**import pandas as pd import numpy as py**

**5/7/24, 5:27 PM Assignment 5 .ipynb - Colab**

[**https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB\_PQe#printMode=true**](https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB_PQe#printMode=true) **1/3**

**data=pd.read\_csv('/content/Social\_Network\_Ads.csv.) df=pd.DataFrame(data)**

**df**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **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()** |  |  |  |  |  |
|  | **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**

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

**X=df[['Age','EstimatedSalary']]**

**Y=df[['Purchased']]**

**X\_train,X\_test,Y\_train,Y\_test=train\_test\_split(X,Y,test\_size=0.3)**

**print(X\_train.shape) print(X\_test.shape) print(Y\_train.shape) print(Y\_test.shape)**

**(280, 2) (120, 2) (280, 1) (120, 1)**

from sklearn.preprocessing import StandardScaler sc=StandardScaler()

**5/7/24, 5:27 PM Assignment 5 .ipynb - Colab**

[**https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB\_PQe#printMode=true**](https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB_PQe#printMode=true) **2/3**

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

X\_test=sc.fit\_transform(X\_test)

from sklearn.linear\_model import LogisticRegression

Reg=LogisticRegression()
  
Reg.fit(X\_train,Y\_train)

/usr/local/lib/python3.10/dist-packages/sklearn/utils/validation.py:1143: DataConversior y = column\_or\_1d(y, warn=True)

• LogisticRegression

LogisticRegression() Y\_pred=Reg.predict(X\_test) print(Y\_pred)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [0 0 | 0 0 | 1 0 | 1 0 | 0 | 0 1 | 0 0 | 1 0 | 0 1 | 0 0 | 0 0 | 0 0 | 1 0 | 1 0 | 1 0 | 0 0 | 0 0 | 0 1 | 1 0 |
| 1 0 | 0 0 | 0 1 | 1 0 | 0 | 1 0 | 0 1 | 0 0 | 0 0 | 1 0 | 0 0 | 0 0 | 0 0 | 1 0 | 1 0 | 1 1 | 0 0 | 1 0 | 0 1 |
| 0 0 | 0 0 | 0 1 | 0 0 | 0 | 0 0 | 1 0 | 0 0 | 1 1 | 0 0 | 1 1 | 0 0 | 0 0 | 1 1 | 0 0 | 0 0 | 1 0 | 0 1 | 0 0 |
| 0 0 | 1 0 | 1 1 | 1 1 | 0] |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

from sklearn.metrics import confusion\_matrix,classification\_report

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

cl\_report=classification\_report(Y\_test,Y\_pred)

print(cm)

[[73 5]

[10 32]]

print(cl\_report)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| 0 | 0.88 | 0.94 | 0.91 | 78 |
| 1 | 0.86 | 0.76 | 0.81 | 42 |
| accuracy |  |  | 0.88 | 120 |
| macro avg | 0.87 | 0.85 | 0.86 | 120 |
| weighted avg | 0.87 | 0.88 | 0.87 | 120 |

from sklearn.metrics import precision\_score,recall\_score,accuracy\_score

ps=precision\_score(Y\_test,Y\_pred)

print(ps)

0.8648648648648649

rs=recall\_score(Y\_test,Y\_pred)

print(rs)

0.7619047619047619

**0.125**

**5/7/24, 5:27 PM Assignment 5 .ipynb - Colab**

**AS=accuracy\_score(Y\_test,Y\_pred)**

**error\_rate=1-AS print(error\_rate)**

[**https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB\_PQe#printMode=true**](https://colab.research.google.com/drive/1Ne6qfsVxcsPJVJfAmg6La8gScnXB_PQe#printMode=true) **3/3**

Start coding or generate with AI.