

Salary_hike

July 18, 2023

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
[2]: df=pd.read_csv(r"C:\Users\Saina\OneDrive\Desktop\Salary_Data.csv")
```

```
[3]: x=df.iloc[:, :-1]
y=df.iloc[:, -1]
```

```
[4]: x
```

```
[4]:      YearsExperience
0              1.1
1              1.3
2              1.5
3              2.0
4              2.2
5              2.9
6              3.0
7              3.2
8              3.2
9              3.7
10             3.9
11             4.0
12             4.0
13             4.1
14             4.5
15             4.9
16             5.1
17             5.3
18             5.9
19             6.0
20             6.8
21             7.1
22             7.9
23             8.2
24             8.7
25             9.0
```

26	9.5
27	9.6
28	10.3
29	10.5

```
[5]: y
```

```
[5]: 0    39343.0
      1    46205.0
      2    37731.0
      3    43525.0
      4    39891.0
      5    56642.0
      6    60150.0
      7    54445.0
      8    64445.0
      9    57189.0
     10    63218.0
     11    55794.0
     12    56957.0
     13    57081.0
     14    61111.0
     15    67938.0
     16    66029.0
     17    83088.0
     18    81363.0
     19    93940.0
     20    91738.0
     21    98273.0
     22   101302.0
     23   113812.0
     24   109431.0
     25   105582.0
     26   116969.0
     27   112635.0
     28   122391.0
     29   121872.0
```

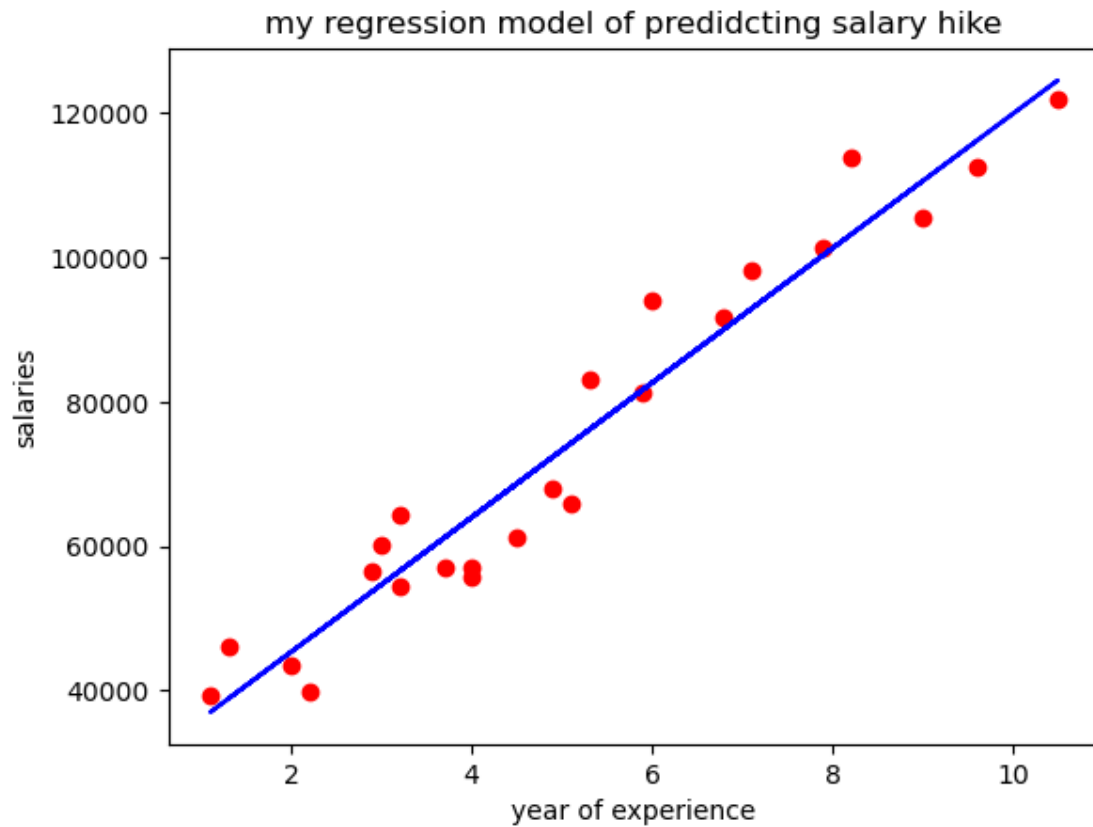
Name: Salary, dtype: float64

```
[6]: from sklearn.model_selection import train_test_split
      x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=0)
```

```
[7]: from sklearn.linear_model import LinearRegression
      regressor=LinearRegression()
      regressor.fit(x_train,y_train)
```

```
[7]: LinearRegression()
```

```
[8]: plt.title('my regression model of predidcting salary hike')
plt.xlabel('year of experience')
plt.ylabel('salaries')
plt.scatter(x_train,y_train,color='red')
plt.plot(x_train,regressor.predict(x_train),color='blue')
plt.show()
```



```
[9]: plt.title('my regression model of predidcting salarie hike')
plt.xlabel('year of experience')
plt.ylabel('salaries')
plt.scatter(x_test,y_test,color='red')
plt.plot(x_train,regressor.predict(x_train),color='blue')
plt.show()
```



```
[10]: regressor.intercept_
```

```
[10]: 26780.099150628186
```

```
[11]: regressor.coef_
```

```
[11]: array([9312.57512673])
```