# simple-linear-regression

### April 4, 2024

## 0.1 To Predict the Salary of a person Using Linear Regression

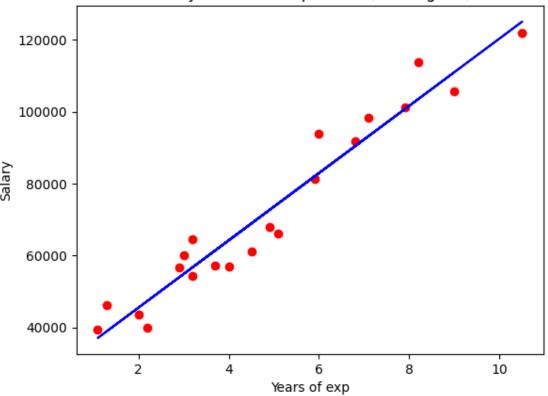
```
[19]: #importing the libraries
      import numpy as np
      import matplotlib.pyplot as plt
      import pandas as pd
[20]: #importing the datasets
      dataset = pd.read_csv(r'C:\Users\ntpc\Desktop\Salary_Data.csv')
[21]: dataset.head()
[21]:
         YearsExperience
                           Salary
                     1.1 39343.0
      1
                     1.3 46205.0
      2
                     1.5 37731.0
                     2.0 43525.0
      3
                     2.2 39891.0
[22]: dataset.shape
[22]: (30, 2)
[23]: \#seperating X \ and y
      X = dataset.iloc[:,:-1].values
      y = dataset.iloc[:,-1].values
[24]: #dividing into train and test
      from sklearn.model_selection import train_test_split
      X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,_
       →random_state=0)
[25]: from sklearn.linear_model import LinearRegression
[26]: regressor = LinearRegression()
[27]: regressor.fit(X_train,y_train)
```

### [27]: LinearRegression()

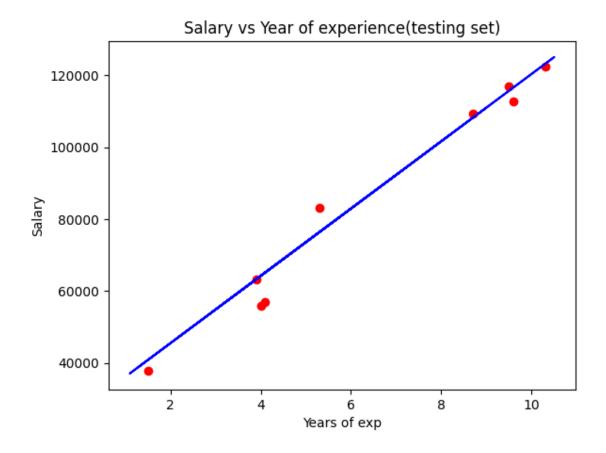
```
[28]: #prediction
y_pred = regressor.predict(X_test)
```

```
[29]: #visualising the training set
plt.scatter(X_train,y_train,color = 'red')
plt.plot(X_train,regressor.predict(X_train),color = 'blue')
plt.title('Salary vs Year of experience(training set)')
plt.xlabel('Years of exp')
plt.ylabel('Salary')
plt.show()
```

## Salary vs Year of experience(training set)



```
[30]: #visualising the testing set
plt.scatter(X_test,y_test,color = 'red')
plt.plot(X_train,regressor.predict(X_train),color = 'blue')
plt.title('Salary vs Year of experience(testing set)')
plt.xlabel('Years of exp')
plt.ylabel('Salary')
plt.show()
```



```
[31]: #checking the score
from sklearn.metrics import r2_score

[32]: r2_score(y_test,y_pred)

[32]: 0.9740993407213511

[]:
```