub3
0            1          Ram      M  90.0  67.0  87.0
1            2        Seema      F  80.0  89.0  86.0
2            3        Vijay      M   NaN  77.0  65.0
3            4         Raju      M  86.0  89.0   NaN
4            5        Rohit      M  65.0  76.0  78.0
    Student_ID Student_Name Gender  Sub1  Sub2  Sub3
0            1          Ram      M  90.0  67.0  87.0
1            2        Seema      F  80.0  89.0  86.0
2            3        Vijay      M  80.0  77.0  65.0
3            4         Raju      M  86.0  89.0  80.0
4            5        Rohit      M  65.0  76.0  78.0
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

II. Create two DataFrames containg Student_ID, Student_Name of male and female students.

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

#print(df.head())

df_male=df[df['Gender']=='M'][['Student_ID','Student_Name','Gender']]

df_female=df[df['Gender']=='F'][['Student_ID','Student_Name','Gender']]

print(df_male.head())

print(df_female.head())

```
= RESTART: C:\Users\Akshit Kothari\Desktop\newwwwwwwwww.py
    Student_ID Student_Name Gender
0            1          Ram      M
2            3        Vijay      M
3            4         Raju      M
4            5        Rohit      M
5            6        Virat      M
    Student_ID Student_Name Gender
1            2        Seema      F
10          11        Reena      F
12          13        Meena      F
14          15        Perry      F
15          16       Smriti      F
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

III.Add a new column in the DataFrame 'Percentage' showing total percentage of each student.

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

df['perc']=df[['Sub1','Sub2','Sub3']].mean(axis=1)

df.to_excel(r"C:\Users\Akshit Kothari\Desktop\students_updated.xlsx",index=False)

print(df.head())

```
= RESTART: C:\Users\Akshit Kothari\Desktop\newwwwwwwwwww.py
   Student_ID Student_Name Gender  Sub1  Sub2  Sub3       perc
0           1          Ram      M  90.0  67.0  87.0  81.333333
1           2        Seema      F  80.0  89.0  86.0  85.000000
2           3        Vijay      M   NaN  77.0  65.0  71.000000
3           4         Raju      M  86.0  89.0   NaN  87.500000
4           5        Rohit      M  65.0  76.0  78.0  73.000000
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

IV. Normalizing the marks of each subject .

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

print(df.head())

df.fillna(80,inplace=True)

for subject in ['Sub1','Sub2','Sub3']:

   df['{subject}_Norm']=(df[subject]-df[subject].min())/(df[subject].max()-df[subject].min())

print(df.head())

```
=========== RESTART: C:\Users\Akshit Kothari\Desktop\newwwwwwwwww.py
    Student_ID Student_Name Gender Sub1   Sub2   Sub3
0   1   Ram   M 90.0   67.0   87.0
1   2   Seema   F 80.0   89.0   86.0
2   3   Vijay   M 80.0   77.0   65.0
3   4   Raju   M 86.0   89.0   80.0
4   5   Rohit   M 65.0   76.0   78.0
    Student_ID Student_Name Gender Sub1   Sub2   Sub3 {subject}_Norm
0   1   Ram   M 90.0   67.0   87.0   0.916667
1   2   Seema   F 80.0   89.0   86.0   0.875000
2   3   Vijay   M 80.0   77.0   65.0   0.000000
3   4   Raju   M 86.0   89.0   80.0   0.625000
4   5   Rohit   M 65.0   76.0   78.0   0.541667
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

**V**.Draw a bar diagram showing number of male and female students in the class.

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt


df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

df.fillna(80,inplace=True)

gen_count=df['Gender'].value_counts()

gen_count.plot(kind='bar',color=['skyblue','pink'])

plt.title("No. Of male and female")

plt.xlabel("gender")

plt.ylabel("Count")

plt.xticks(rotation=0)

plt.show()
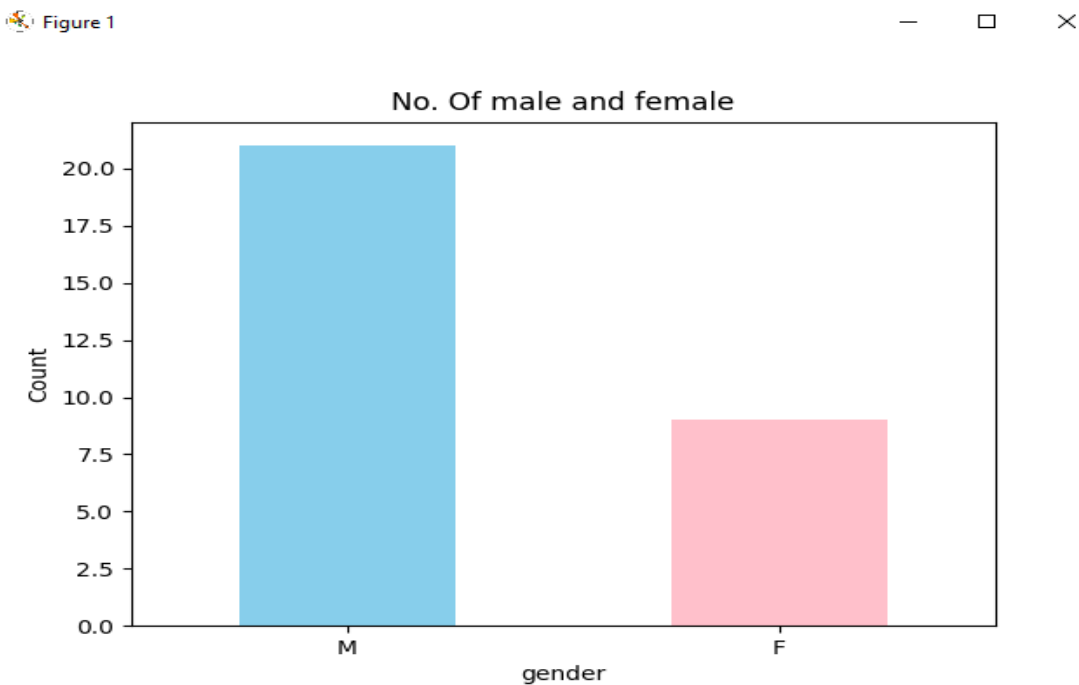
Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

**VI**.Draw a pie chart showing the number of students having percentage

(a) > =60 (b) > =50 and < 60 (c) < 50

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read_excel(r"C:\Users\Akshit Kothari\Desktop\student.xlsx")

df.fillna(80,inplace=True)

df['perc']=df[['Sub1','Sub2','Sub3']].mean(axis=1)

df.to_excel(r"C:\Users\Akshit Kothari\Desktop\students_updated.xlsx",index=False)

high=df[df['perc']>=60].shape[0]

med=df[(df['perc']>=50) & (df['perc']<60)].shape[0]

low=df[df['perc']<50].shape[0]

labels=['>=60%','50%-59%','<50%']

sizes=[high,med,low]

color=['green','yellow','red']

plt.pie(sizes,labels=labels,colors=color,autopct='%1.1f%%',startangle=140)

plt.axis('equal')

plt.show()
```
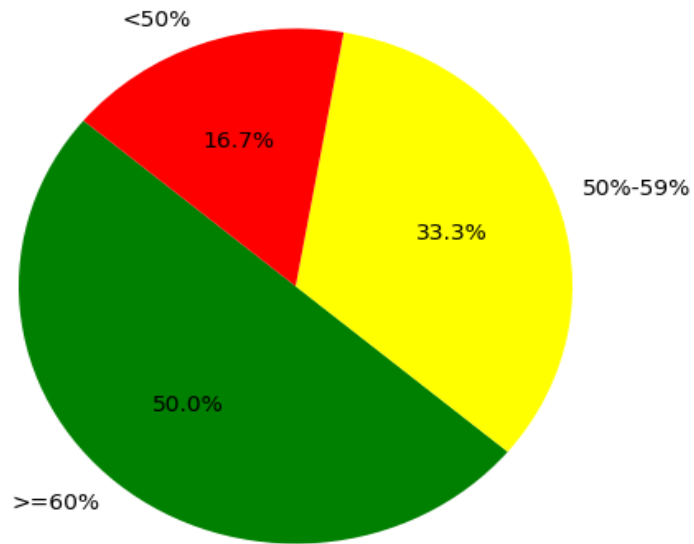
Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

Name : Akshit Kothari

Section : BCA A1

Roll no : 07 (2221118)

**PROBLEM STATEMENT (19):-** Create a .txt file in your directory with three lines as follows:

Hi how are you?

I am fine.

I hope that you are also fine.

I. Display the content of the file as a string.

II. Display each line as an element of a list.

III. Display the number of characters in the file.

IV. Number of characters in first line.

V. 2nd to 5th characters of second last line.

**PROGRAM:-**

```
f=open("p.txt","w")

f.write("hi how are you?\n")

f.write("I am fine.\n")

f.write("I hope you are also fine.")

f.close()




f=open("p.txt","r")

content=f.read()

print("content as a string : ",content)

f.close()




f=open("p.txt","r")

lines=f.readlines()

print("content as list: ",lines)




count=len(content)
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

```python
print("Total character in file: ",count)



first_line=len(lines[0].strip())

print("Character in first line: ",first_line)



second_line=lines[-2].strip()

print("2nd to 5th Character of second line: ",second_line)
```

Output:

```
= RESTART: C:/Users/Akshit Kothari/Desktop/kjhkj.py
content as a string :  hi how are you?
I am fine.
I hope you are also fine.
content as list:  ['hi how are you?\n', 'I am fine.\n', 'I hope you are also fin
e.']
Total character in file:  52
Character in first line:  15
2nd to 5th Character of second line:  I am fine.
```

Name : Akshit Kothari
Section : BCA A1
Roll no : 07 (2221118)

VI. Rename the file

VII. Delete the file

```python
import os

f=open("p.txt","w")

f.write("hi how are you?\n")

f.write("I am fine.\n")

f.write("I hope you are also fine.")

f.close()


new_file="renamep.txt"

os.rename("p.txt",new_file)

print(f"file renamed to {new_file}")


os.remove(new_file)

print(f"file {new_file} deleted.")
```

Output:

```
= RESTART: C:/Users/Akshit Kothari/Desktop/kjhkj.py
file renamed to renamep.txt
file renamep.txt deleted.
```