

Ex.No 4

Write a python program to implement linear models to approximate the given data

Aim:

To implement linear model in order to approximate the given data.

Procedure:

1. Import the necessary libraries

```
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd
```

2. Load the dataset

```
data_set= pd.read_csv('Salary_Data.csv')
data_set
```

3. Extract dependent and independent variable from the dataset

```
x= data_set.iloc[:, :-1].values
y= data_set.iloc[:, 1].values
x,y
```

4. Construct training and test dataset

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 1/3, random_state=0)
x_train,x_test,y_train,y_test
```

5. Fit the simple linear regression to training set

```
from sklearn.linear_model import LinearRegression
regressor= LinearRegression()
regressor.fit(x_train, y_train)
```

6. Predict the test set result

```
y_pred= regressor.predict(x_test)
x_pred= regressor.predict(x_train)
```

7. Visualize the training set result

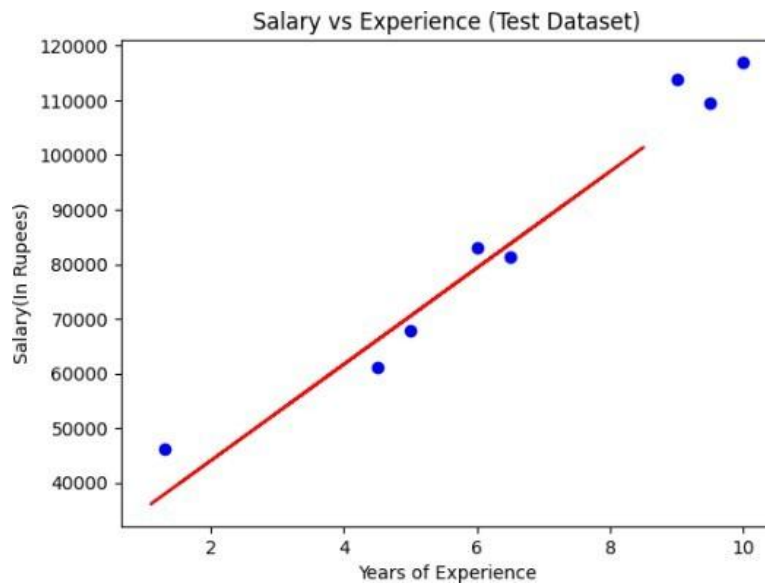
```
mtp.scatter(x_train, y_train, color="green")
mtp.plot(x_train, x_pred, color="red")
mtp.title("Salary vs Experience (Training Dataset)")
mtp.xlabel("Years of Experience")
mtp.ylabel("Salary(In Rupees)")
mtp.show()
```

8. Visualize the test set result

```
mtp.scatter(x_test, y_test, color="blue")
mtp.plot(x_train, x_pred, color="red")
mtp.title("Salary vs Experience (Test Dataset)")
mtp.xlabel("Years of Experience")
mtp.ylabel("Salary(In Rupees)")
mtp.show()
```

Sample Output:





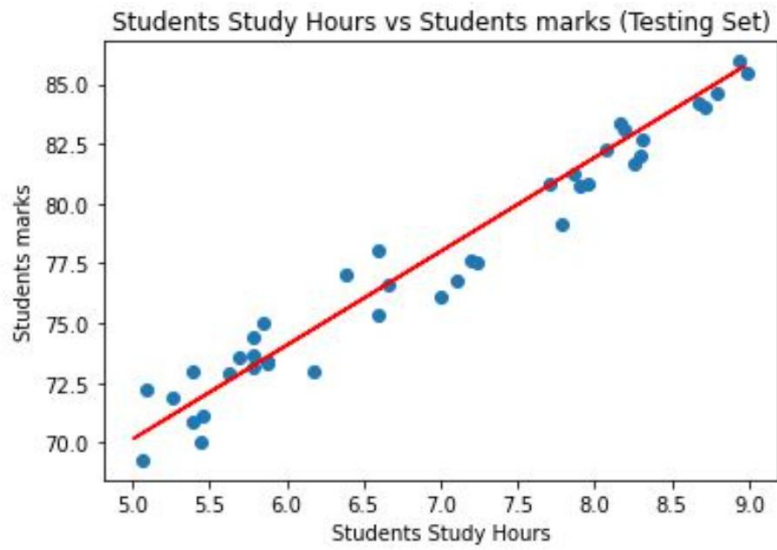
Data Set:

Roll No	Student Name	Hindi	English	Math	Physics
1	RAM	20	10	14	18
2	ASHOK	21	12	14	12
3	MANOJ	33	15	7	14
4	RAJESH	15	14	8	16
5	RANJANA	14	17	10	13
6	POOJA	16	8	20	17
7	MAHESH	18	19	3	10
8	ASHUTOSH	19	20	7	14
9	ANIL	22	13	8	12
10	PREM	26	12	10	11

CODING:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
df = pd.read_csv("Student_Marks_Data.csv")
print("Dataset Preview:")
print(df.head())
X = df.iloc[:, 2].values.reshape(-1, 1)
y = df.iloc[:, 5].values
data_combined = pd.DataFrame(np.hstack((X, y.reshape(-1, 1))), columns=['X', 'y'])
data_combined.dropna(inplace=True)
X = data_combined['X'].values.reshape(-1, 1)
y = data_combined['y'].values
model = LinearRegression()
model.fit(X, y)
X_test = np.linspace(X.min(), X.max(), 100).reshape(-1, 1)
y_pred = model.predict(X_test)
plt.scatter(X, y, label='Actual Data')
plt.plot(X_test, y_pred, label='Regression Line')
plt.xlabel('Hindi Marks')
plt.ylabel('Physics Marks')
plt.title('Linear Regression: Hindi vs Physics')
plt.legend()
plt.show()
```

Output:



MARK ALLOCATION	
Conduct of Experiment(30)	
Record Observation (20)	
Viva (10)	
Total (60)	
Signature of the Faculty with Date	

Result:

Thus the linear model is implemented to approximate the given data.