

Salary Prediction Machine Learning Model Based on Years Of Experience

Simple Linear Regression Model Practice Vivek Chauhan

```
In [5]: # first of all install the machine Learning library which name is scikit-Learn
```

```
!pip install scikit-learn
```

```
Requirement already satisfied: scikit-learn in c:\users\vivek chauhan\anaconda3\lib\site-packages (1.5.1)
Requirement already satisfied: numpy>=1.19.5 in c:\users\vivek chauhan\anaconda3\lib\site-packages (from scikit-learn) (1.26.4)
Requirement already satisfied: scipy>=1.6.0 in c:\users\vivek chauhan\anaconda3\lib\site-packages (from scikit-learn) (1.13.1)
Requirement already satisfied: joblib>=1.2.0 in c:\users\vivek chauhan\anaconda3\lib\site-packages (from scikit-learn) (1.4.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in c:\users\vivek chauhan\anaconda3\lib\site-packages (from scikit-learn)
(3.5.0)
```

```
In [1]: # Upload the necessary Libraries to work with data
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
from sklearn.model_selection import train_test_split
```

```
In [2]: # Load the dataset
```

```
data = pd.read_csv("Salary Data.csv")
data.head(3)
```

Out[2]:

	YearsExperience	Salary
0	1.1	39343.0
1	1.3	46205.0
2	1.5	37731.0

In [3]:

```
# check the dataset is null or not
```

```
data.isnull().sum()
```

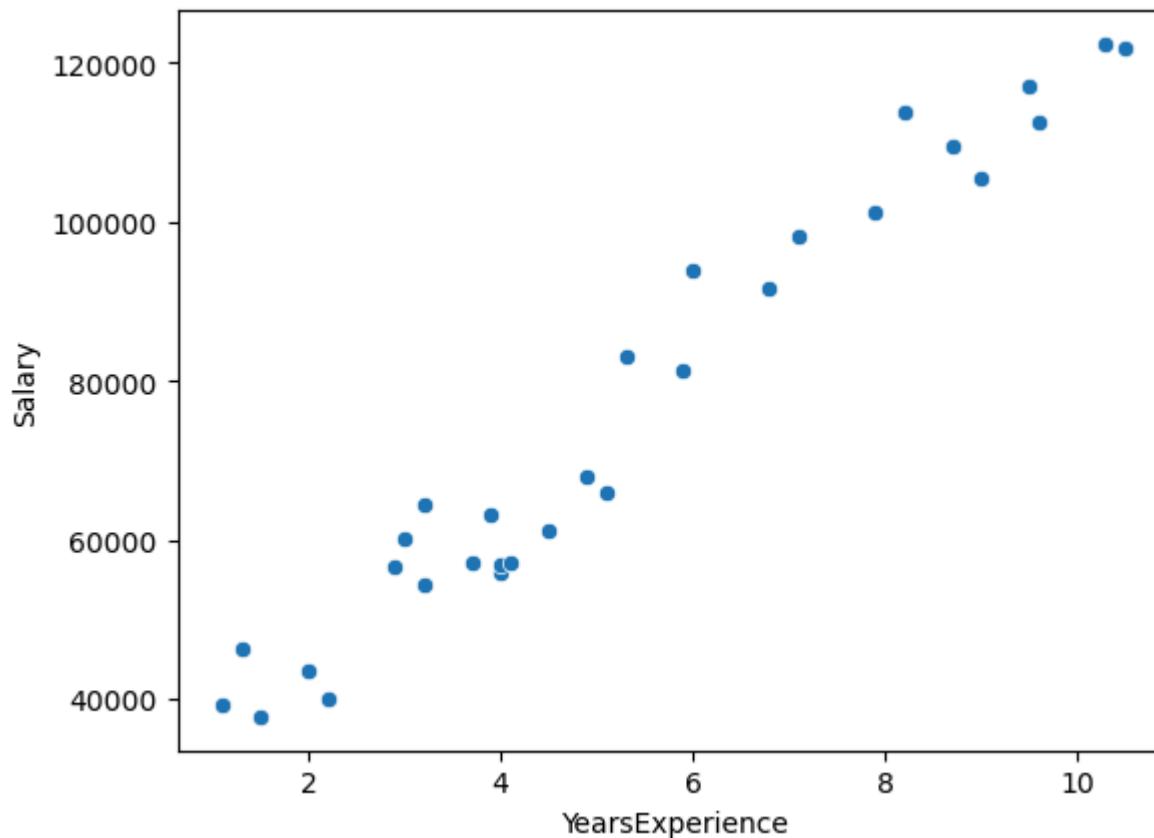
Out[3]:

```
YearsExperience      0
Salary              0
dtype: int64
```

In [4]:

```
# check our data is linearly increasing or not
```

```
sns.scatterplot(x = "YearsExperience",y = "Salary",data=data)
plt.show()
```



As you can see our data is linearly increasing means if one column data is increasing then y column data is increasing that's why we used here Simple-Linear Regressing Technique.

Split the Dataset

In [5]: `# split the data for input and output`

```
x = data[["YearsExperience"]]
y = data["Salary"]
```

Assign train & test model

In [151...]

```
# Let's train and test our data

x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.51,random_state = 51)
```

Upload the necessary library to work with LinearRegression Model

In [152...]

```
# upload the Linear regression module

from sklearn.linear_model import LinearRegression
```

Fit the Dataset

In [153...]

```
lr = LinearRegression()
lr.fit(x_train,y_train)
```

Out[153...]

```
▼ LinearRegression ⓘ ?
```

```
LinearRegression()
```

Pridict the salary based on years of experience & check the accuracy of our model.

In [154...]

```
lr.predict([[1.1]])
```

```
Out[154... array([34464.01665906])
```

```
In [155... # Let's check the accuracy score of our model
```

```
lr.score(x_train,y_train)*100
```

```
Out[155... 98.11107318211484
```

Cross check the value of line or you can say that cross check the value of our model.

```
In [157... # check the  $y = mx + c$  line
```

```
lr.intercept_ # for c value means the angle of theta
```

```
Out[157... 23998.28982199912
```

```
In [166... # check the  $y = mx + c$  line
```

```
lr.coef_[0] # for m value mens how the line is cross/brekonw the x-axis Line
```

```
Out[166... 9514.297124600633
```

```
In [168... # Put the value in Line formula like  $y = mx + c$ 
```

```
line_y = 9514.297124600633 * 1.1 + 23998.28982199912  
line_y
```

```
Out[168... 34464.01665905982
```