NEURAL NETWORK & DEEP LEARNING

ASSIGNMENT 4

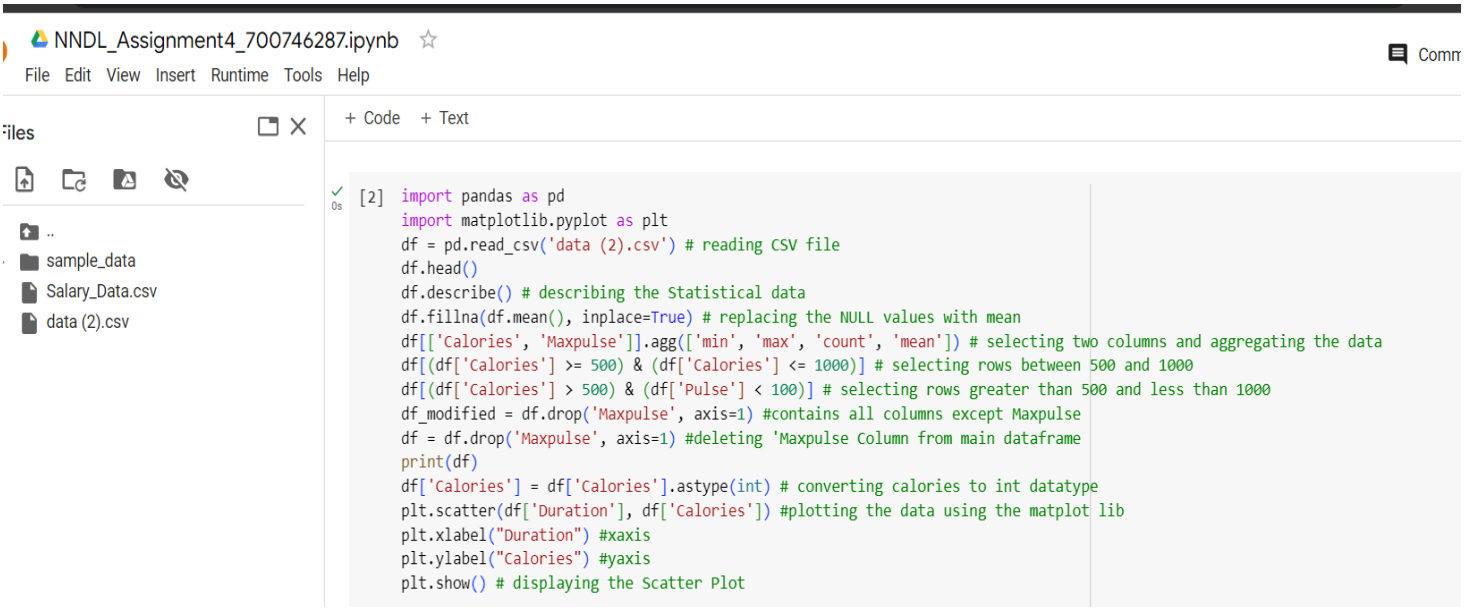
Name : SAI SNUSHA NAKKA

Student ID : 700746287

Git hub Link: https://github.com/NSnusha/NNDL_Assignment4

Video link:
https://drive.google.com/file/d/1V1HRWKh446Ffiz5HJT06j7l6mNgZt64Q/view?usp=sharing

1. 1. Data Manipulation a. Read the provided CSV file 'data.csv'. b. https://drive.google.com/drive/folders/1h8C3mLsso-R-sIOLsvoYwPLzy2fJ4IOF?usp=sharing c. Show the basic statistical description about the data. d. Check if the data has null values. i. Replace the null values with the mean e. Select at least two columns and aggregate the data using: min, max, count,mean. f. Filter the dataframe to select the rows with calories values between 500 and1000. g. Filter the dataframe to select the rows with calories values > 500 and pulse



NNDL_Assignment4_700746287.ipynb

colab.research.google.com/drive/1y3E0GuV7JC-4EaG7-OoKntZP6NZ0MG41#scrollTo=S2niCjzZMPji

NNDL_Assignment4_700746287.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

.. sample_data Salary_Data.csv data (2).csv

Disk 81.43 GB available

+ Code + Text

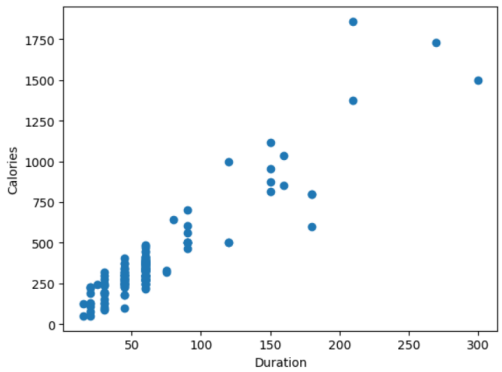
RAM Disk

```
plt.show() # displaying the Scatter Plot
```

[2]

```
     Duration  Pulse  Calories
0          60    110     409.1
1          60    117     479.0
2          60    103     340.0
3          45    109     282.4
4          45    117     406.0
..        ...    ...       ...
164        60    105     290.8
165        60    110     300.0
166        60    115     310.2
167        75    120     320.4
168        75    125     330.4

[169 rows x 3 columns]
```



1s completed at 5:04 PM

2. Linear Regression a) Import the given “Salary_Data.csv” b) Split the data in train_test partitions, such that 1/3 of the data is reserved as test subset. c) Train and predict the model. d) Calculate the mean_squared error e) Visualize both train and test data using scatter plot.

```
import pandas as pd
from sklearn.metrics import mean_squared_error
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split

salariesData = pd.read_csv('Salary_Data.csv') #importing data from the CSV file

#splitting the data into training and testing
X = salariesData.iloc[:, :-1].values
Y= salariesData.iloc[:, 1].values

#splitting 1/3 of the data as testsubset
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 1/3, random_state = 0)

# Fitting Simple Linear Regression to the training set
reg = LinearRegression()
reg.fit(X_train, Y_train)

# Predicting the Test set result
pred = reg.predict(X_test)

# Calculating the Mean_squared_error
mse = mean_squared_error(Y_test, pred)

#Visualising the Training set results and Test set results
plt.scatter(X_train, Y_train, color = 'blue')
plt.scatter(X_test, Y_test, color = 'red')
plt.title('Salary Data')
plt.xlabel('Experience (Years)')
plt.ylabel('Salary')
plt.show()
```