Machine Learning

Multiple Linear Regression

## ▾ Step-1 Import Libraries

```
1   import pandas as pd
2   import numpy as np
3   import seaborn as sns
4   import matplotlib.pyplot as plt
```

## ▾ Step-2 Import dataset

```
1   import pandas as pd
2   df = pd.read_csv("ml_data_salary.csv")
3   df.head()
4
```

|   | age | distance | YearsExperience | Salary |
|---|-----|----------|-----------------|--------|
| **0** | 31.1 | 77.75 | 1.1 | 39343 |
| **1** | 31.3 | 78.25 | 1.3 | 46205 |
| **2** | 31.5 | 78.75 | 1.5 | 37731 |
| **3** | 32.0 | 80.00 | 2.0 | 43525 |
| **4** | 32.2 | 80.50 | 2.2 | 39891 |

## ▾ Step-3 Define dependent and independent variables

```
1 X = df[["age","distance", "YearsExperience"]]
2 y = df["Salary"]
```

## ▾ Step-4 Fit Linear Regression Model

```
1 from sklearn.linear_model import LinearRegression
2 model = LinearRegression()
3 model = model.fit(X, y)
4 model
```

```
▾ LinearRegression
LinearRegression()
```

## ▾ model.coef

```
1 model.coef_
```

```
array([-3.00216193e+15,  1.18788781e+15,  3.24424072e+13])
```

## ▾ Step-5 Evaluating Model Fitness

```
1 print(model.score(X, y))
```

```
0.9569960750337954
```

## ▾ Step-6 Prediction of unknown values

```
1 model.predict([[31.1,77.75,1.1]])
2
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression was
  warnings.warn(
array([36209.375])
```

## ▾ Step-7 Prediction accuracy score

```
 1 import numpy as np
 2 import pandas as pd
 3 from sklearn.model_selection import train_test_split
 4 from sklearn.linear_model import LinearRegression
 5 from sklearn.metrics import r2_score
 6 data = pd.read_csv('ml_data_salary.csv') # Replace 'your_dataset.csv' with your actual dataset filename
 7 X = data.drop('distance', axis=1) # Replace 'target_variable' with the column name of your target variable
 8 y = data['distance']
 9 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=43)
10 model = LinearRegression()
11 model.fit(X_train, y_train)
12 y_pred = model.predict(X_test)
13 accuracy = r2_score(y_test, y_pred)
14 print("Accuracy score:", accuracy)
```

```
Accuracy score: 1.0
```