CS-23334 FUNDAMENTALS OF DATA SCIENCE

ABENANTHAN P 240701005

EXPERIMENT 7

7. Experiment to understand Linear Regression for a given data set.

Aim:

To understand and evluate Linear Regression Model for predicting the continuous dependent variable

Description:

Understand the Linear regression for the dataset given.

Algorithm:

Step 1: Define the Problem and Select Features

Step 2: Split the Dataset into Training and Testing Sets

Step 3: Train the Linear Regression Model

Step 4: Evaluate Model Performance Using Metrics

Step 5: Visualize Predictions and Residuals

Code With Output:

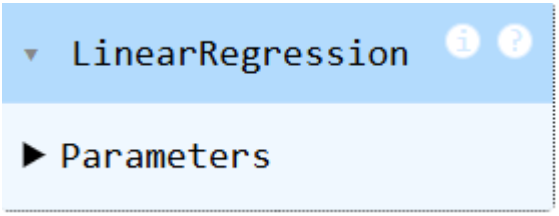
```
import numpy as np
import pandas as pd
df=pd.read_csv(r'D:\REC 2nd Year\Data Science\Data Sets\Linear
Regression Dataset.csv')

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   YearsExperience  30 non-null     float64
 1   Salary           30 non-null     int64
dtypes: float64(1), int64(1)
memory usage: 612.0 bytes
```

```
df.dropna(inplace=True)

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   YearsExperience  30 non-null     float64
 1   Salary           30 non-null     int64
dtypes: float64(1), int64(1)
memory usage: 612.0 bytes
```

```
features=df.iloc[:, [0]].values
label=df.iloc[:, 1].values

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(features,label,test_siz
e=0.2,random_state=23)

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



```
model.score(x_train,y_train)
```

```
0.9603182547438908
```

```
model.score(x_test,y_test)
```

```
0.9184170849214232
```

```
model.coef_
```

```
array([9281.30847068])
```

```
model.intercept_
```

```
np.float64(27166.736828912777)
```

```
import pickle
pickle.dump(model,open('SalaryPred.model','wb'))

model=pickle.load(open('SalaryPred.model','rb'))

yr_of_exp=float(input("Enter Years of Experience: "))
yr_of_exp_NP=np.array([[yr_of_exp]])
Salary=model.predict(yr_of_exp_NP)
```

```
Enter Years of Experience:  44
```

```
print("Estimated Salary for {} years of experience is {}:
" .format(yr_of_exp,Salary))
```

```
Estimated Salary for 44.0 years of experience is [435544.30953887]:
```

Result:

Thus evaluating Linear Regression for a dataset is completed using a python program.