NAME : Nitesh Yadav

Subject : DATA SCIENCE

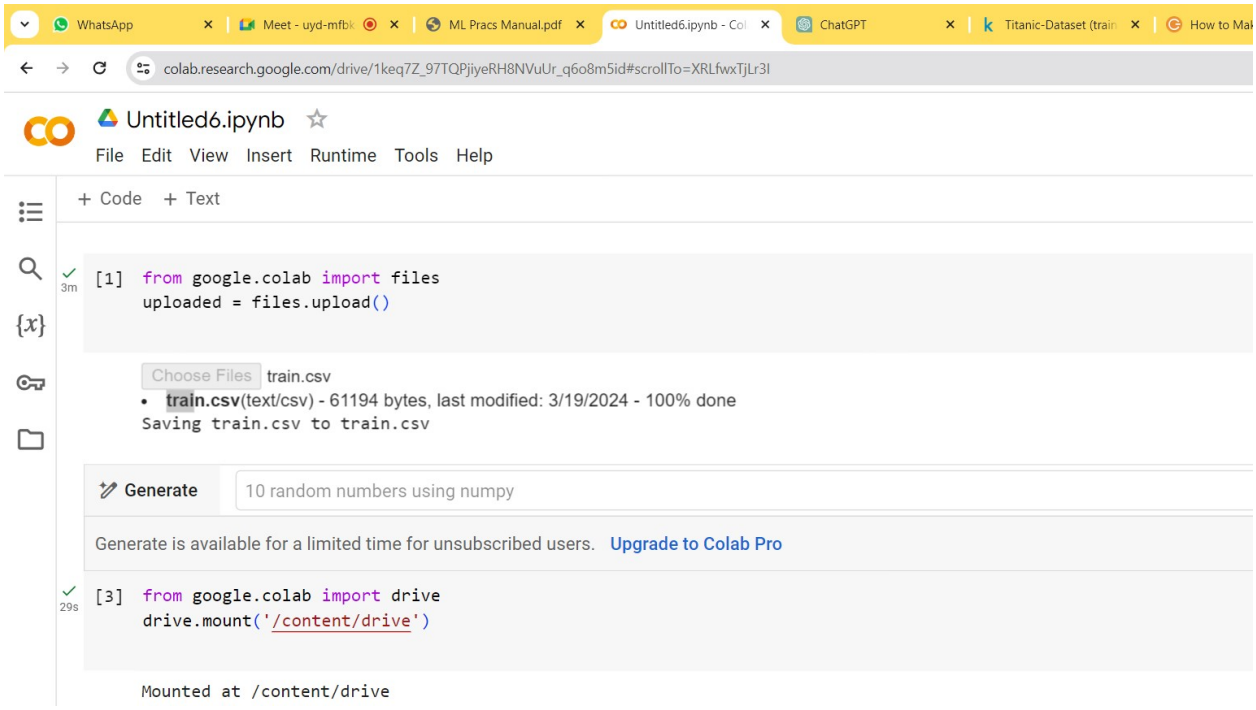
Course :TYIT

Seat NO. : 310921060

Practical No. 1

Q1. Importing data

To import a dataset into a Google Colab notebook, you have several options depending on where your dataset is located and the format it's in. Here's how you can import datasets from different sources:



1. Upload from Local File System:

- If the dataset is on your local machine, you can upload it directly to your Colab notebook. You can use the following code to upload a file:

```python
from google.colab import files
uploaded = files.upload()
```

- This will prompt you to select files from your local machine and upload them to the current working directory in your Colab notebook.

2. Mount Google Drive:

- If your dataset is stored in Google Drive, you can mount your Google Drive in the notebook and access the files. Follow these steps:

```python
from google.colab import drive
drive.mount('/content/drive')
```

```

- After running this code, you'll need to authenticate and provide an authorization code. Once mounted, you can access your files under `/content/drive/My Drive/`.

3. Direct Download from URL:

- If the dataset is available via a direct URL, you can download it using `wget` or `curl`. For example:

```python
!wget <URL>
```

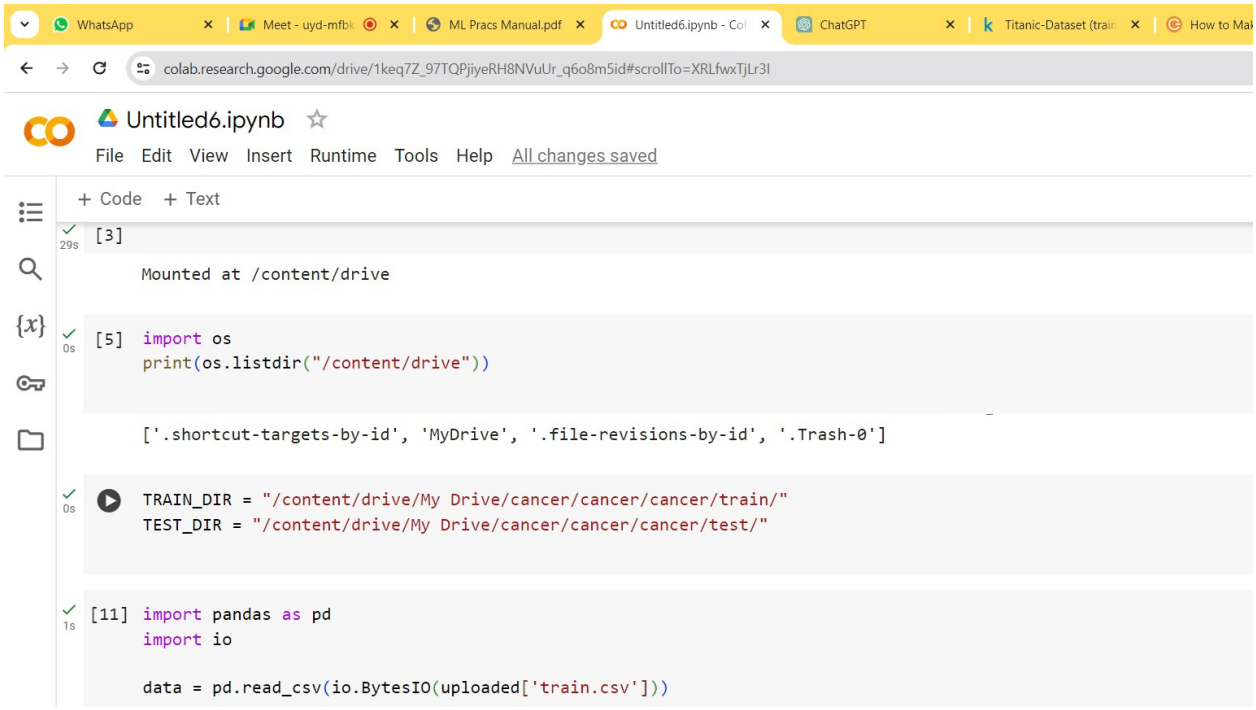
or

```python
!curl -O <URL>
```

- Replace `<URL>` with the actual URL of the dataset.

4. Using APIs or Libraries:
```

- If your dataset is available through APIs or specific libraries like `torchvision` (for image datasets), `tensorflow_datasets` (for TensorFlow datasets), or `scikit-learn` (for built-in datasets), you can directly load it using their respective methods.

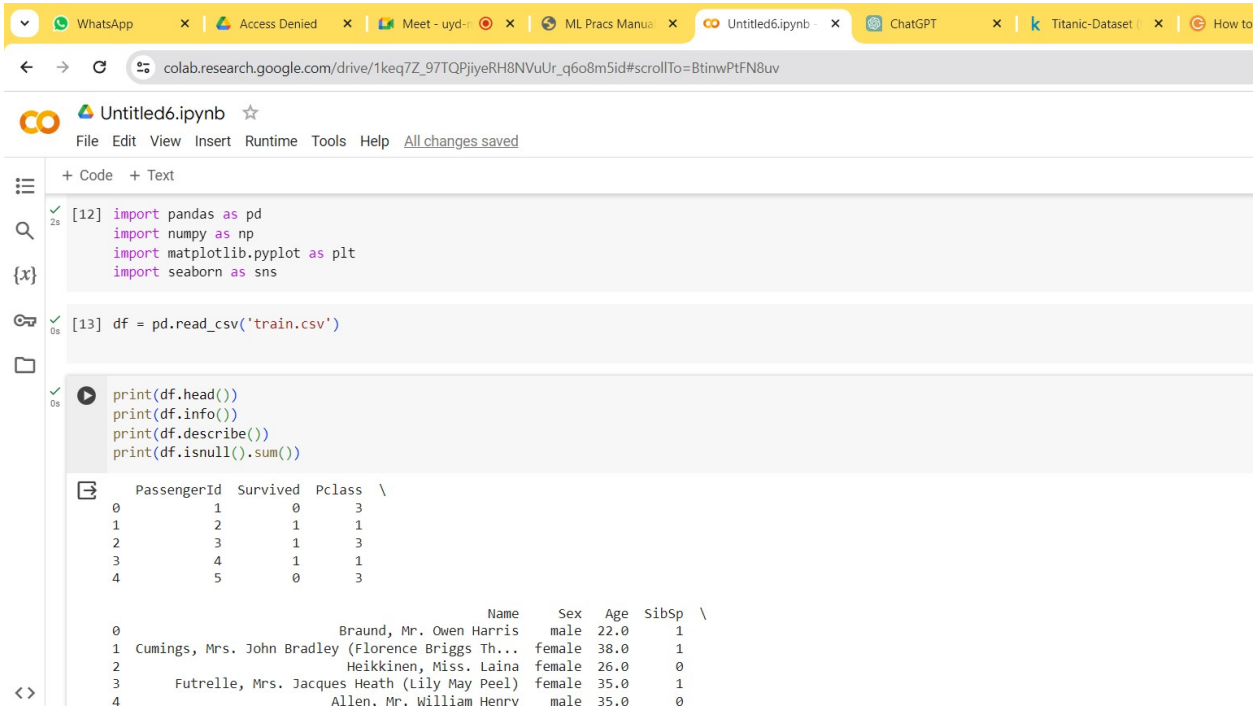


Choose the method that suits your dataset's location and format, and then proceed with further processing or analysis as needed in your Colab notebook.

Practical No. 2

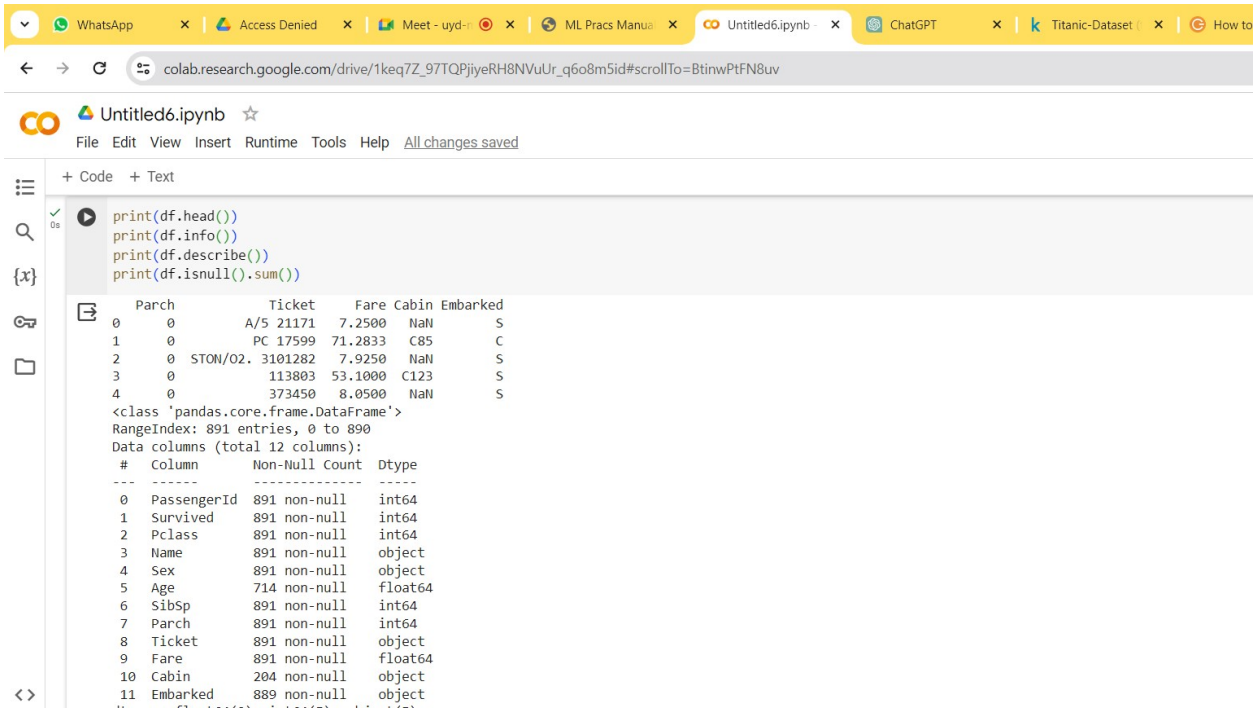
To summerize the dataset in colab notebook

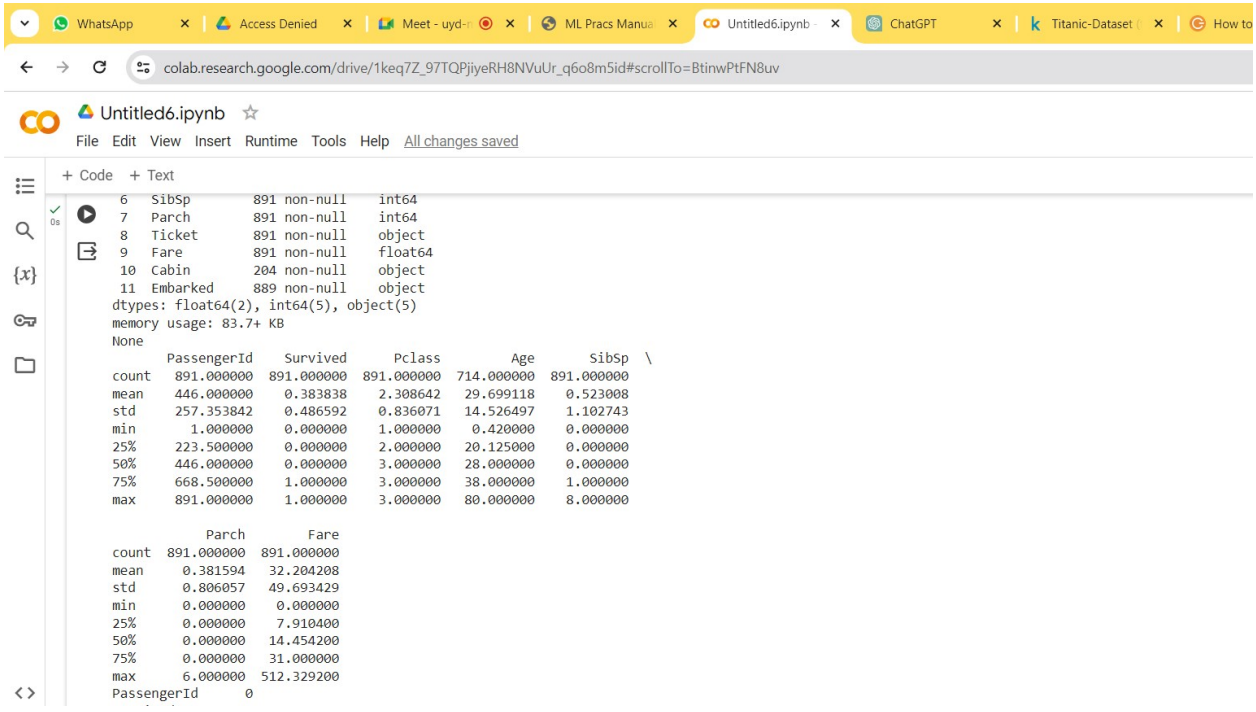
Import Necessary Libraries: Make sure you have libraries like Pandas, NumPy, and any visualization libraries installed.

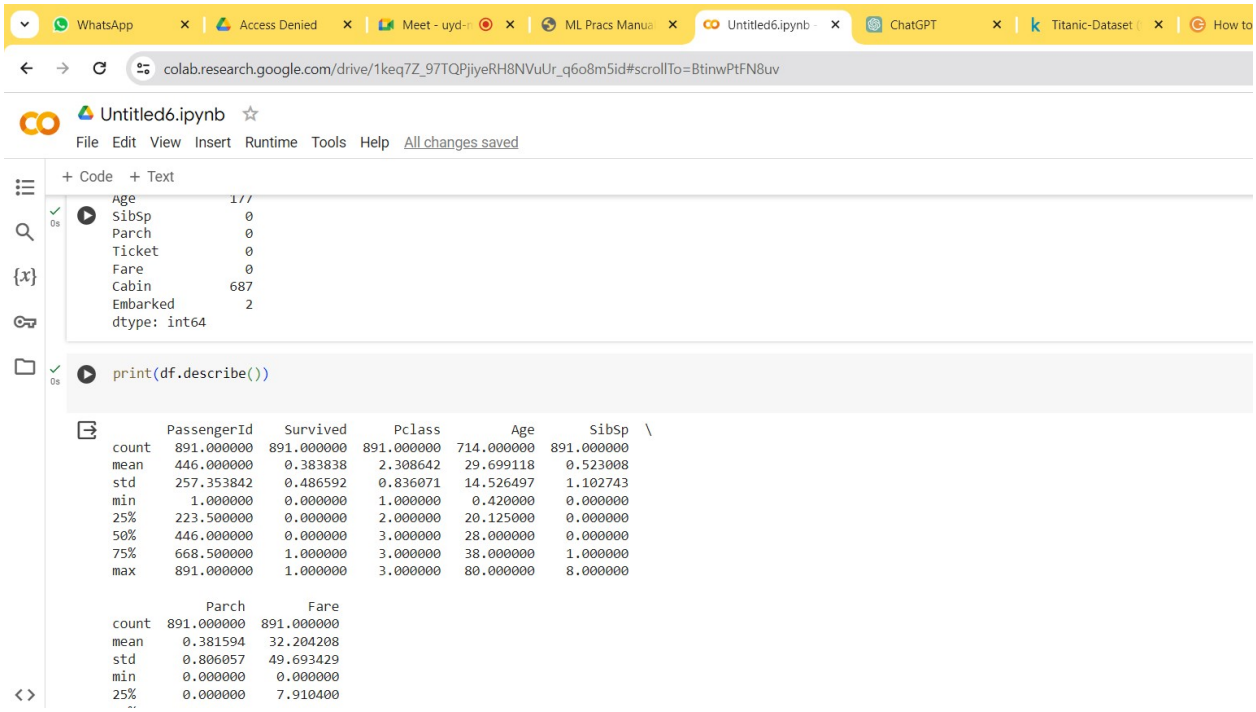


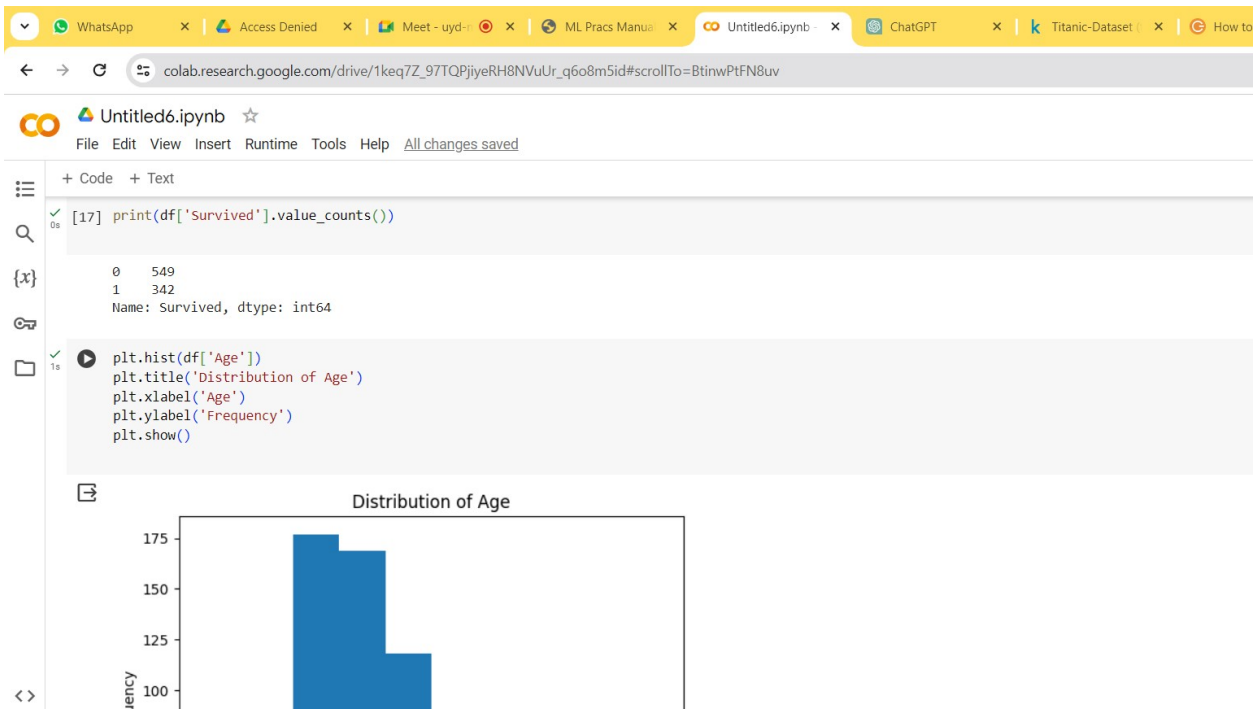
Load the Dataset: Load your dataset into a Pandas DataFrame.

Exploratory Data Analysis (EDA): Perform basic exploratory data analysis to understand the structure, content, and summary statistics of your dataset. Some common tasks include:

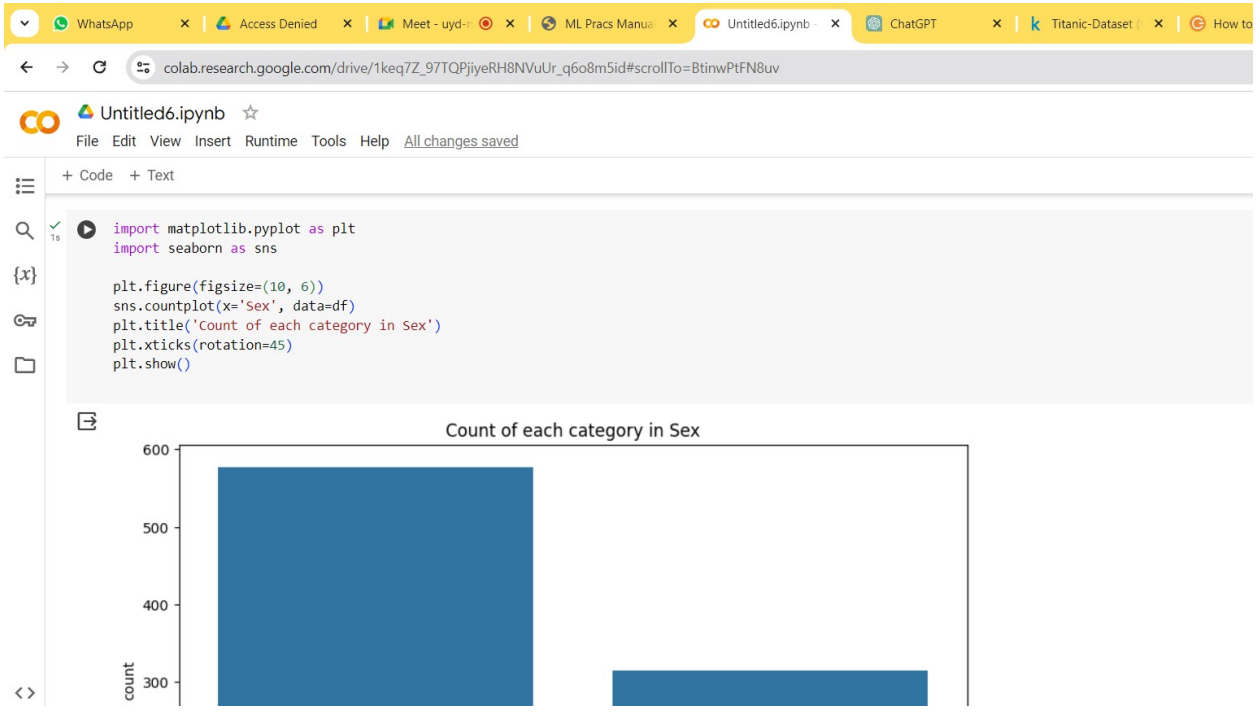

WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
print(df.head())
print(df.info())
print(df.describe())
print(df.isnull().sum())
Parch Ticket Fare Cabin Embarked
0 0 A/5 21171 7.2500 NaN S
1 0 PC 17599 71.2833 C85 C
2 0 STON/O2. 3101282 7.9250 NaN S
3 0 113803 53.1000 C123 S
4 0 373450 8.0500 NaN S
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
Column Non-Null Count Dtype
0 PassengerId 891 non-null int64
1 Survived 891 non-null int64
2 Pclass 891 non-null int64
3 Name 891 non-null object
4 Sex 891 non-null object
5 Age 714 non-null float64
6 SibSp 891 non-null int64
7 Parch 891 non-null int64
8 Ticket 891 non-null object
9 Fare 891 non-null float64
10 Cabin 204 non-null object
11 Embarked 889 non-null object

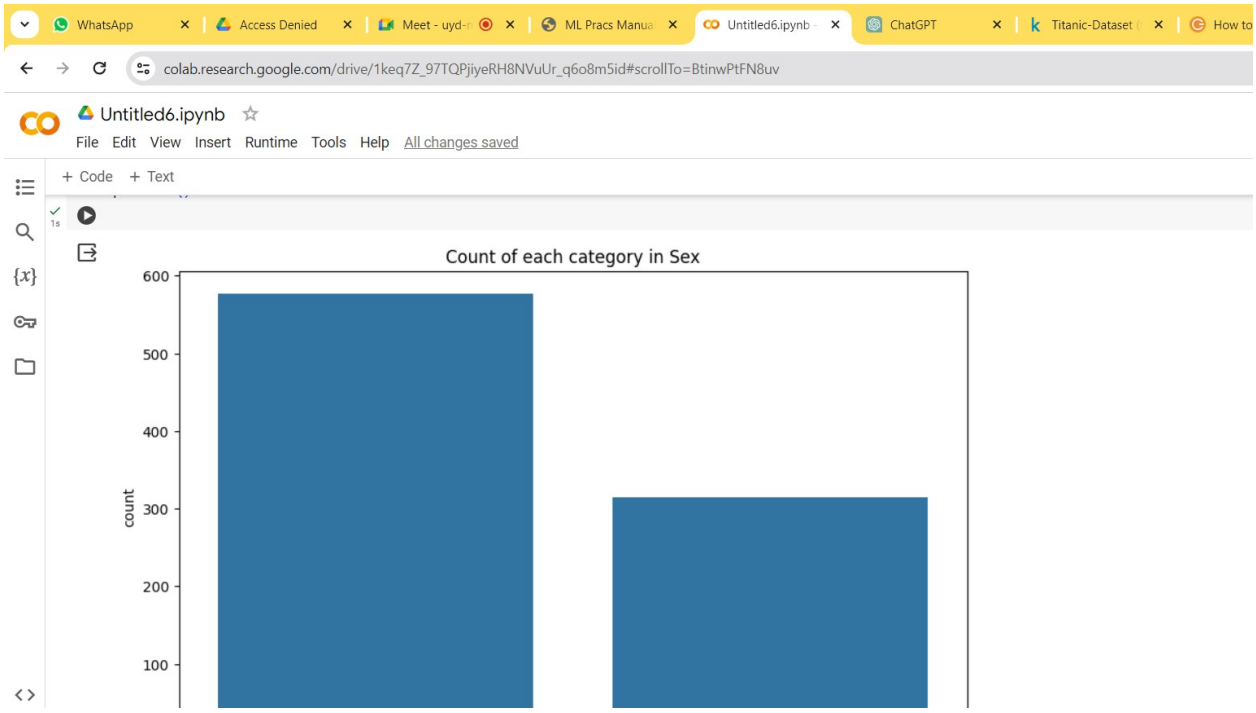

WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
6 SibSp 891 non-null int64
7 Parch 891 non-null int64
8 Ticket 891 non-null object
9 Fare 891 non-null float64
10 Cabin 204 non-null object
11 Embarked 889 non-null object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
None
PassengerId Survived Pclass Age SibSp \
count 891.000000 891.000000 891.000000 714.000000 891.000000
mean 446.000000 0.383838 2.308642 29.699118 0.523008
std 257.353842 0.486592 0.836071 14.526497 1.102743
min 1.000000 0.000000 1.000000 0.420000 0.000000
25% 223.500000 0.000000 2.000000 20.125000 0.000000
50% 446.000000 0.000000 3.000000 28.000000 0.000000
75% 668.500000 1.000000 3.000000 38.000000 1.000000
max 891.000000 1.000000 3.000000 80.000000 8.000000
Parch Fare
count 891.000000 891.000000
mean 0.381594 32.204208
std 0.806057 49.693429
min 0.000000 0.000000
25% 0.000000 7.910400
50% 0.000000 14.454200
75% 0.000000 31.000000
max 6.000000 512.329200
PassengerId 0


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
Age 177
SibSp 0
Parch 0
Ticket 0
Fare 0
Cabin 687
Embarked 2
dtype: int64
print(df.describe())
PassengerId Survived Pclass Age SibSp \
count 891.000000 891.000000 891.000000 714.000000 891.000000
mean 446.000000 0.383838 2.308642 29.699118 0.523008
std 257.353842 0.486592 0.836071 14.526497 1.102743
min 1.000000 0.000000 1.000000 0.420000 0.000000
25% 223.500000 0.000000 2.000000 20.125000 0.000000
50% 446.000000 0.000000 3.000000 28.000000 0.000000
75% 668.500000 1.000000 3.000000 38.000000 1.000000
max 891.000000 1.000000 3.000000 80.000000 8.000000
Parch Fare
count 891.000000 891.000000
mean 0.381594 32.204208
std 0.806057 49.693429
min 0.000000 0.000000
25% 0.000000 7.910400


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[17] print(df['Survived'].value_counts())
0 549
1 342
Name: Survived, dtype: int64
plt.hist(df['Age'])
plt.title('Distribution of Age')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
Distribution of Age
175
150
125
100

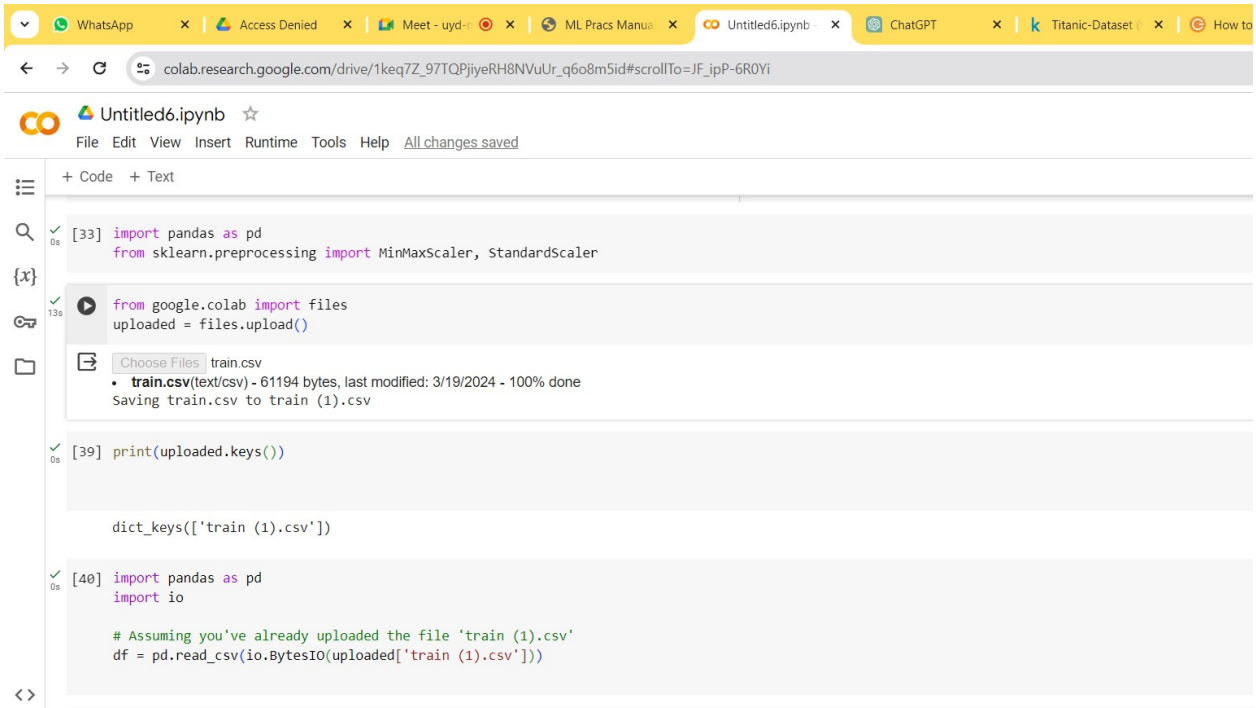

WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
plt.ylabel('Frequency')
plt.show()
Distribution of Age
Frequency
0
25
50
75
100
125
150
175
0
10
20
30
40
50
60
70
80

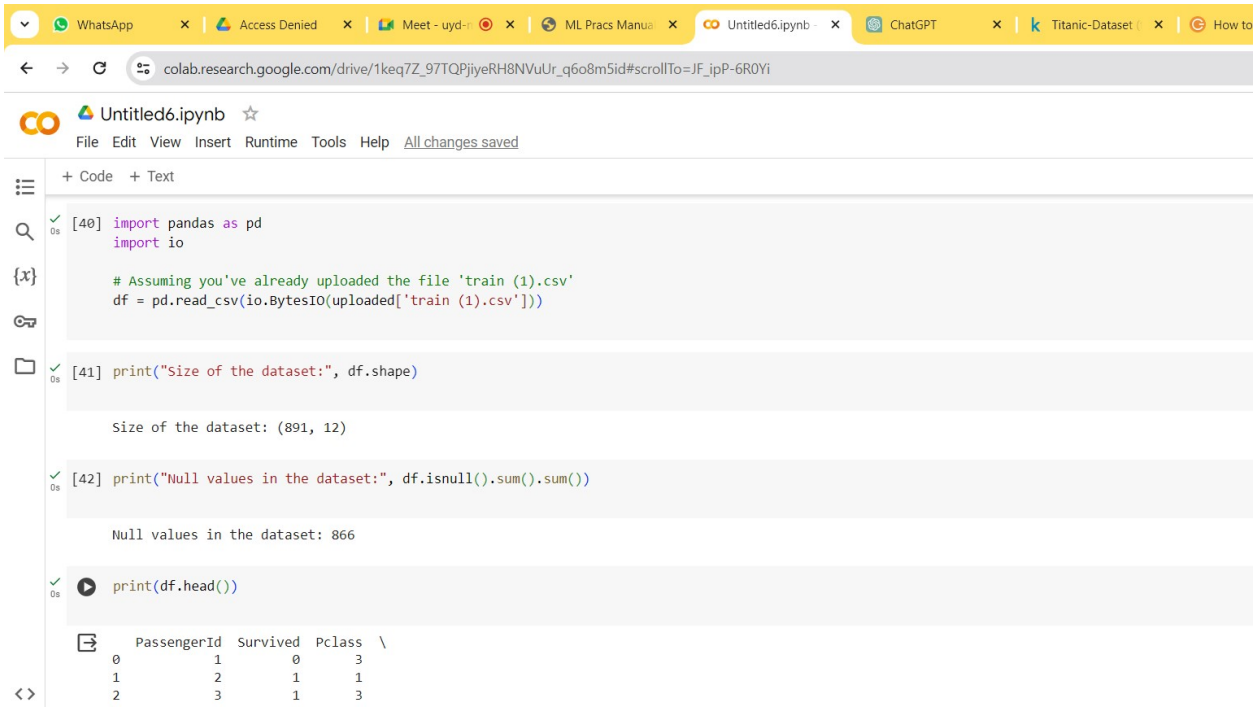

WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=BtinwPtFN8uv
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize=(10, 6))
sns.countplot(x='Sex', data=df)
plt.title('Count of each category in Sex')
plt.xticks(rotation=45)
plt.show()
Count of each category in Sex
count
300
400
500
600

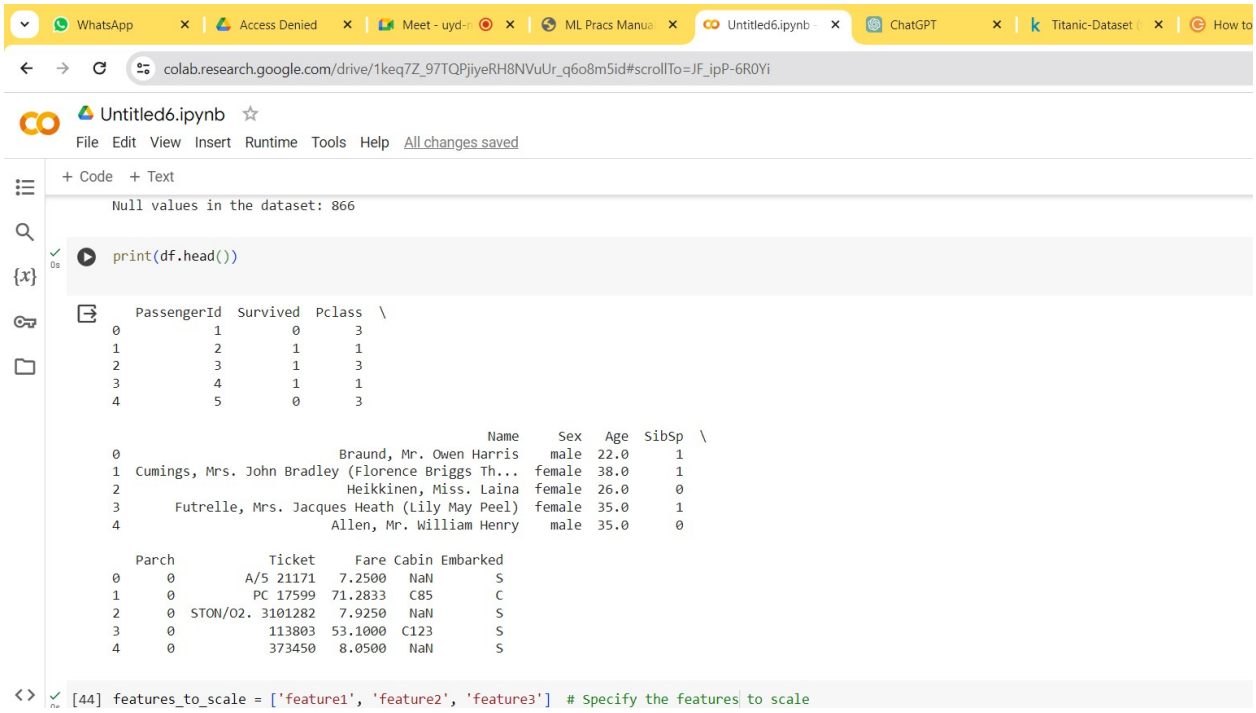


Practical No. 3

To perform Feature Scaling on the dataset


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[33] import pandas as pd
from sklearn.preprocessing import MinMaxScaler, StandardScaler
13s
from google.colab import files
uploaded = files.upload()
Choose Files train.csv
train.csv(text/csv) - 61194 bytes, last modified: 3/19/2024 - 100% done
Saving train.csv to train (1).csv
[39] print(uploaded.keys())
dict_keys(['train (1).csv'])
[40] import pandas as pd
import io
# Assuming you've already uploaded the file 'train (1).csv'
df = pd.read_csv(io.BytesIO(uploaded['train (1).csv']))


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[40] import pandas as pd
import io
# Assuming you've already uploaded the file 'train (1).csv'
df = pd.read_csv(io.BytesIO(uploaded['train (1).csv']))
[41] print("Size of the dataset:", df.shape)
Size of the dataset: (891, 12)
[42] print("Null values in the dataset:", df.isnull().sum().sum())
Null values in the dataset: 866
print(df.head())
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3



WhatsApp | Access Denied | Meet - uyd-n | ML Pracs Manua | Untitled6.ipynb | ChatGPT | Titanic-Dataset | How to

colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi

Untitled6.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
Null values in the dataset: 866
```

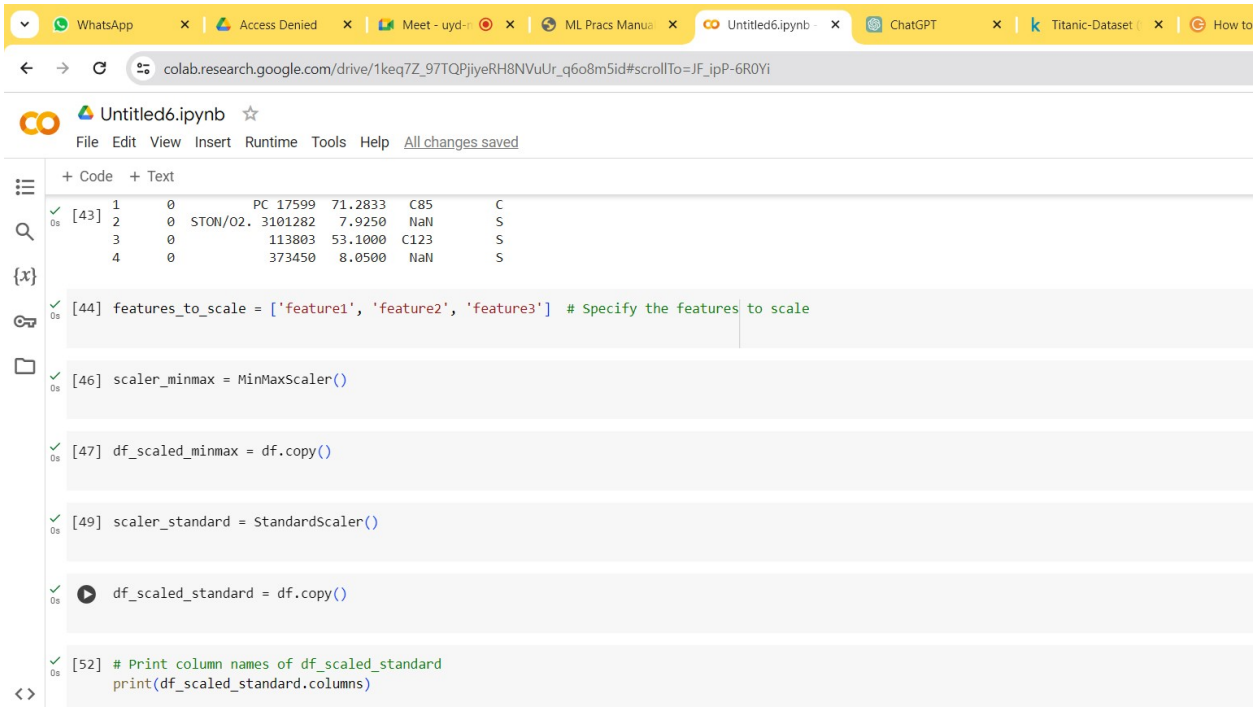
```
print(df.head())
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0            113803  53.1000  C123        S
4      0            373450   8.0500   NaN        S
```

```
[44] features_to_scale = ['feature1', 'feature2', 'feature3']  # Specify the features to scale
```



WhatsApp | Access Denied | Meet - uyd-n | ML Pracs Manua | Untitled6.ipynb | ChatGPT | Titanic-Dataset | How to

colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi

Untitled6.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[43]  1      0          PC 17599  71.2833   C85        C
      2      0  STON/O2. 3101282   7.9250   NaN        S
      3      0            113803  53.1000  C123        S
      4      0            373450   8.0500   NaN        S
```

```
[44] features_to_scale = ['feature1', 'feature2', 'feature3']  # Specify the features to scale
```

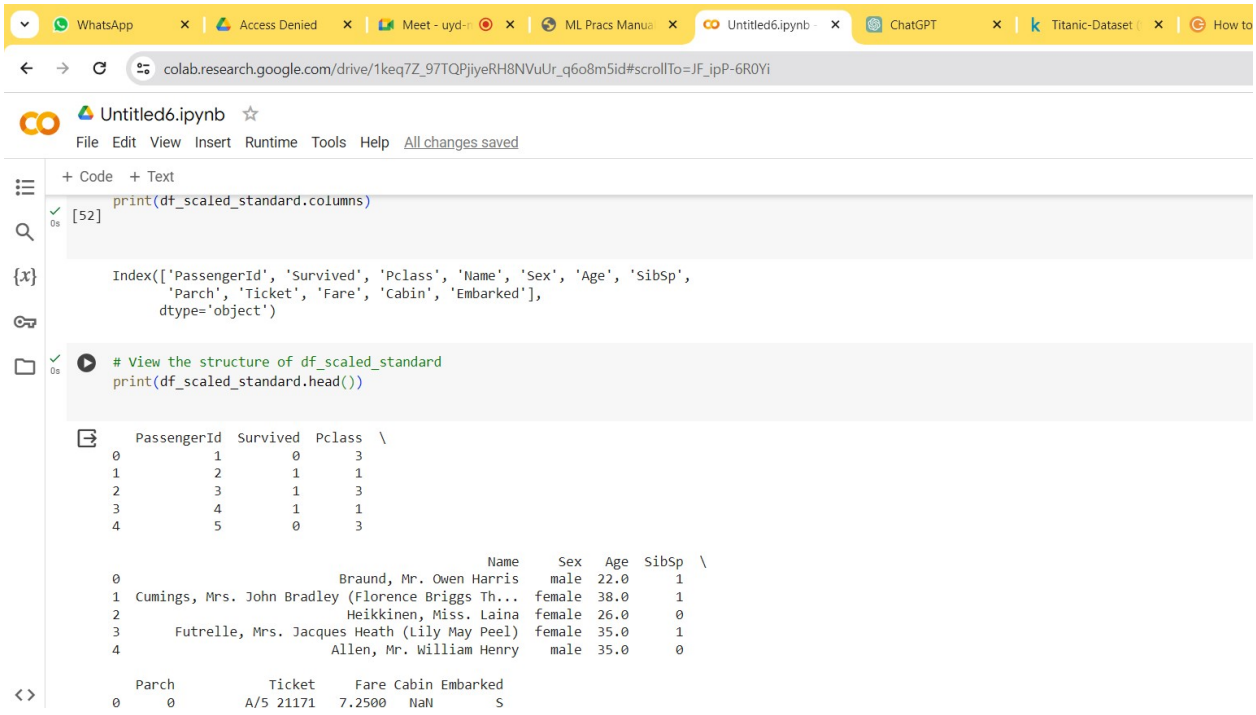
```
[46] scaler_minmax = MinMaxScaler()
```

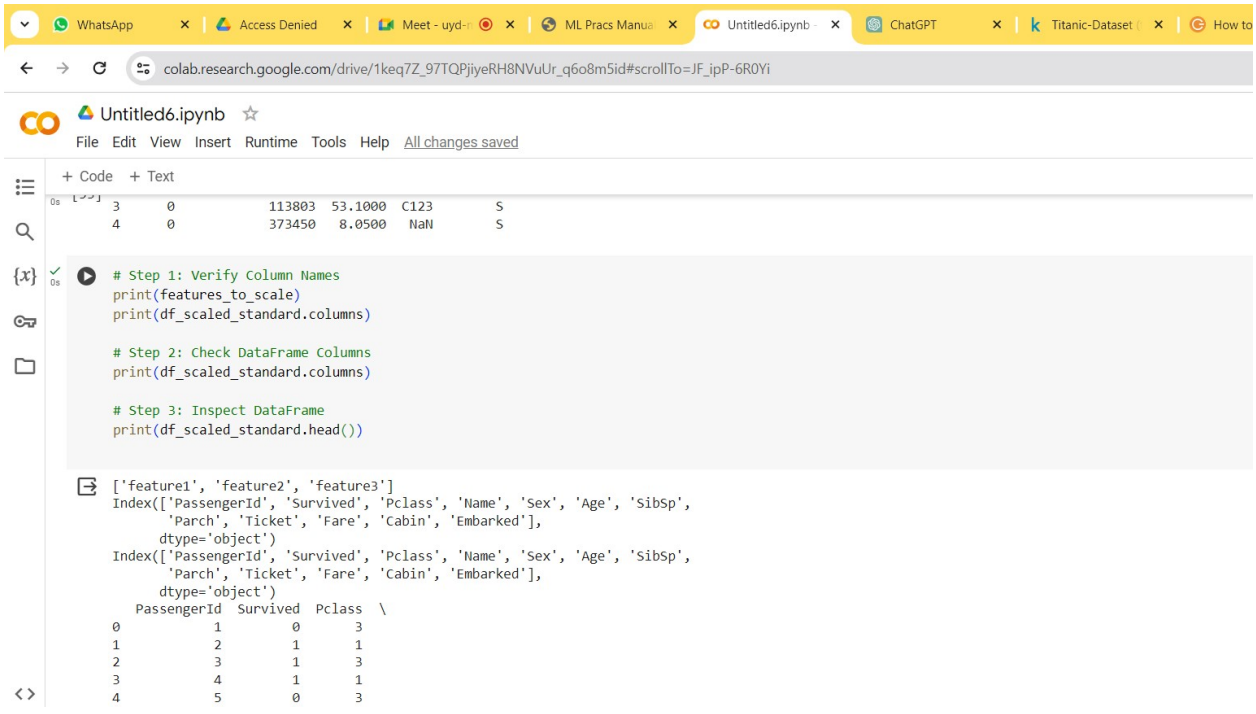
```
[47] df_scaled_minmax = df.copy()
```

```
[49] scaler_standard = StandardScaler()
```

```
df_scaled_standard = df.copy()
```

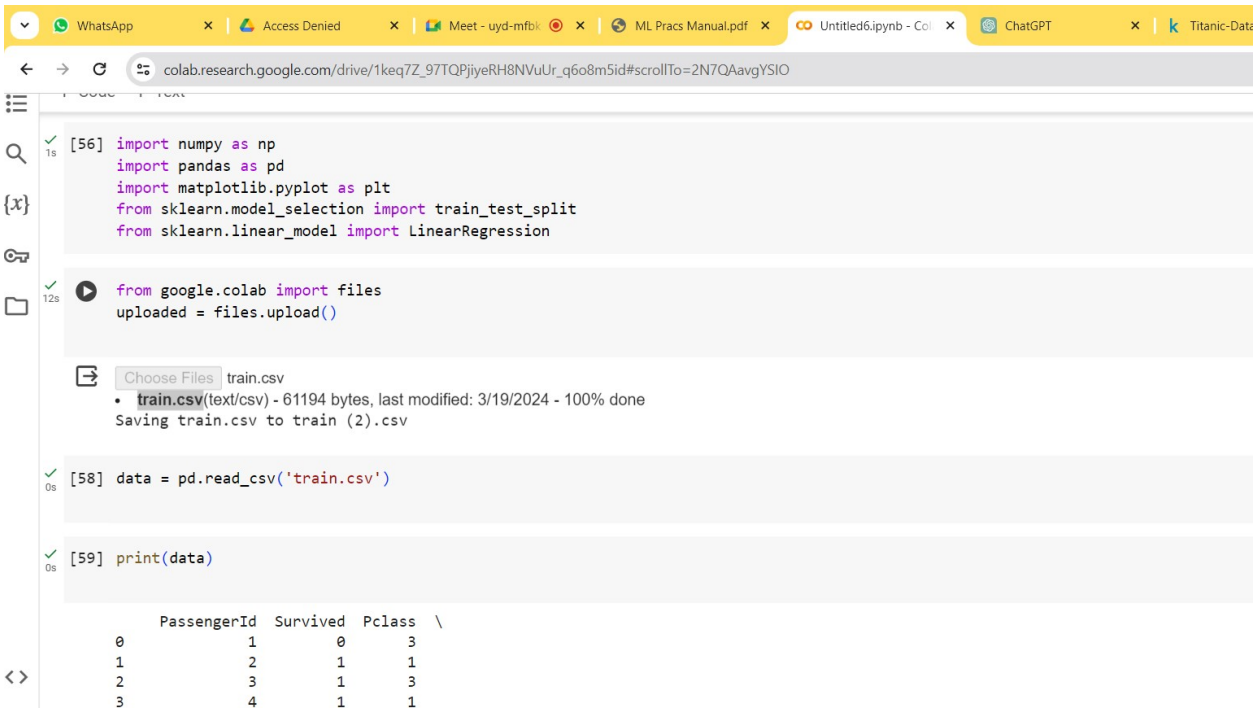
```
[52] # Print column names of df_scaled_standard
     print(df_scaled_standard.columns)
```


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[52]
print(df_scaled_standard.columns)
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
dtype='object')
# View the structure of df_scaled_standard
print(df_scaled_standard.head())
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3
3 4 1 1
4 5 0 3
Name Sex Age SibSp \
0 Braund, Mr. Owen Harris male 22.0 1
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
2 Heikkinen, Miss. Laina female 26.0 0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1
4 Allen, Mr. William Henry male 35.0 0
Parch Ticket Fare Cabin Embarked
0 0 A/5 21171 7.2500 NaN S


WhatsApp
Access Denied
Meet - uyd-
ML Pracs Manua
Untitled6.ipynb
ChatGPT
Titanic-Dataset
How to
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=JF_ipP-6R0Yi
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
3 0 113803 53.1000 C123 S
4 0 373450 8.0500 NaN S
# Step 1: Verify Column Names
print(features_to_scale)
print(df_scaled_standard.columns)
# Step 2: Check DataFrame Columns
print(df_scaled_standard.columns)
# Step 3: Inspect DataFrame
print(df_scaled_standard.head())
['feature1', 'feature2', 'feature3']
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
dtype='object')
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
dtype='object')
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3
3 4 1 1
4 5 0 3

Practical No. 4

To perform linear regression using google



```
[59]
886      0           211536  13.0000   NaN        S
887      0           112053  30.0000   B42        S
888      2       W./C. 6607  23.4500   NaN        S
889      0           111369  30.0000  C148        C
890      0           370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

```
# Check if column names exist in data
print("Column names in data:", data.columns)

# Print the first few rows of data to inspect its structure
print(data.head())
```

```
Column names in data: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
       'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
```

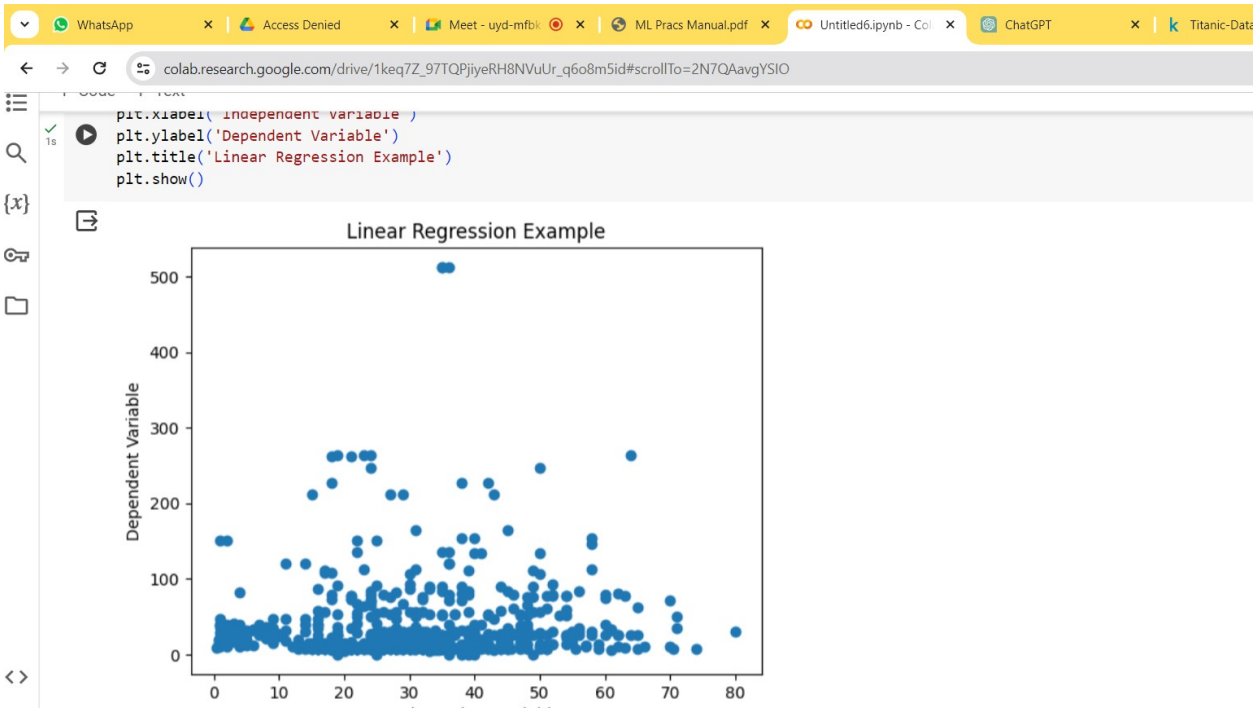
WhatsApp | Access Denied | Meet - uyd-mfbk | ML Pracs Manual.pdf | Untitled6.ipynb - Col | ChatGPT | Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=2N7QAavgYSIO

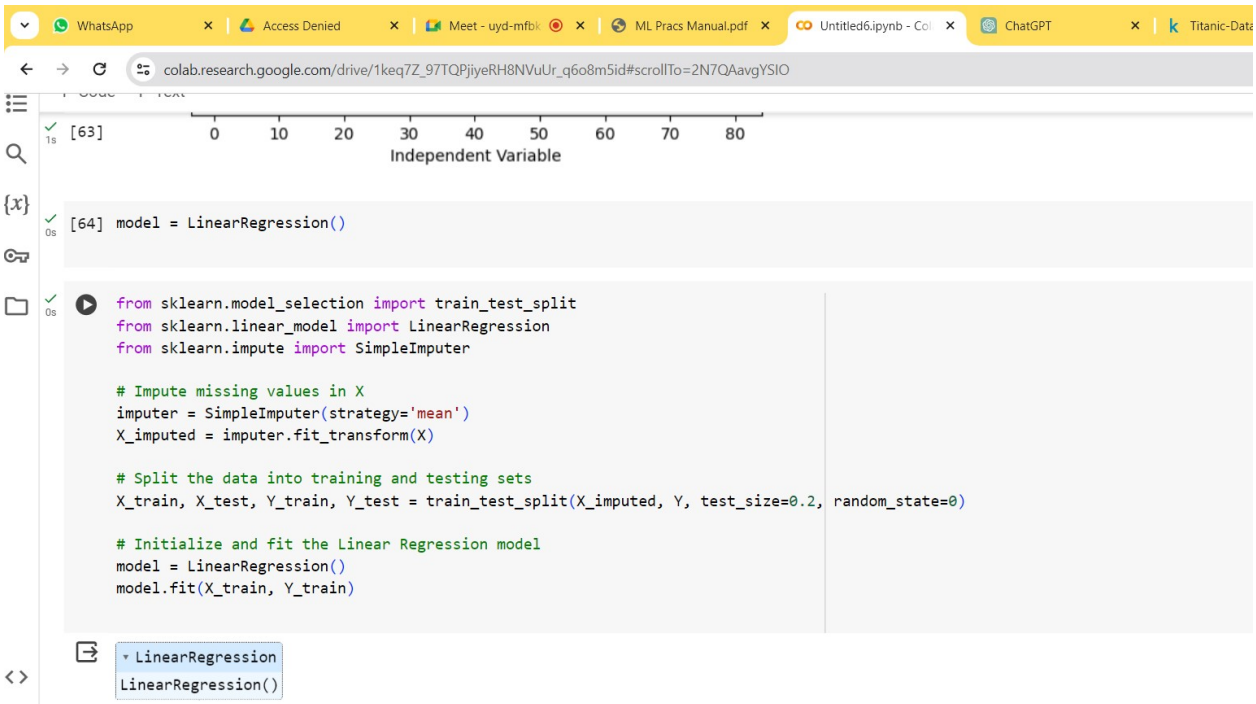
```
[61]
                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

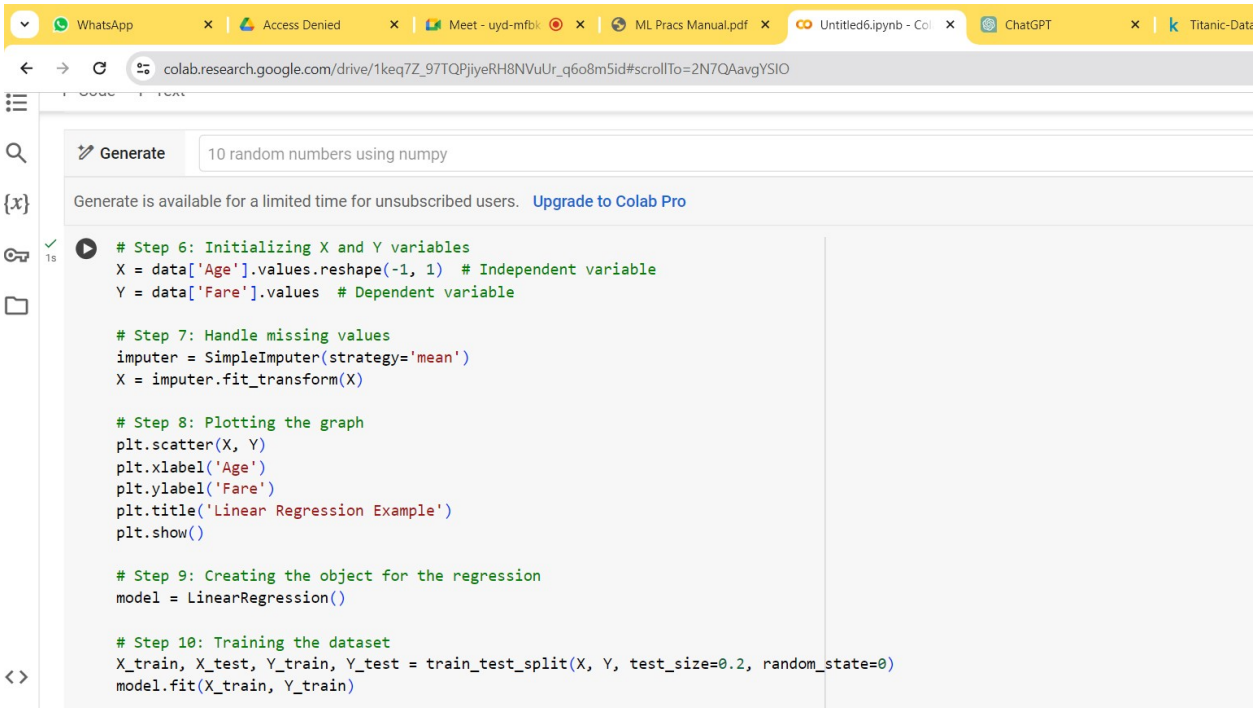
   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0            113803  53.1000  C123        S
4      0            373450   8.0500   NaN        S
```

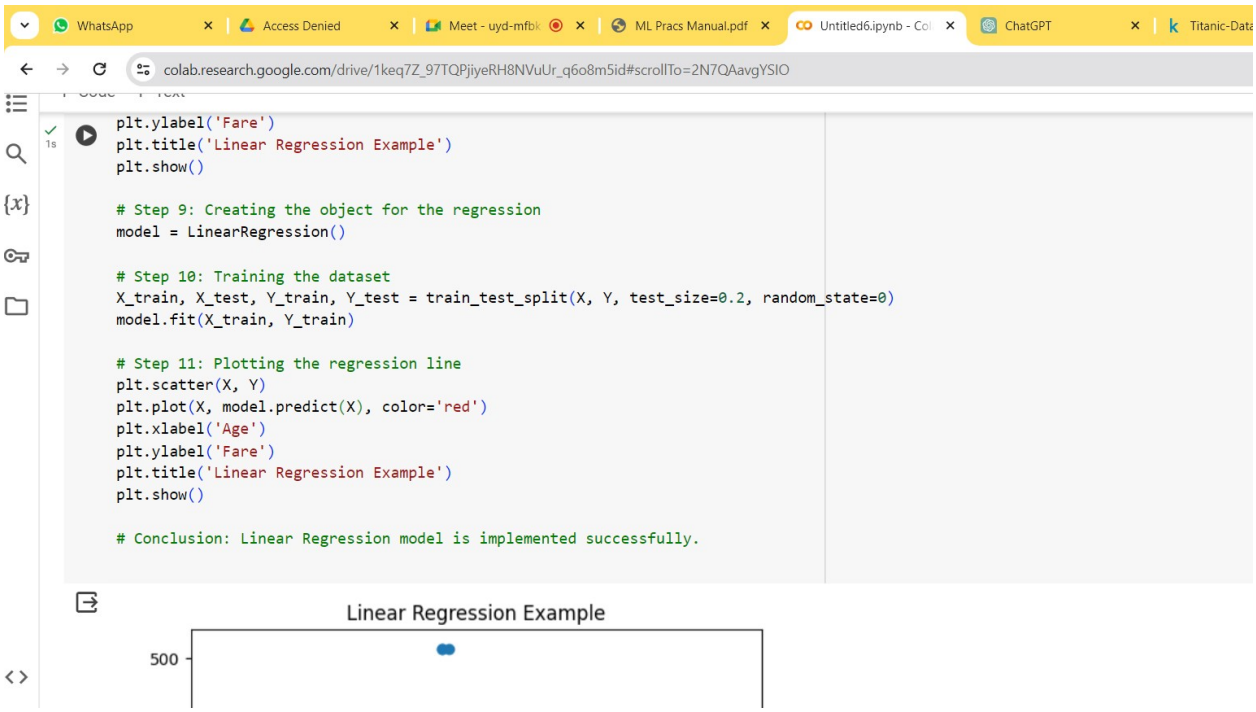
```
[62] X = data['Age'].values.reshape(-1, 1)
     Y = data['Fare'].values
```

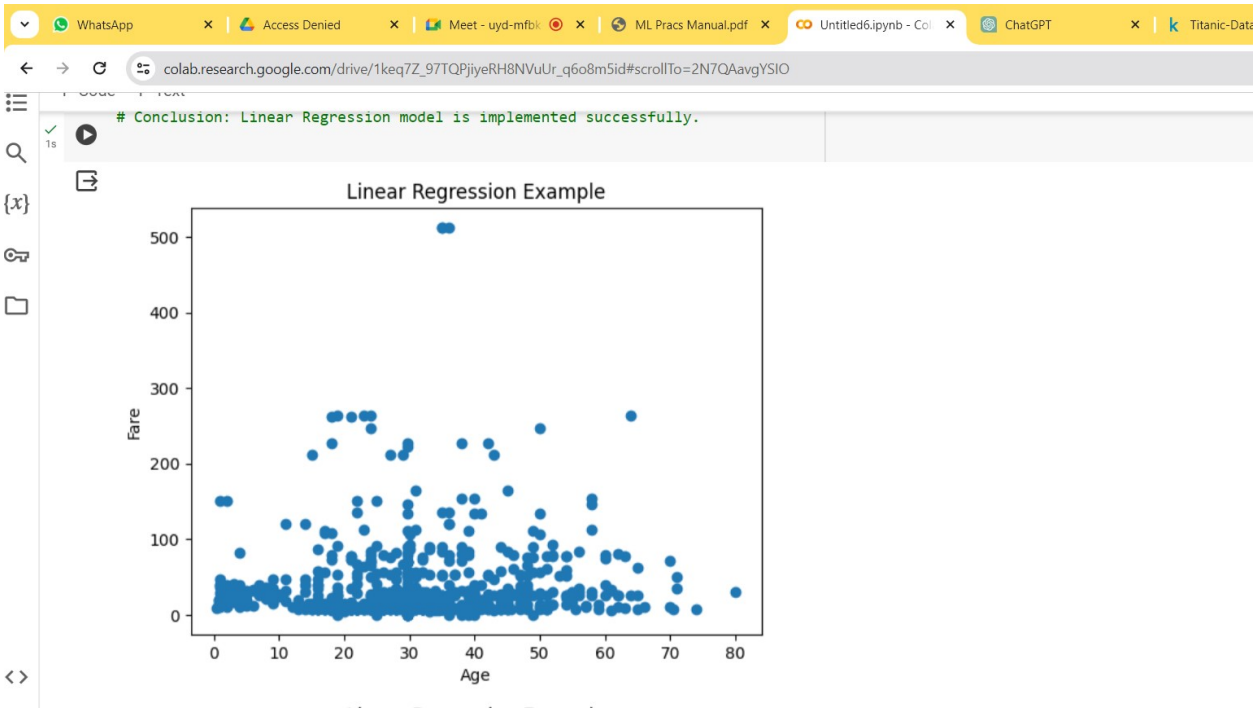
```
plt.scatter(X, Y)
plt.xlabel('Independent Variable')
plt.ylabel('Dependent Variable')
plt.title('Linear Regression Example')
plt.show()
```

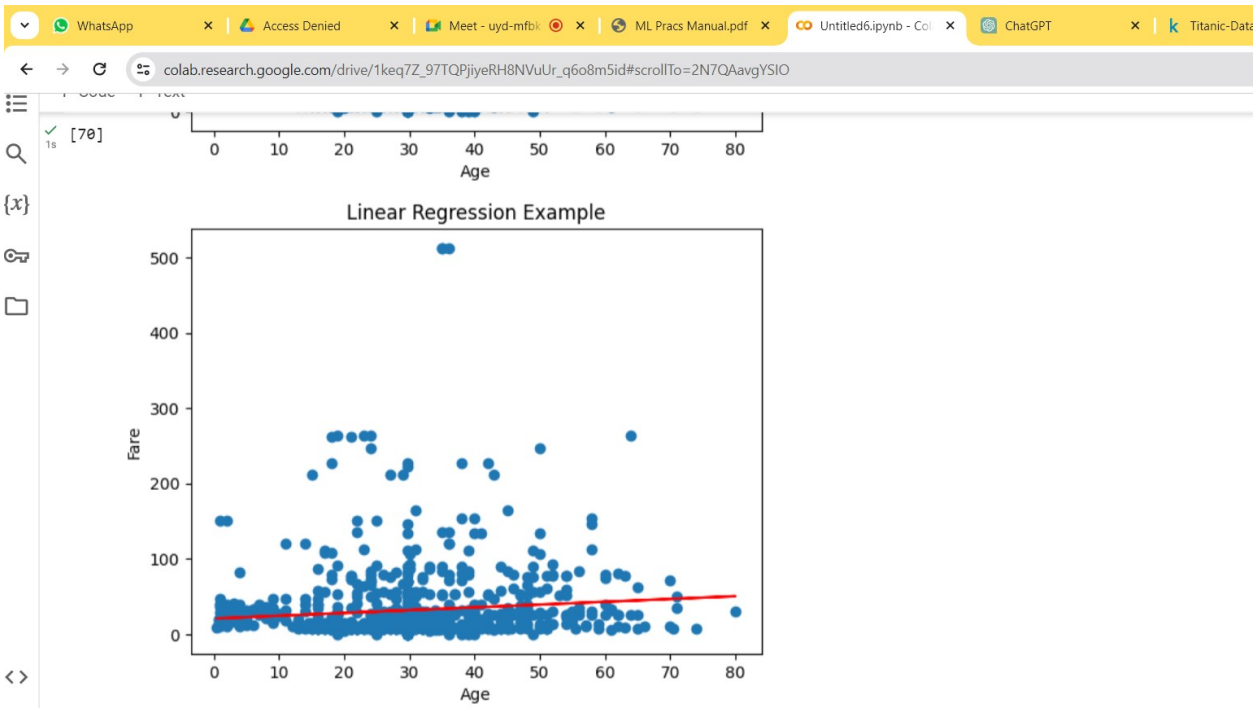

WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=2N7QAavgYSIO
plt.xlabel('Independent Variable')
plt.ylabel('Dependent Variable')
plt.title('Linear Regression Example')
plt.show()
Linear Regression Example
Dependent Variable
0
100
200
300
400
500
0
10
20
30
40
50
60
70
80


WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=2N7QAavgYSIO
[63]
0
10
20
30
40
50
60
70
80
Independent Variable
[64] model = LinearRegression()
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.impute import SimpleImputer
# Impute missing values in X
imputer = SimpleImputer(strategy='mean')
X_imputed = imputer.fit_transform(X)
# Split the data into training and testing sets
X_train, X_test, Y_train, Y_test = train_test_split(X_imputed, Y, test_size=0.2, random_state=0)
# Initialize and fit the Linear Regression model
model = LinearRegression()
model.fit(X_train, Y_train)
LinearRegression
LinearRegression()


WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=2N7QAavgYSIO
Generate
10 random numbers using numpy
Generate is available for a limited time for unsubscribed users.
Upgrade to Colab Pro
# Step 6: Initializing X and Y variables
X = data['Age'].values.reshape(-1, 1) # Independent variable
Y = data['Fare'].values # Dependent variable
# Step 7: Handle missing values
imputer = SimpleImputer(strategy='mean')
X = imputer.fit_transform(X)
# Step 8: Plotting the graph
plt.scatter(X, Y)
plt.xlabel('Age')
plt.ylabel('Fare')
plt.title('Linear Regression Example')
plt.show()
# Step 9: Creating the object for the regression
model = LinearRegression()
# Step 10: Training the dataset
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=0)
model.fit(X_train, Y_train)

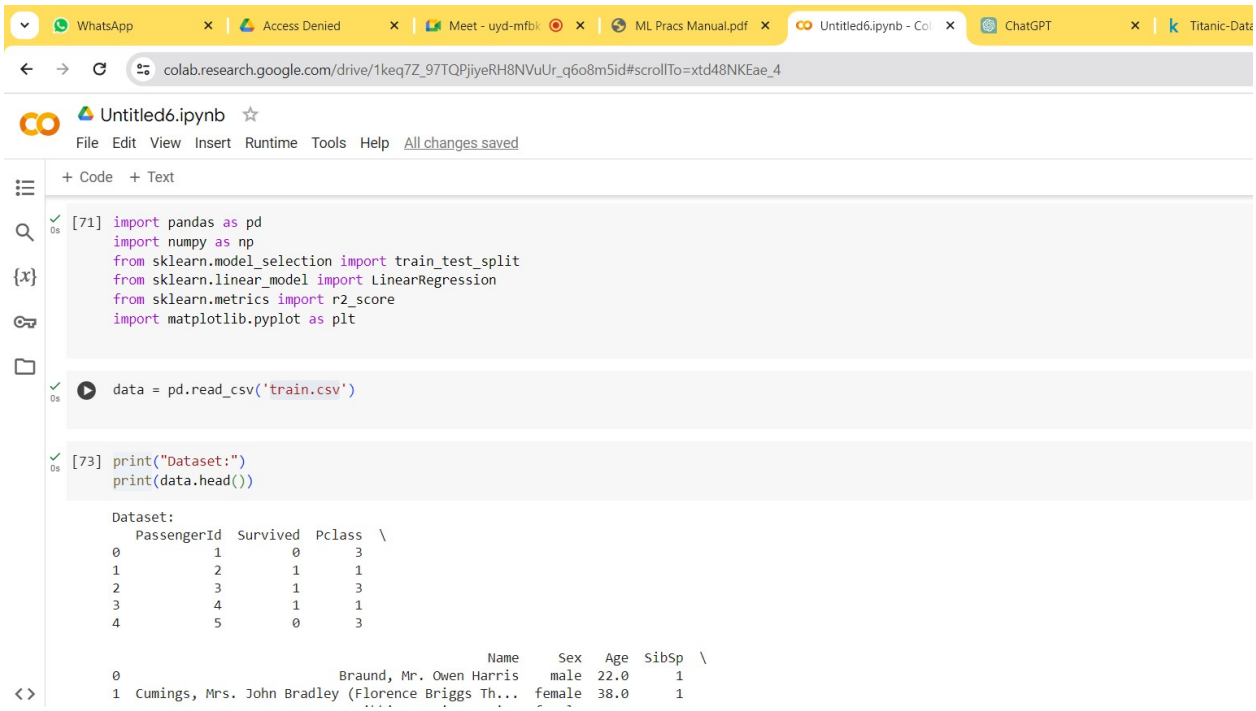

WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=2N7QAavgYSIO
plt.ylabel('Fare')
plt.title('Linear Regression Example')
plt.show()
# Step 9: Creating the object for the regression
model = LinearRegression()
# Step 10: Training the dataset
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=0)
model.fit(X_train, Y_train)
# Step 11: Plotting the regression line
plt.scatter(X, Y)
plt.plot(X, model.predict(X), color='red')
plt.xlabel('Age')
plt.ylabel('Fare')
plt.title('Linear Regression Example')
plt.show()
# Conclusion: Linear Regression model is implemented successfully.
Linear Regression Example
500

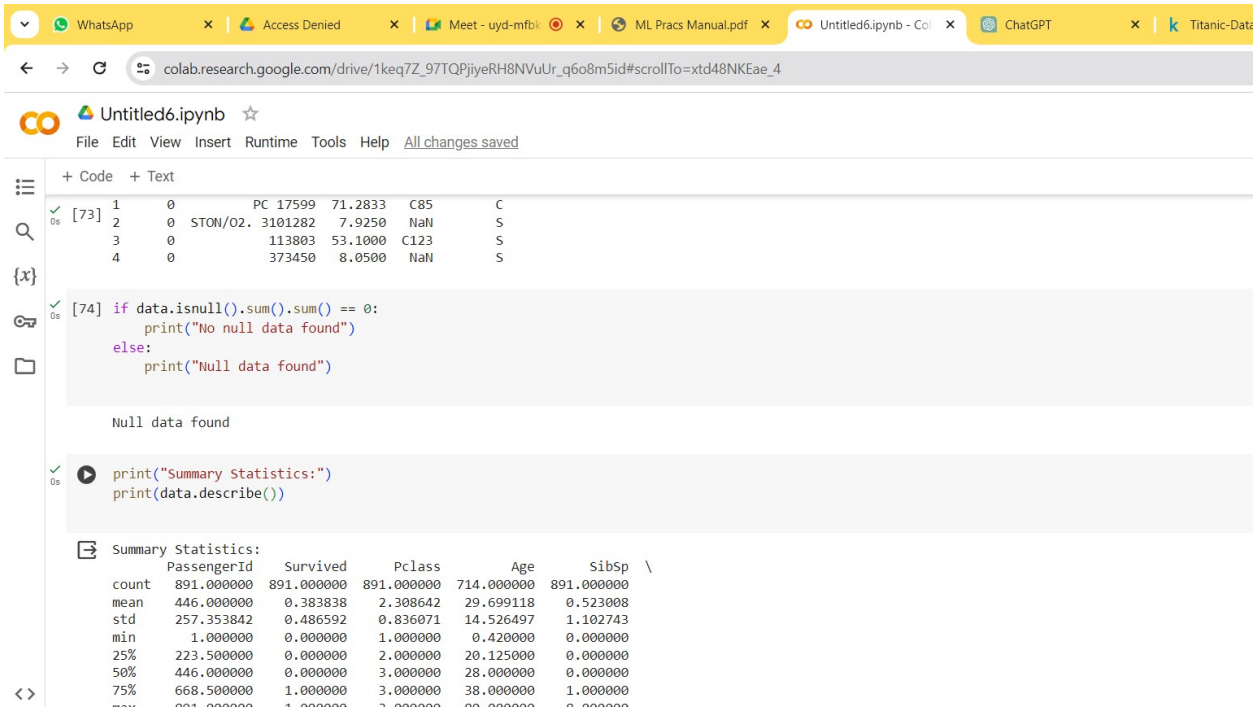


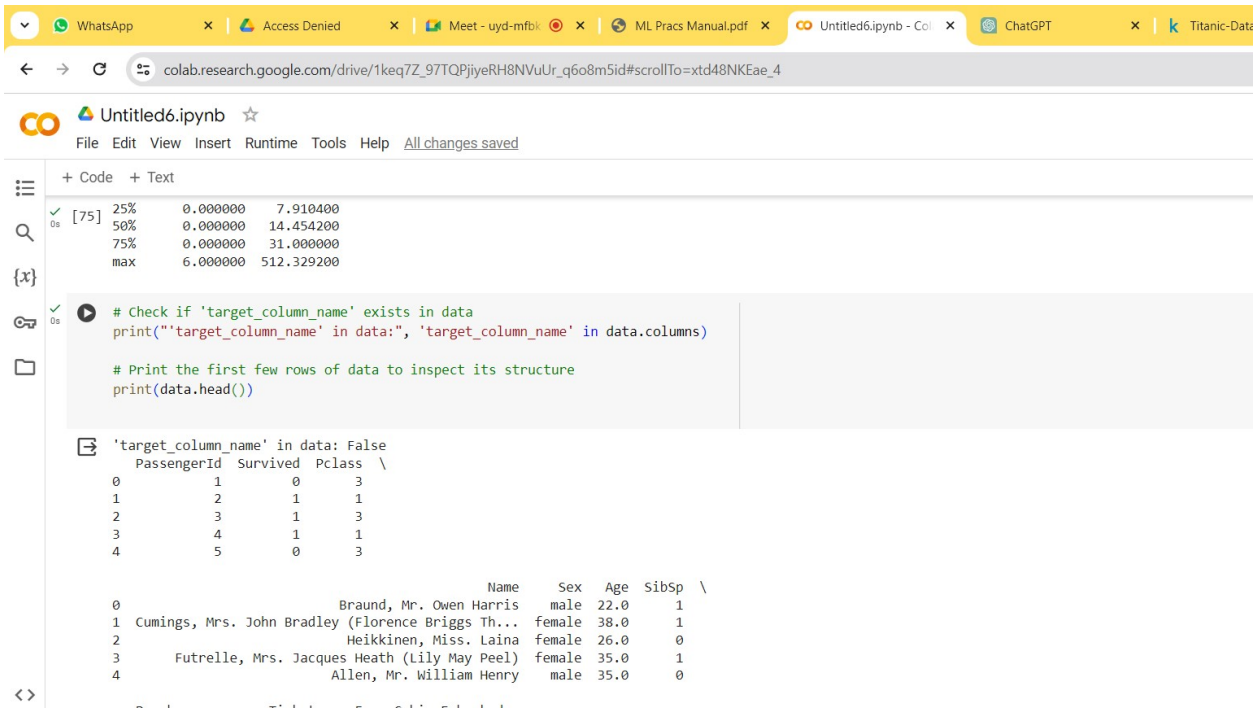


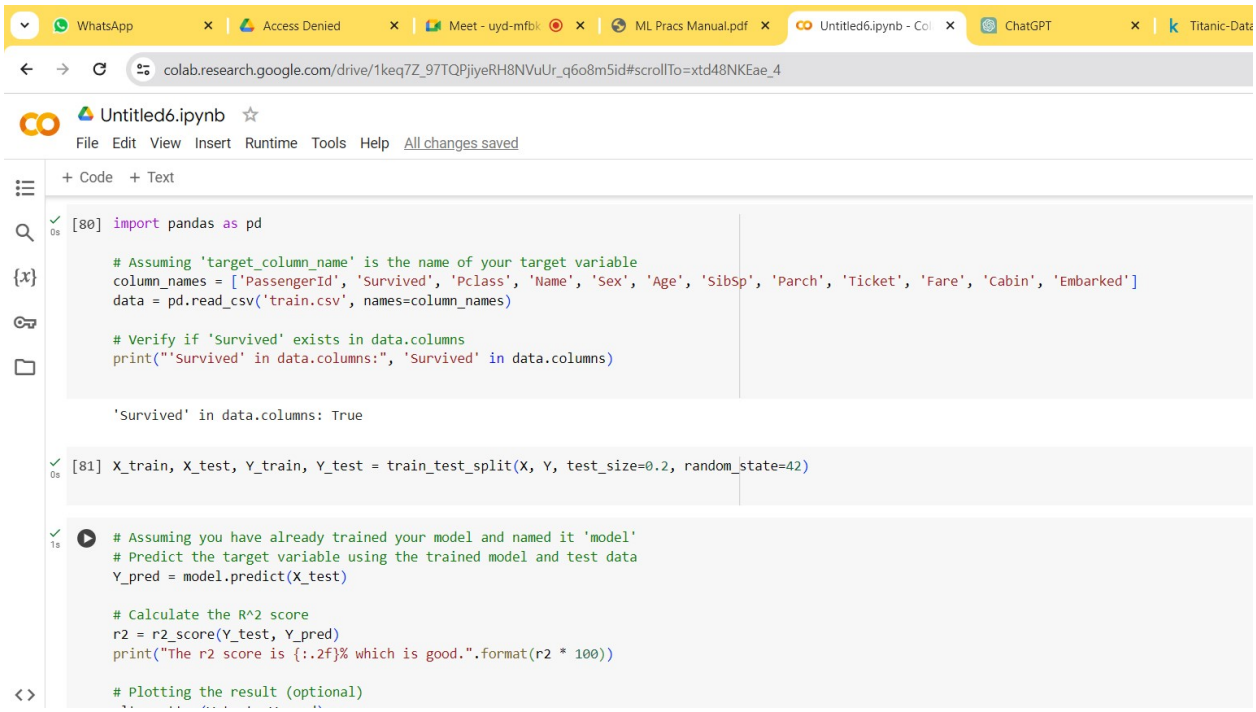
Practical No. 5

To perform Multiple Linear Regression using google colab



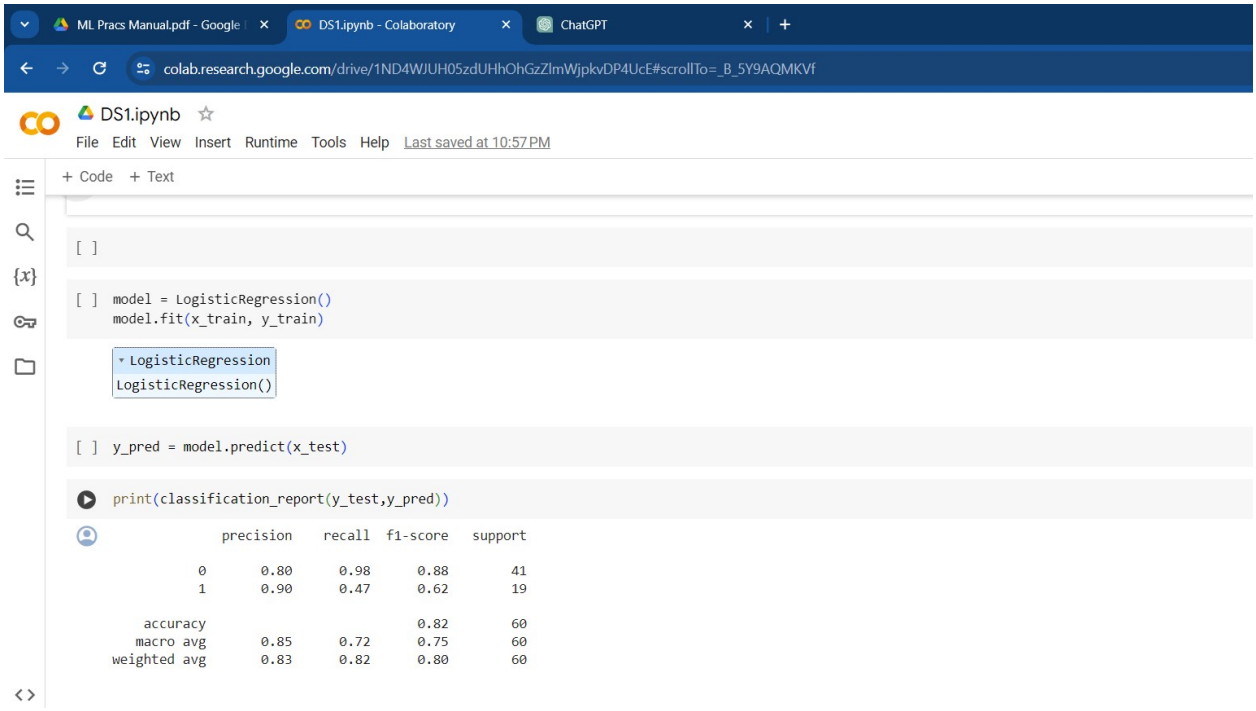


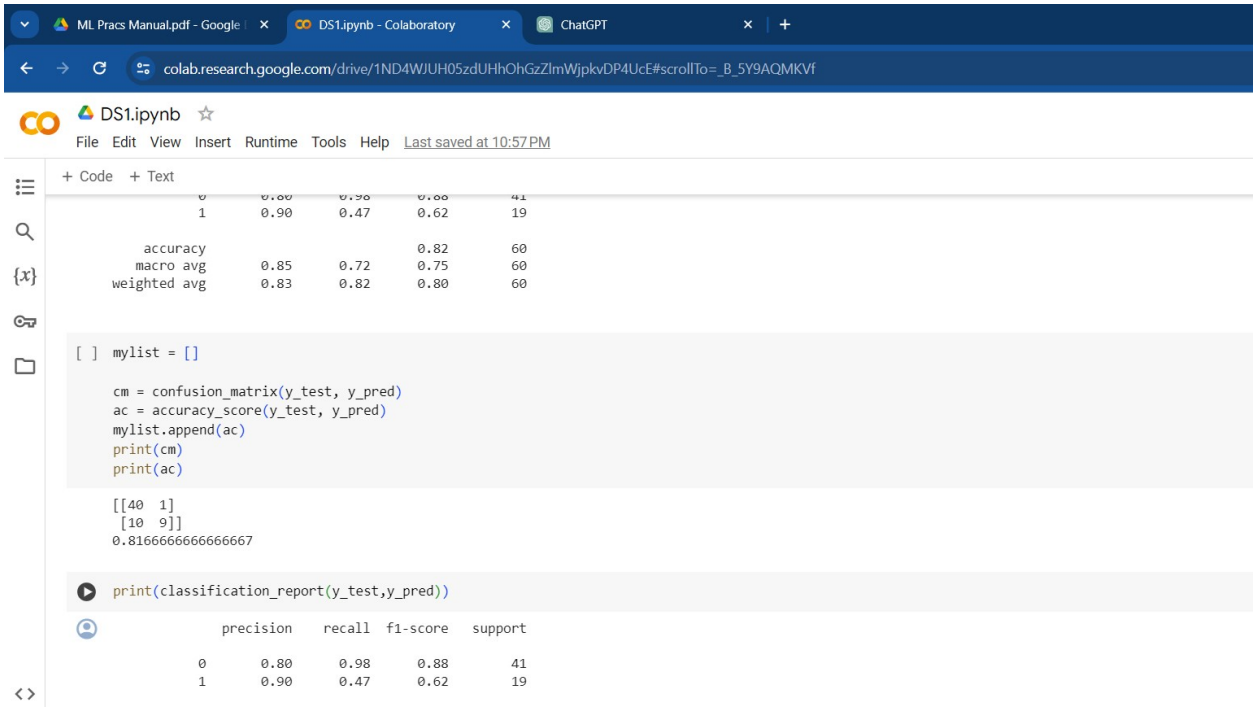

WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=xtd48NKEae_4
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[75] 25% 0.000000 7.910400
50% 0.000000 14.454200
75% 0.000000 31.000000
max 6.000000 512.329200
# Check if 'target_column_name' exists in data
print("'target_column_name' in data:", 'target_column_name' in data.columns)
# Print the first few rows of data to inspect its structure
print(data.head())
'target_column_name' in data: False
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3
3 4 1 1
4 5 0 3
Name Sex Age SibSp \
0 Braund, Mr. Owen Harris male 22.0 1
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
2 Heikkinen, Miss. Laina female 26.0 0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1
4 Allen, Mr. William Henry male 35.0 0

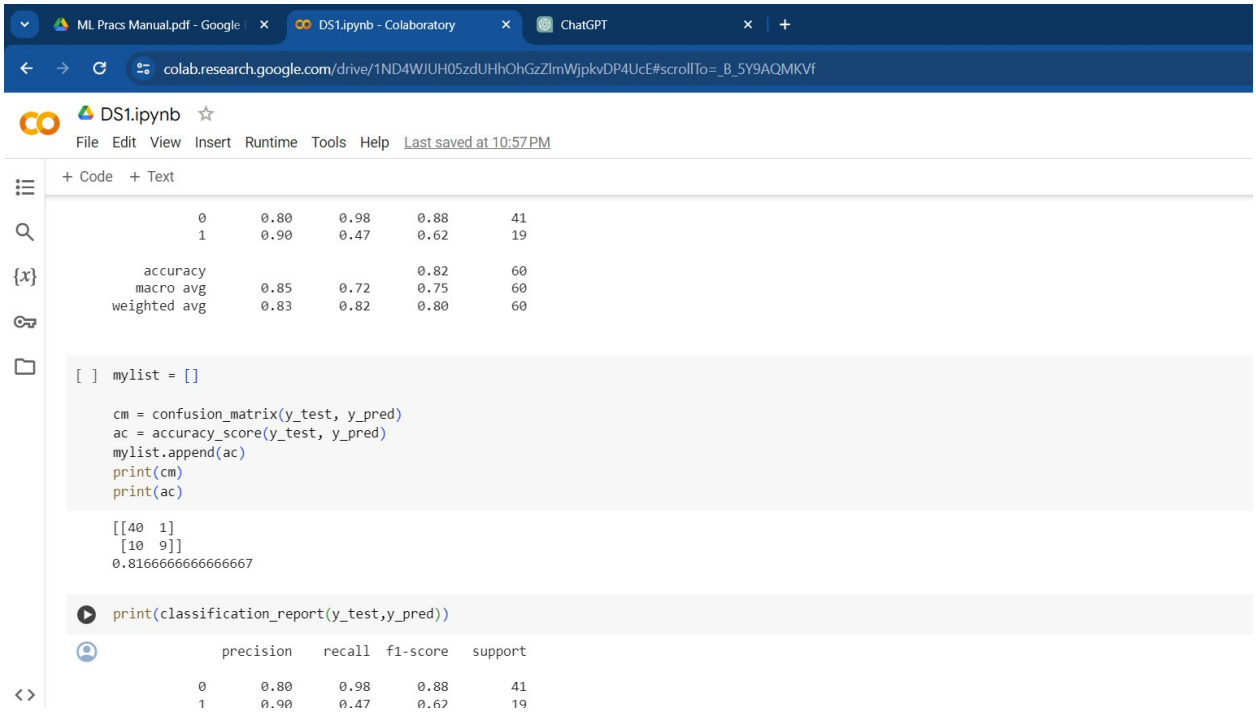

WhatsApp
Access Denied
Meet - uyd-mfbk
ML Pracs Manual.pdf
Untitled6.ipynb - Col
ChatGPT
Titanic-Data
colab.research.google.com/drive/1keq7Z_97TQPjiyeRH8NVuUr_q6o8m5id#scrollTo=xtd48NKEae_4
Untitled6.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[80] import pandas as pd
# Assuming 'target_column_name' is the name of your target variable
column_names = ['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp', 'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked']
data = pd.read_csv('train.csv', names=column_names)
# Verify if 'Survived' exists in data.columns
print("'Survived' in data.columns:", 'Survived' in data.columns)
'Survived' in data.columns: True
[81] X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)
# Assuming you have already trained your model and named it 'model'
# Predict the target variable using the trained model and test data
Y_pred = model.predict(X_test)
# Calculate the R^2 score
r2 = r2_score(Y_test, Y_pred)
print("The r2 score is {:.2f}% which is good.".format(r2 * 100))
# Plotting the result (optional)

Practical No. 6

To perform Multiple Logistic Regression using google colab

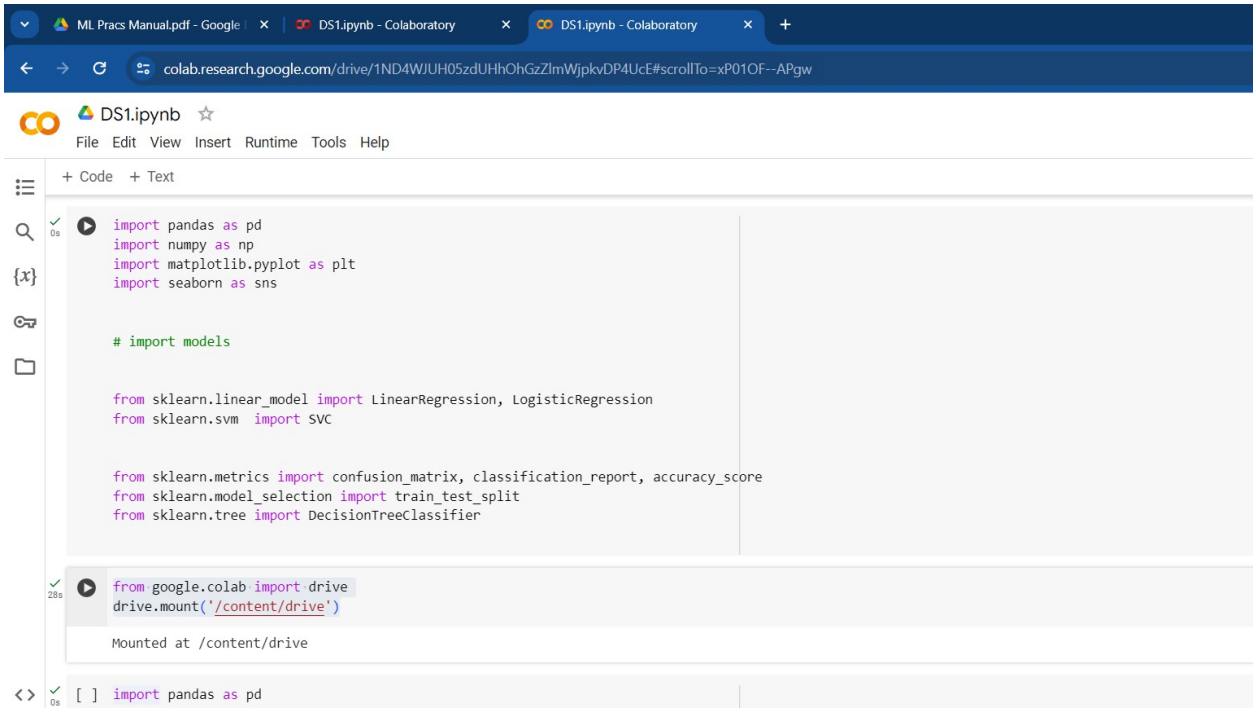

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=_B_5Y9AQMKVf
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
[ ]
[ ] model = LogisticRegression()
model.fit(x_train, y_train)
LogisticRegression
LogisticRegression()
[ ] y_pred = model.predict(x_test)
print(classification_report(y_test,y_pred))
precision recall f1-score support
0 0.80 0.98 0.88 41
1 0.90 0.47 0.62 19
accuracy 0.82 60
macro avg 0.85 0.72 0.75 60
weighted avg 0.83 0.82 0.80 60

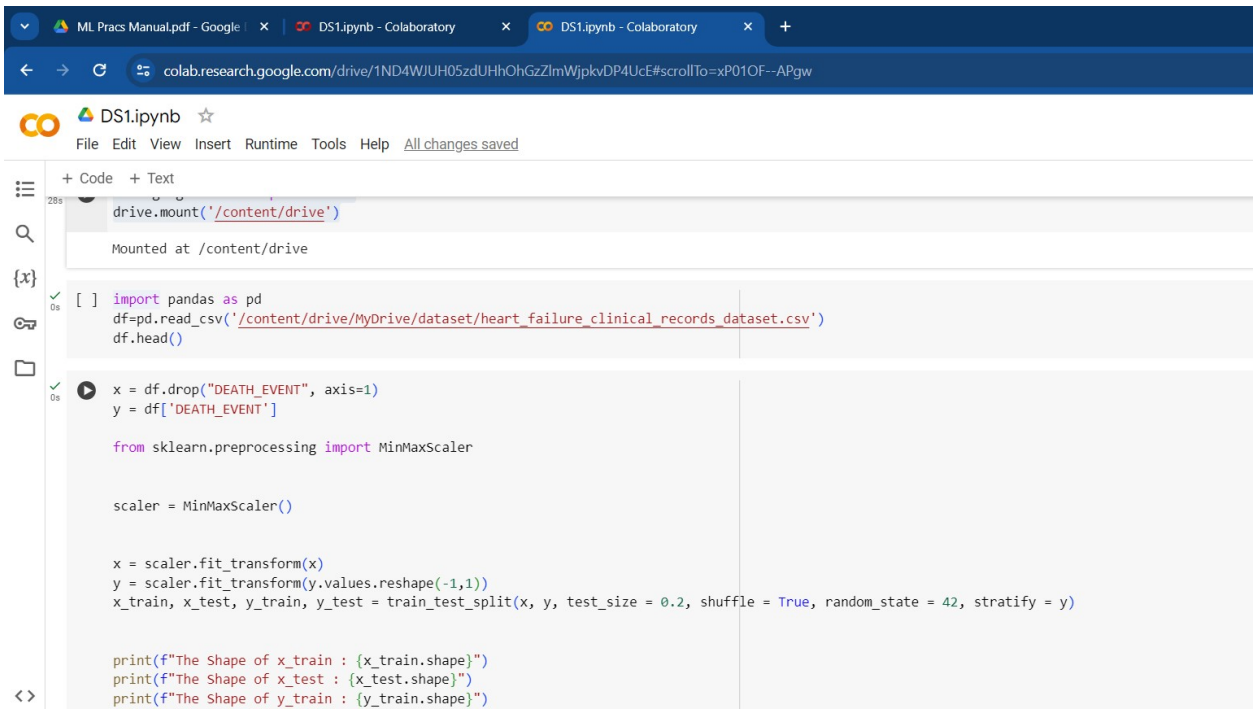

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=_B_5Y9AQMKVf
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
1 0.90 0.47 0.62 19
accuracy 0.82 60
macro avg 0.85 0.72 0.75 60
weighted avg 0.83 0.82 0.80 60
[ ] mylist = []
cm = confusion_matrix(y_test, y_pred)
ac = accuracy_score(y_test, y_pred)
mylist.append(ac)
print(cm)
print(ac)
[[40 1]
[10 9]]
0.8166666666666667
print(classification_report(y_test,y_pred))
precision recall f1-score support
0 0.80 0.98 0.88 41
1 0.90 0.47 0.62 19

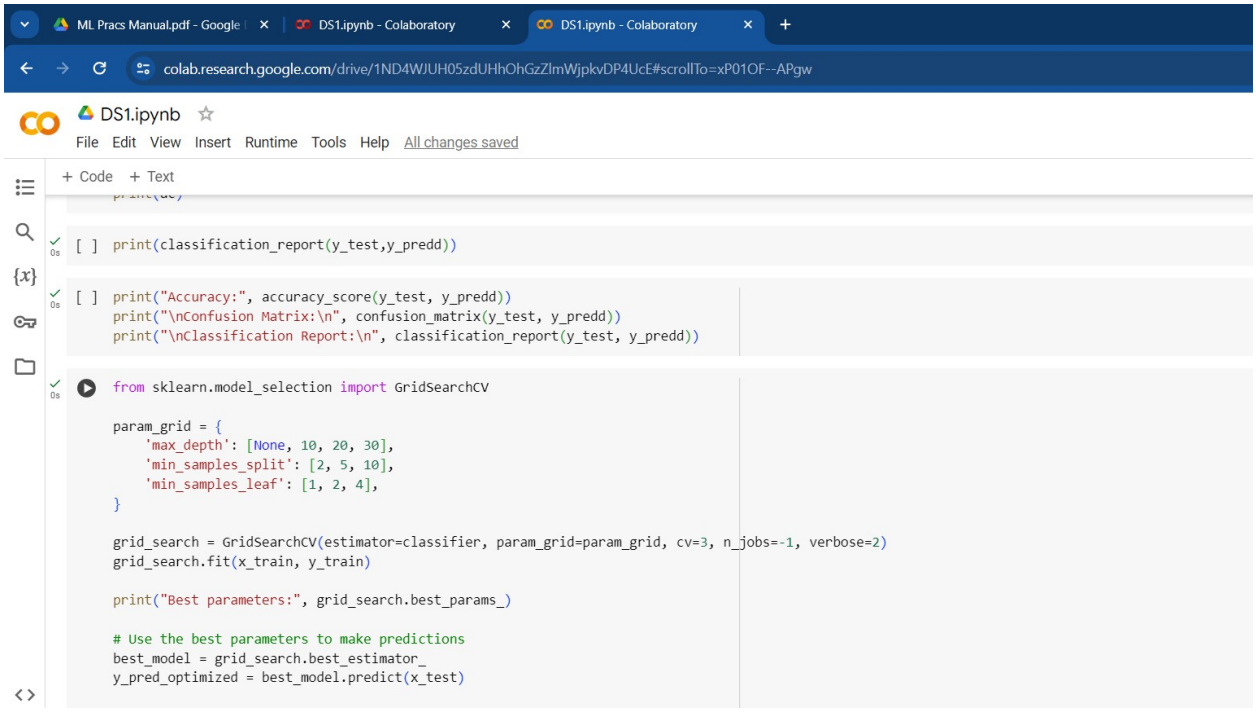


Practical No. 7

To build Decision Tree Model using google

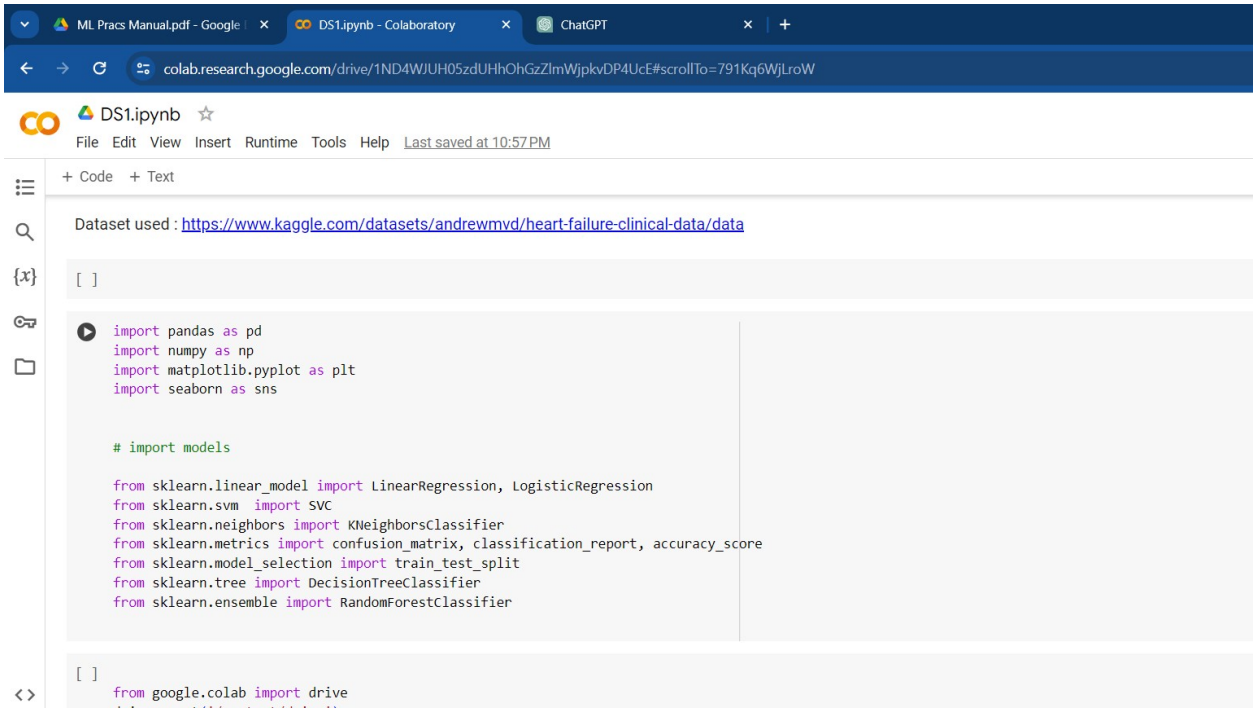

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
DS1.ipynb - Colaboratory
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=xP01OF--APgw
DS1.ipynb
File Edit View Insert Runtime Tools Help
+ Code + Text
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# import models
from sklearn.linear_model import LinearRegression, LogisticRegression
from sklearn.svm import SVC
from sklearn.metrics import confusion_matrix, classification_report, accuracy_score
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from google.colab import drive
drive.mount('/content/drive')
Mounted at /content/drive
[ ] import pandas as pd

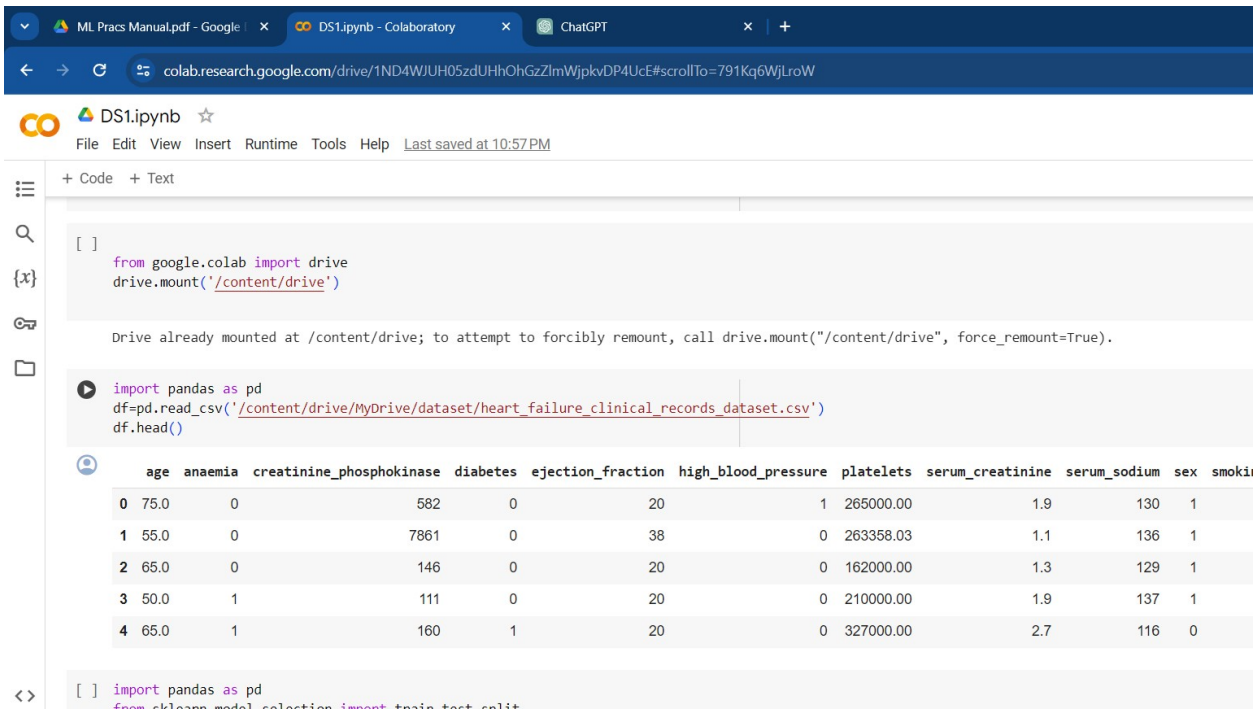

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
DS1.ipynb - Colaboratory
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=xP01OF--APgw
DS1.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
drive.mount('/content/drive')
Mounted at /content/drive
[ ] import pandas as pd
df=pd.read_csv('/content/drive/MyDrive/dataset/heart_failure_clinical_records_dataset.csv')
df.head()
x = df.drop("DEATH_EVENT", axis=1)
y = df['DEATH_EVENT']
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
x = scaler.fit_transform(x)
y = scaler.fit_transform(y.values.reshape(-1,1))
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, shuffle = True, random_state = 42, stratify = y)
print(f"The Shape of x_train : {x_train.shape}")
print(f"The Shape of x_test : {x_test.shape}")
print(f"The Shape of y_train : {y_train.shape}")



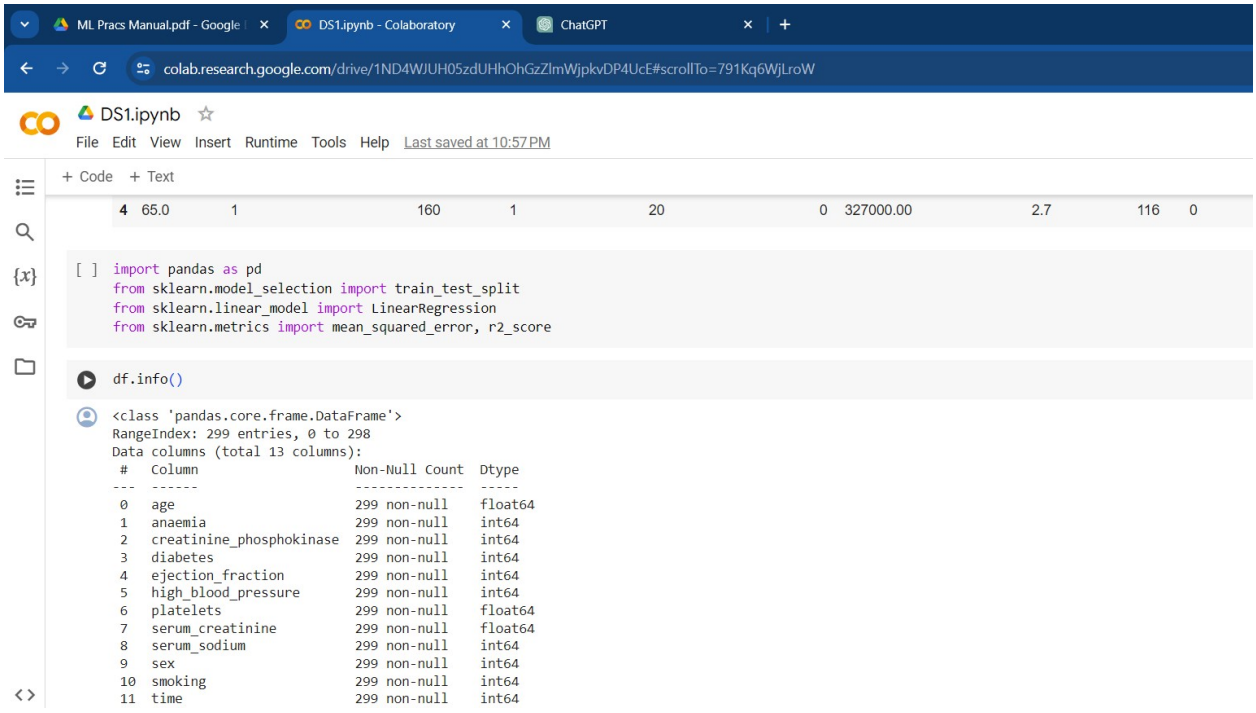
Practical No. 8

To build Suport Vector Machine model (SVM) using google colab



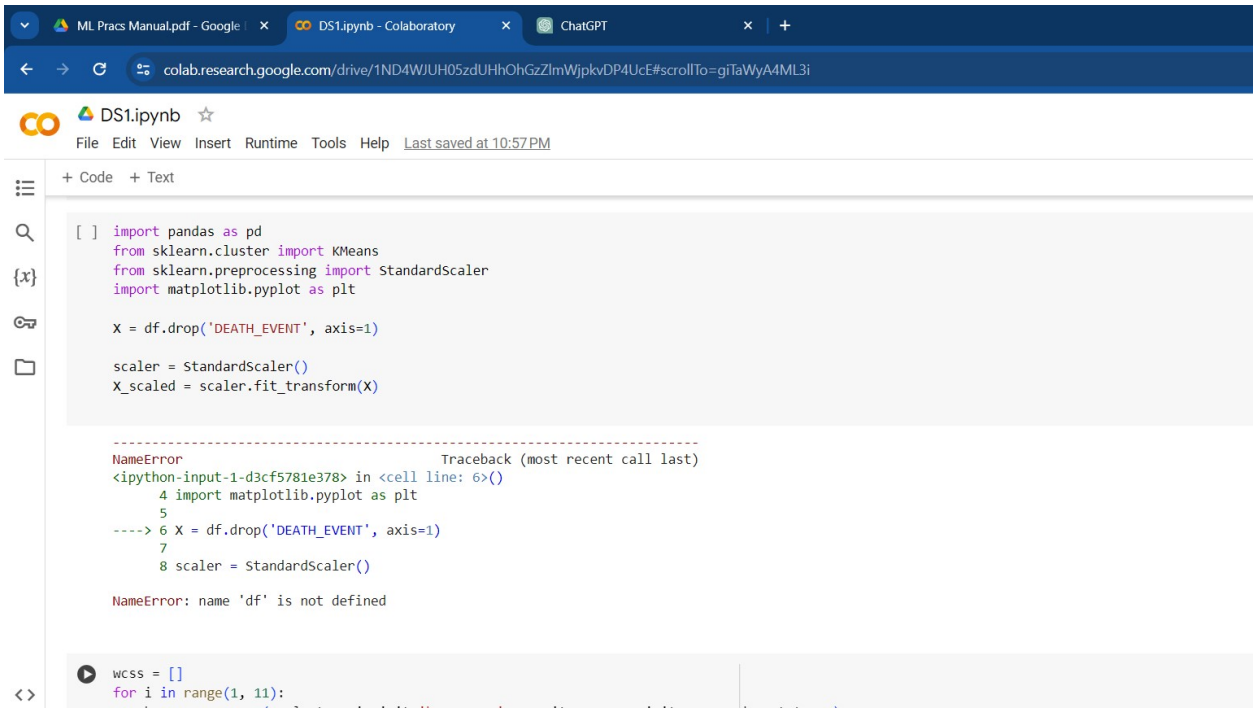


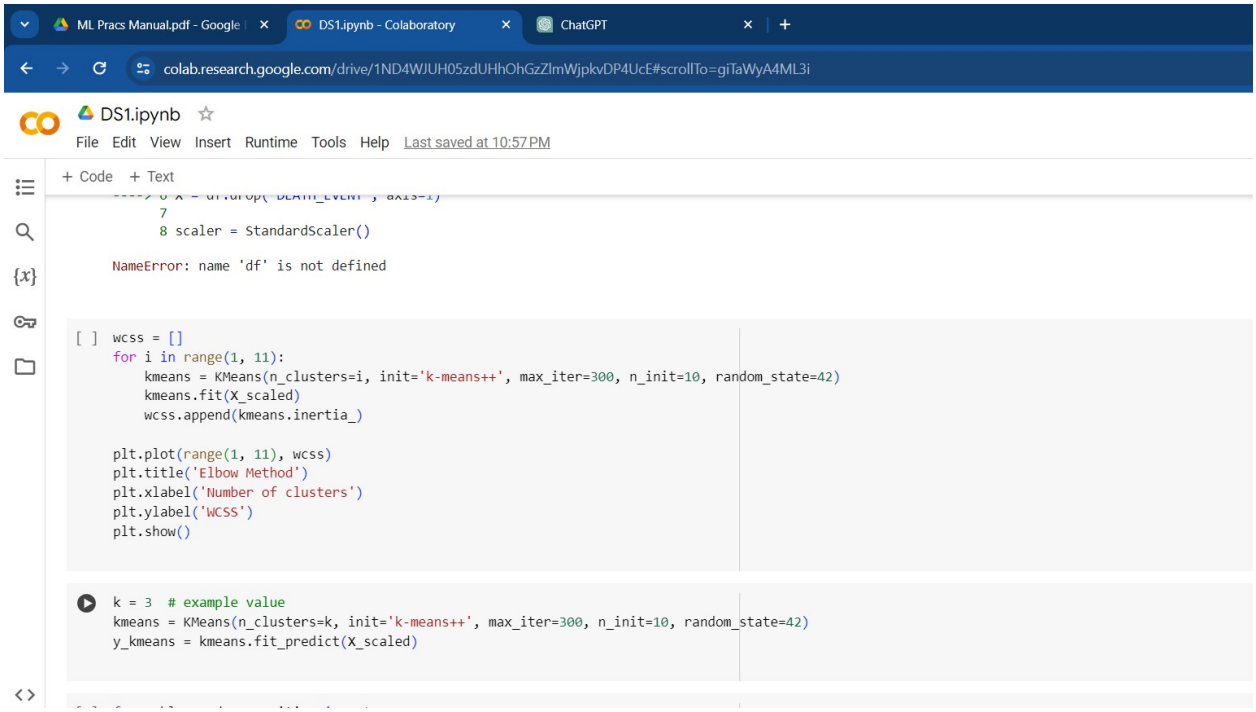
	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoki
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	

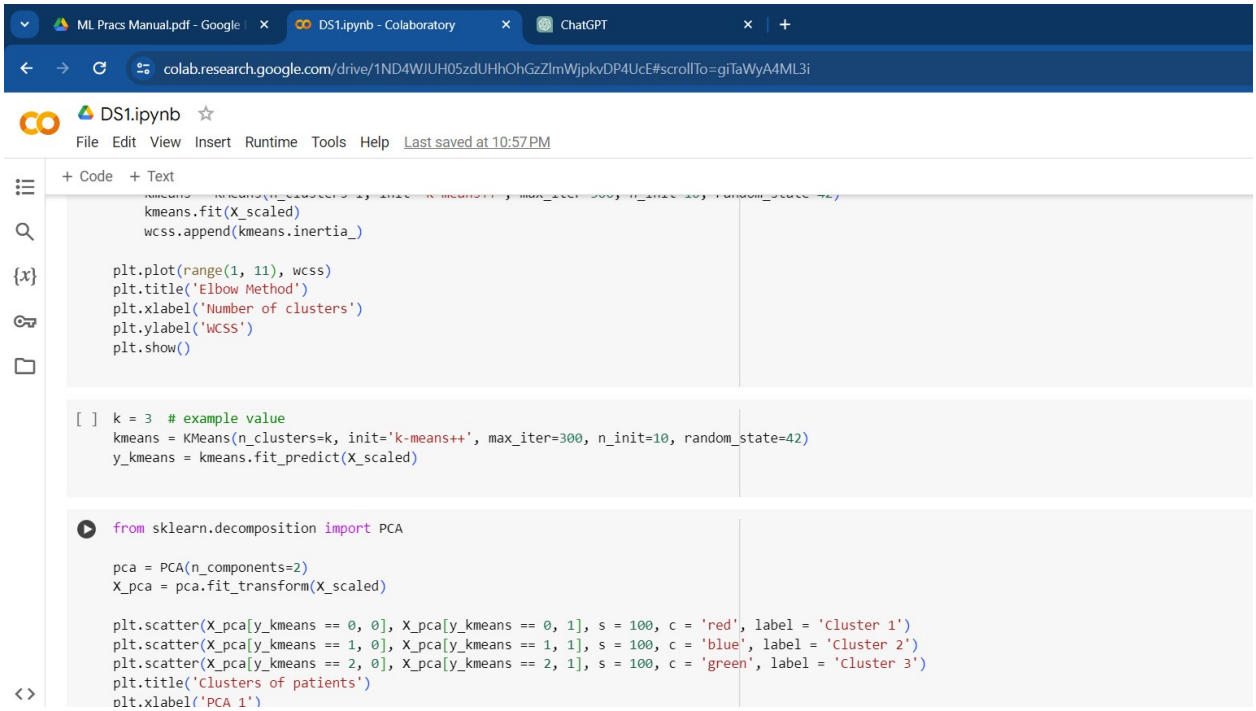


Practical No. 9

perform clustering using KMeans clustering

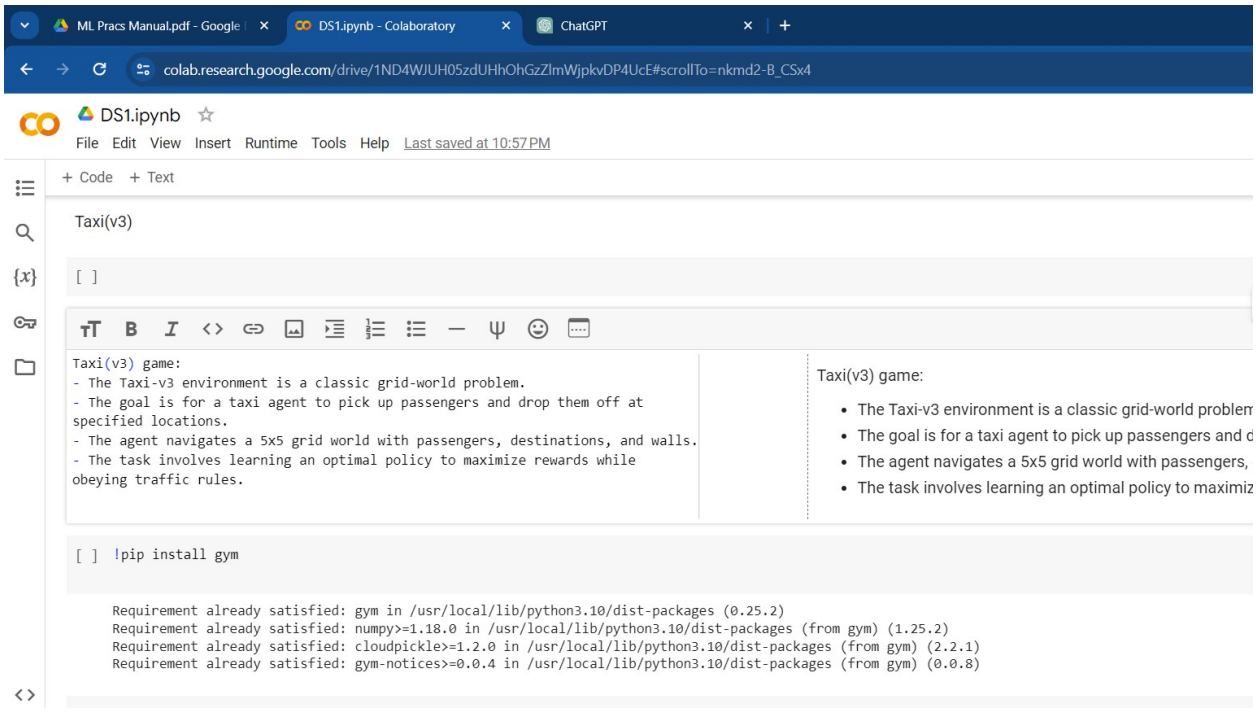


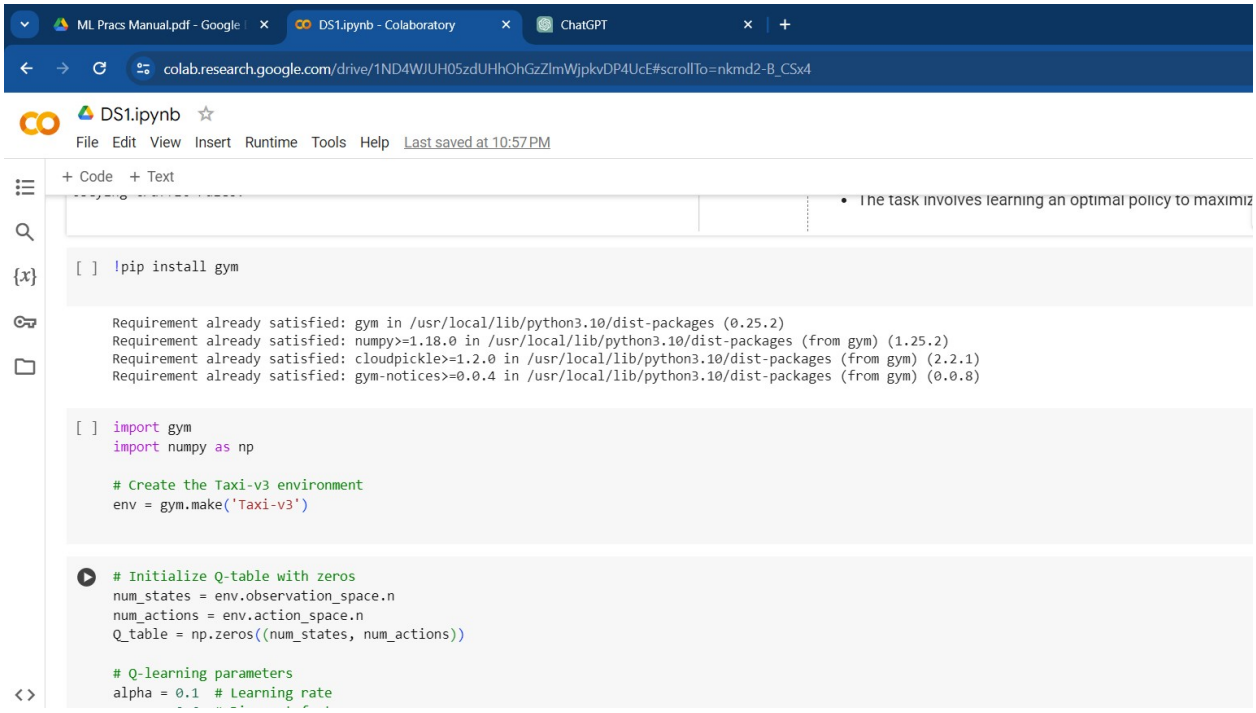

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=giTaWyA4ML3i
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
7
8 scaler = StandardScaler()
NameError: name 'df' is not defined
[ ] wcss = []
for i in range(1, 11):
kmeans = KMeans(n_clusters=i, init='k-means++', max_iter=300, n_init=10, random_state=42)
kmeans.fit(X_scaled)
wcss.append(kmeans.inertia_)
plt.plot(range(1, 11), wcss)
plt.title('Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()
k = 3 # example value
kmeans = KMeans(n_clusters=k, init='k-means++', max_iter=300, n_init=10, random_state=42)
y_kmeans = kmeans.fit_predict(X_scaled)

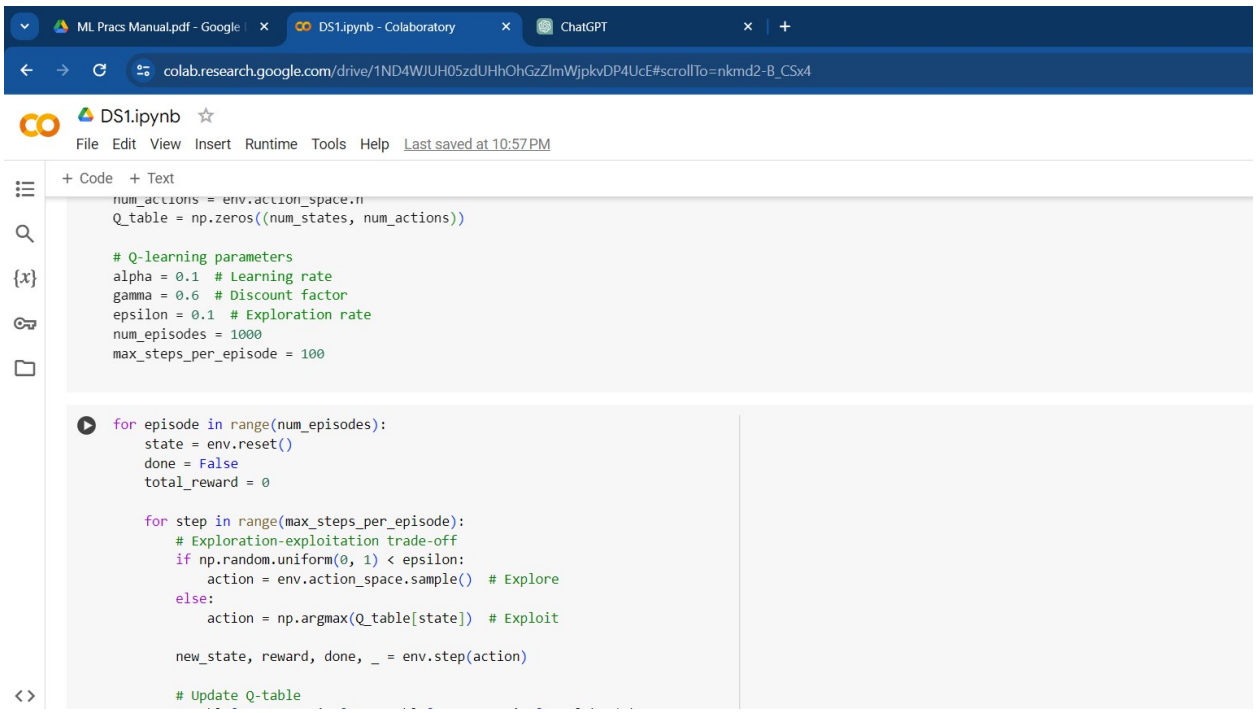

ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=giTaWyA4ML3i
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
kmeans.fit(X_scaled)
wcss.append(kmeans.inertia_)
plt.plot(range(1, 11), wcss)
plt.title('Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()
[ ] k = 3 # example value
kmeans = KMeans(n_clusters=k, init='k-means++', max_iter=300, n_init=10, random_state=42)
y_kmeans = kmeans.fit_predict(X_scaled)
from sklearn.decomposition import PCA
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_scaled)
plt.scatter(X_pca[y_kmeans == 0, 0], X_pca[y_kmeans == 0, 1], s = 100, c = 'red', label = 'Cluster 1')
plt.scatter(X_pca[y_kmeans == 1, 0], X_pca[y_kmeans == 1, 1], s = 100, c = 'blue', label = 'Cluster 2')
plt.scatter(X_pca[y_kmeans == 2, 0], X_pca[y_kmeans == 2, 1], s = 100, c = 'green', label = 'Cluster 3')
plt.title('Clusters of patients')
plt.xlabel('PCA 1')

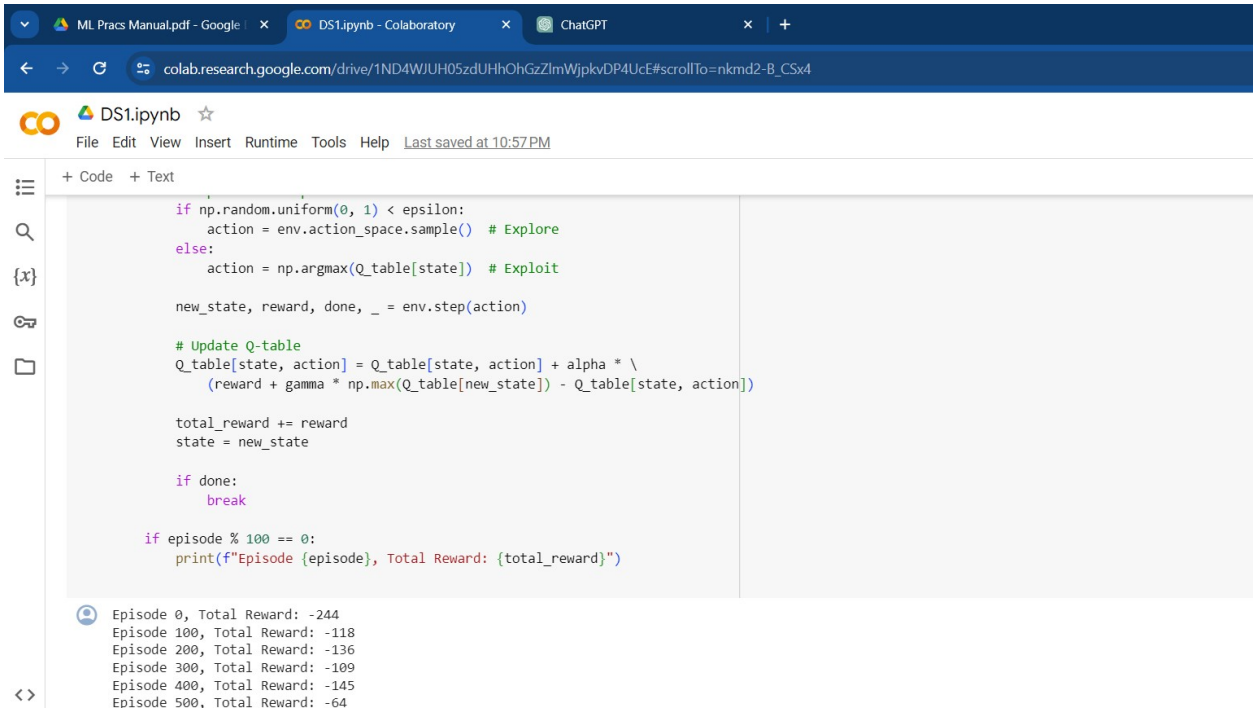
Practical No. 10 Aim:

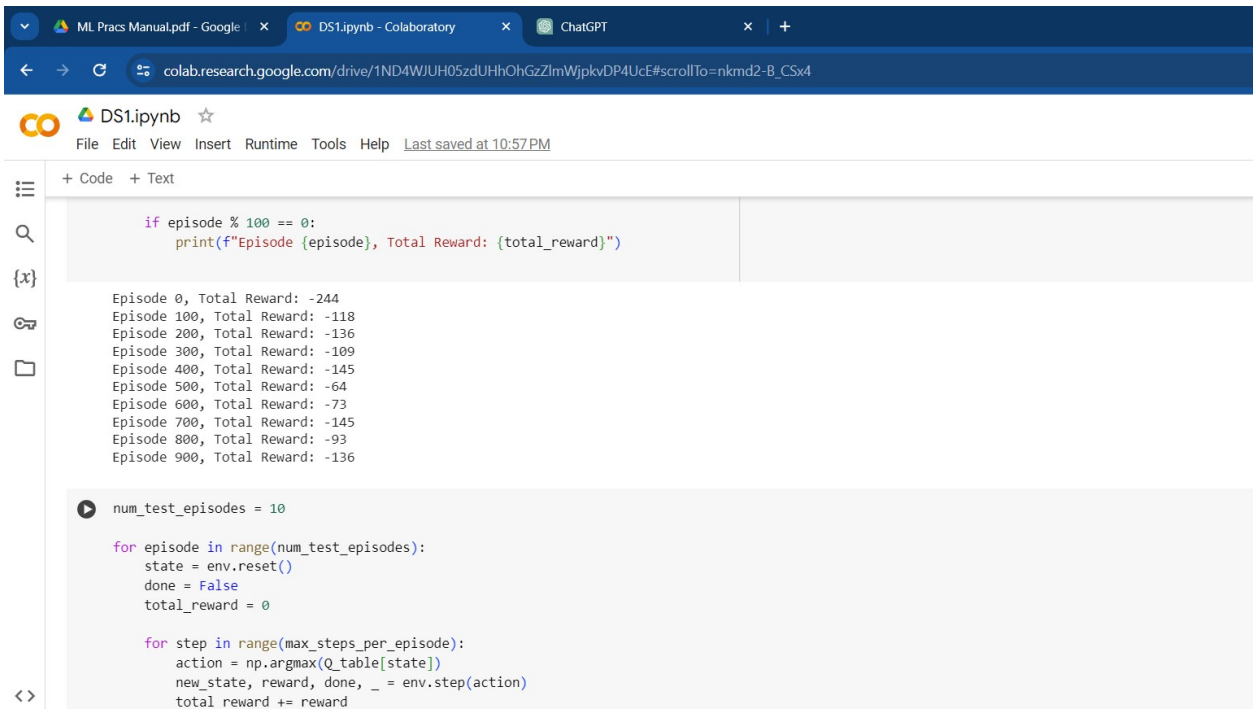
To perform reinforcement learning

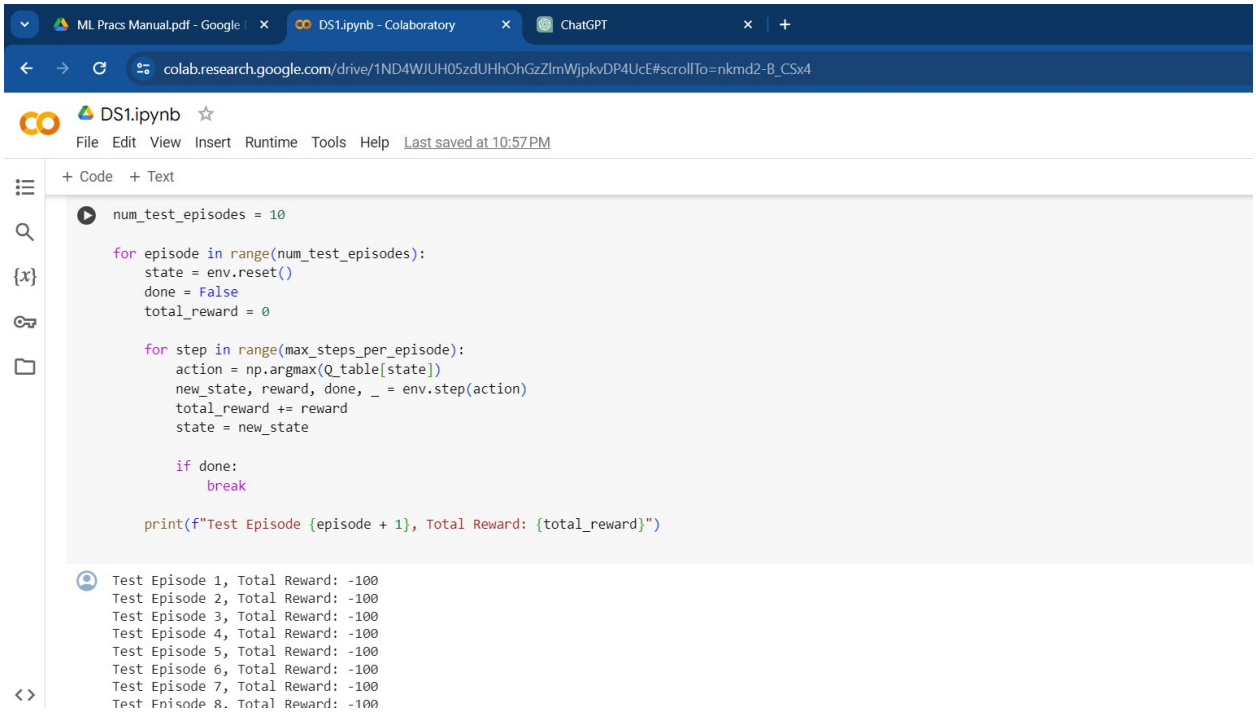



ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=nkmd2-B_CSx4
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
• The task involves learning an optimal policy to maximiz
[ ] !pip install gym
Requirement already satisfied: gym in /usr/local/lib/python3.10/dist-packages (0.25.2)
Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym) (1.25.2)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym) (2.2.1)
Requirement already satisfied: gym-notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym) (0.0.8)
[ ] import gym
import numpy as np
# Create the Taxi-v3 environment
env = gym.make('Taxi-v3')
# Initialize Q-table with zeros
num_states = env.observation_space.n
num_actions = env.action_space.n
Q_table = np.zeros((num_states, num_actions))
# Q-learning parameters
alpha = 0.1 # Learning rate


ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=nkmd2-B_CSx4
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
num_actions = env.action_space.n
Q_table = np.zeros((num_states, num_actions))
# Q-learning parameters
alpha = 0.1 # Learning rate
gamma = 0.6 # Discount factor
epsilon = 0.1 # Exploration rate
num_episodes = 1000
max_steps_per_episode = 100
for episode in range(num_episodes):
state = env.reset()
done = False
total_reward = 0
for step in range(max_steps_per_episode):
# Exploration-exploitation trade-off
if np.random.uniform(0, 1) < epsilon:
action = env.action_space.sample() # Explore
else:
action = np.argmax(Q_table[state]) # Exploit
new_state, reward, done, _ = env.step(action)
Update Q-table


ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=nkmd2-B_CSx4
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
if np.random.uniform(0, 1) < epsilon:
action = env.action_space.sample() # Explore
else:
action = np.argmax(Q_table[state]) # Exploit
new_state, reward, done, _ = env.step(action)
Update Q-table
Q_table[state, action] = Q_table[state, action] + alpha * \
(reward + gamma * np.max(Q_table[new_state]) - Q_table[state, action])
total_reward += reward
state = new_state
if done:
break
if episode % 100 == 0:
print(f"Episode {episode}, Total Reward: {total_reward}")
Episode 0, Total Reward: -244
Episode 100, Total Reward: -118
Episode 200, Total Reward: -136
Episode 300, Total Reward: -109
Episode 400, Total Reward: -145
Episode 500, Total Reward: -64


ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=nkmd2-B_CSx4
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
if episode % 100 == 0:
print(f"Episode {episode}, Total Reward: {total_reward}")
Episode 0, Total Reward: -244
Episode 100, Total Reward: -118
Episode 200, Total Reward: -136
Episode 300, Total Reward: -109
Episode 400, Total Reward: -145
Episode 500, Total Reward: -64
Episode 600, Total Reward: -73
Episode 700, Total Reward: -145
Episode 800, Total Reward: -93
Episode 900, Total Reward: -136
num_test_episodes = 10
for episode in range(num_test_episodes):
state = env.reset()
done = False
total_reward = 0
for step in range(max_steps_per_episode):
action = np.argmax(Q_table[state])
new_state, reward, done, _ = env.step(action)
total_reward += reward


ML Pracs Manual.pdf - Google
DS1.ipynb - Colaboratory
ChatGPT
colab.research.google.com/drive/1ND4WJUH05zdUHhOhGzZlmWjpkvDP4UcE#scrollTo=nkmd2-B_CSx4
DS1.ipynb
File Edit View Insert Runtime Tools Help Last saved at 10:57 PM
+ Code + Text
num_test_episodes = 10
for episode in range(num_test_episodes):
state = env.reset()
done = False
total_reward = 0
for step in range(max_steps_per_episode):
action = np.argmax(Q_table[state])
new_state, reward, done, _ = env.step(action)
total_reward += reward
state = new_state
if done:
break
print(f"Test Episode {episode + 1}, Total Reward: {total_reward}")
Test Episode 1, Total Reward: -100
Test Episode 2, Total Reward: -100
Test Episode 3, Total Reward: -100
Test Episode 4, Total Reward: -100
Test Episode 5, Total Reward: -100
Test Episode 6, Total Reward: -100
Test Episode 7, Total Reward: -100
Test Episode 8, Total Reward: -100