PRACTICAL NO : 05

DATA ANALYTICS 2

**CODE :**

import pandas as pd

df=pd.read\_csv(r'E:\DSBDA\DSBDA Datasets\Social\_Network\_Ads.csv')

print(df)

df['Gender']

df.isnull()

df.dtypes

df['Gender']=df['Gender'].map({'Male':1,'Female':0})

df['Gender']

x=df.drop(['Age'],axis=1)

y=df['Age']

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

xtrain,xtest,ytrain,ytest=train\_test\_split(x,y,test\_size=0.25,random\_state=0)

from sklearn.preprocessing import StandardScaler

st\_x=StandardScaler()

xtrain=st\_x.fit\_transform(xtrain)

xtest=st\_x.transform(xtest)

from sklearn.linear\_model import LogisticRegression

classifier=LogisticRegression(random\_state=0)

classifier.fit(xtrain, ytrain)

y\_pred=classifier.predict(xtest)

y\_pred

from sklearn.metrics import confusion\_matrix

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

cm

OUTPUT :

runfile('E:/DSBDA/dsbdapr5.py', wdir='E:/DSBDA')

User ID Gender Age EstimatedSalary Purchased

0 15624510 Male 19 19000 0

1 15810944 Male 35 20000 0

2 15668575 Female 26 43000 0

3 15603246 Female 27 57000 0

4 15804002 Male 19 76000 0

.. ... ... ... ... ...

395 15691863 Female 46 41000 1

396 15706071 Male 51 23000 1

397 15654296 Female 50 20000 1

398 15755018 Male 36 33000 0

399 15594041 Female 49 36000 1

[400 rows x 5 columns]

df.isnull()

Out[3]:

User ID Gender Age EstimatedSalary Purchased

0 False False False False False

1 False False False False False

2 False False False False False

3 False False False False False

4 False False False False False

.. ... ... ... ... ...

395 False False False False False

396 False False False False False

397 False False False False False

398 False False False False False

399 False False False False False

[400 rows x 5 columns]

df.dtypes

Out[4]:

User ID int64

Gender int64

Age int64

EstimatedSalary int64

Purchased int64

dtype: object

cm

Out[5]:

array([[0, 0, 0, ..., 0, 0, 0],

[0, 0, 0, ..., 0, 0, 0],

[0, 0, 0, ..., 0, 0, 0],

...,

[0, 0, 0, ..., 0, 0, 0],

[0, 0, 0, ..., 0, 0, 0],

[0, 0, 0, ..., 0, 0, 0]], dtype=int64)

**OUTPUT :**

