Task

On

Machine Learning

**Course**: Artificial Intelligence

(Machine Learning & Deep Learning)

Name: Abbas Shafi

Father Name: Dr Muhammad Shafi

Week: 09

Submitted to: Sir Syed Nazir Afridi



National Vocational & Technical Training Commission

National Center For Big Data & Cloud Computing

University of Engineering & Technology

Peshawar

[1]:

**import numpy as np import pandas as pd**

**import matplotlib.pyplot as plt**

# Simple Linear Regression

[4]:

# Importing the Dataset

|  |  |  |  |
| --- | --- | --- | --- |
| [4]: |  | YearsExperience | Salary |
|  | 0 | 1.1 | 39343.0 |
|  | 1 | 1.3 | 46205.0 |
|  | 2 | 1.5 | 37731.0 |
|  | 3 | 2.0 | 43525.0 |
|  | 4 | 2.2 | 39891.0 |
|  | 5 | 2.9 | 56642.0 |
|  | 6 | 3.0 | 60150.0 |
|  | 7 | 3.2 | 54445.0 |
|  | 8 | 3.2 | 64445.0 |
|  | 9 | 3.7 | 57189.0 |
|  | 10 | 3.9 | 63218.0 |
|  | 11 | 4.0 | 55794.0 |
|  | 12 | 4.0 | 56957.0 |
|  | 13 | 4.1 | 57081.0 |
|  | 14 | 4.5 | 61111.0 |
|  | 15 | 4.9 | 67938.0 |
|  | 16 | 5.1 | 66029.0 |
|  | 17 | 5.3 | 83088.0 |
|  | 18 | 5.9 | 81363.0 |
|  | 19 | 6.0 | 93940.0 |
|  | 20 | 6.8 | 91738.0 |
|  | 21 | 7.1 | 98273.0 |

|  |  |
| --- | --- |
| 22 | 7.9 101302.0 |
| 23 | 8.2 113812.0 |
| 24 | 8.7 109431.0 |
| 25 | 9.0 105582.0 |
| 26 | 9.5 116969.0 |
| 27 | 9.6 112635.0 |
| 28 | 10.3 122391.0 |
| 29 | 10.5 121872.0 |

[5]:

[6]:

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 |

[7]:

[8]:

Salary

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

# Spliting the Dataset into train\_test\_split

[9]:

# Training the S Linear Regression

[9]: LinearRegression()

# Prediction

[10]:

[10]: array([[ 36239.85597261],

[ 34336.22684101],

[ 66697.92207822],

[ 59083.40555182],

[ 91445.10078903],

[ 80975.14056523],

[101915.06101284],

[ 52420.70359122]])

[11]:

[12]:

*#Visualizing the training set result*

*#y\_test contain the real salary and y\_pred contain the predicted salary #we will plot the red plots for real salaries and blue straight line for*␣

*‹→predicted salaries*

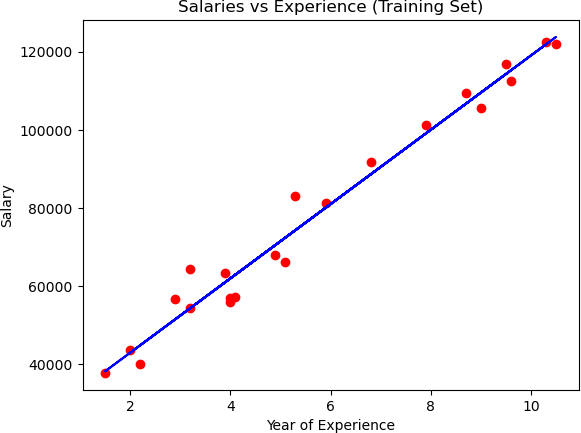
plt.scatter(x\_train,y\_train,color='red') plt.plot(x\_train,regress.predict(x\_train),color='blue') plt.title("Salaries vs Experience (Training Set)") plt.xlabel("Year of Experience")

plt.ylabel("Salary") plt.show()

Salary

|  |  |
| --- | --- |
| 1 | 46205.0 |
| 0 | 39343.0 |
| 14 | 61111.0 |
| 9 | 57189.0 |
| 21 | 98273.0 |
| 19 | 93940.0 |
| 23 | 113812.0 |
| 6 | 60150.0 |

# Visualization



[13]:



[14]:

[14]: array([[9518.145658]])

[15]:

[15]: array([23866.26661721])

[17]:

# Checking Accuracy

[17]: 0.8877872430356781

[18]:

[20]:

Mean Sqr error is: 71895163.96158665

# Multiple Linear Regression

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| [20]: | R&D Spend | Administration | Marketing Spend | State | Profit |
|  | 0 165349.20 | 136897.80 | 471784.10 | New York | 192261.83 |
|  | 1 162597.70 | 151377.59 | 443898.53 | California | 191792.06 |
|  | 2 153441.51 | 101145.55 | 407934.54 | Florida | 191050.39 |
|  | 3 144372.41 | 118671.85 | 383199.62 | New York | 182901.99 |
|  | 4 142107.34 | 91391.77 | 366168.42 | Florida | 166187.94 |
|  | 5 131876.90 | 99814.71 | 362861.36 | New York | 156991.12 |
|  | 6 134615.46 | 147198.87 | 127716.82 | California | 156122.51 |
|  | 7 130298.13 | 145530.06 | 323876.68 | Florida | 155752.60 |
|  | 8 120542.52 | 148718.95 | 311613.29 | New York | 152211.77 |
|  | 9 123334.88 | 108679.17 | 304981.62 | California | 149759.96 |
|  | 10 101913.08 | 110594.11 | 229160.95 | Florida | 146121.95 |
|  | 11 100671.96 | 91790.61 | 249744.55 | California | 144259.40 |
|  | 12 93863.75 | 127320.38 | 249839.44 | Florida | 141585.52 |
|  | 13 91992.39 | 135495.07 | 252664.93 | California | 134307.35 |
|  | 14 119943.24 | 156547.42 | 256512.92 | Florida | 132602.65 |
|  | 15 114523.61 | 122616.84 | 261776.23 | New York | 129917.04 |
|  | 16 78013.11 | 121597.55 | 264346.06 | California | 126992.93 |
|  | 17 94657.16 | 145077.58 | 282574.31 | New York | 125370.37 |
|  | 18 91749.16 | 114175.79 | 294919.57 | Florida | 124266.90 |
|  | 19 86419.70 | 153514.11 | 0.00 | New York | 122776.86 |
|  | 20 76253.86 | 113867.30 | 298664.47 | California | 118474.03 |
|  | 21 78389.47 | 153773.43 | 299737.29 | New York | 111313.02 |
|  | 22 73994.56 | 122782.75 | 303319.26 | Florida | 110352.25 |
|  | 23 67532.53 | 105751.03 | 304768.73 | Florida | 108733.99 |
|  | 24 77044.01 | 99281.34 | 140574.81 | New York | 108552.04 |
|  | 25 64664.71 | 139553.16 | 137962.62 | California | 107404.34 |
|  | 26 75328.87 | 144135.98 | 134050.07 | Florida | 105733.54 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 27 | 72107.60 | 127864.55 | 353183.81 | New York | 105008.31 |
| 28 | 66051.52 | 182645.56 | 118148.20 | Florida | 103282.38 |
| 29 | 65605.48 | 153032.06 | 107138.38 | New York | 101004.64 |
| 30 | 61994.48 | 115641.28 | 91131.24 | Florida | 99937.59 |
| 31 | 61136.38 | 152701.92 | 88218.23 | New York | 97483.56 |
| 32 | 63408.86 | 129219.61 | 46085.25 | California | 97427.84 |
| 33 | 55493.95 | 103057.49 | 214634.81 | Florida | 96778.92 |
| 34 | 46426.07 | 157693.92 | 210797.67 | California | 96712.80 |
| 35 | 46014.02 | 85047.44 | 205517.64 | New York | 96479.51 |
| 36 | 28663.76 | 127056.21 | 201126.82 | Florida | 90708.19 |
| 37 | 44069.95 | 51283.14 | 197029.42 | California | 89949.14 |
| 38 | 20229.59 | 65947.93 | 185265.10 | New York | 81229.06 |
| 39 | 38558.51 | 82982.09 | 174999.30 | California | 81005.76 |
| 40 | 28754.33 | 118546.05 | 172795.67 | California | 78239.91 |
| 41 | 27892.92 | 84710.77 | 164470.71 | Florida | 77798.83 |
| 42 | 23640.93 | 96189.63 | 148001.11 | California | 71498.49 |
| 43 | 15505.73 | 127382.30 | 35534.17 | New York | 69758.98 |
| 44 | 22177.74 | 154806.14 | 28334.72 | California | 65200.33 |
| 45 | 1000.23 | 124153.04 | 1903.93 | New York | 64926.08 |
| 46 | 1315.46 | 115816.21 | 297114.46 | Florida | 49490.75 |
| 47 | 0.00 | 135426.92 | 0.00 | California | 42559.73 |
| 48 | 542.05 | 51743.15 | 0.00 | New York | 35673.41 |
| 49 | 0.00 | 116983.80 | 45173.06 | California | 14681.40 |

[21]:

[22]:

*#OneHotEncoding*

**from sklearn.compose import** ColumnTransformer

**from sklearn.preprocessing import** OneHotEncoder

*#encoder to encode type is OneHotEncoding and which index or column*

ct=ColumnTransformer([('encoder',OneHotEncoder(),[3])], remainder="passthrough") x=np.array(ct.fit\_transform(x))

[23]:

[24]:

[24]: LinearRegression()

[25]:

[26]:

|  |  |
| --- | --- |
| [ 74061.28471131 | 46009.23798761 99637.26360757 155786.53229374 |
| 127636.76349539 | 192765.18597815 63906.99972423 54935.14415852 |
| 84532.35238007 | 109460.29711992] |

[27]:

[28]:

[ 90708.19 42559.73 103282.38 149759.96 134307.35 192261.83 65200.33

49490.75 81005.76 108733.99]

|  |  |
| --- | --- |
| [[ 74061.28 | 90708.19] |
| [ 46009.24 | 42559.73] |
| [ 99637.26 | 103282.38] |
| [155786.53 | 149759.96] |
| [127636.76 | 134307.35] |
| [192765.19 | 192261.83] |
| [ 63907. | 65200.33] |
| [ 54935.14 | 49490.75] |
| [ 84532.35 | 81005.76] |
| [109460.3 | 108733.99]] |

[29]:

[29]: 0.9783259006626401

# Polynomial Regression

[32]:

|  |  |  |  |
| --- | --- | --- | --- |
| [32]: | Position | Level | Salary |
|  | 0 Business Analyst | 1 | 45000 |
|  | 1 Junior Consultant | 2 | 50000 |
|  | 2 Senior Consultant | 3 | 60000 |
|  | 3 Manager | 4 | 80000 |
|  | 4 Country Manager | 5 | 110000 |
|  | 5 Region Manager | 6 | 150000 |
|  | 6 Partner | 7 | 200000 |

[33]:

7 Senior Partner 8 300000 8 C-level 9 500000

9 CEO 10 1000000

[34]:

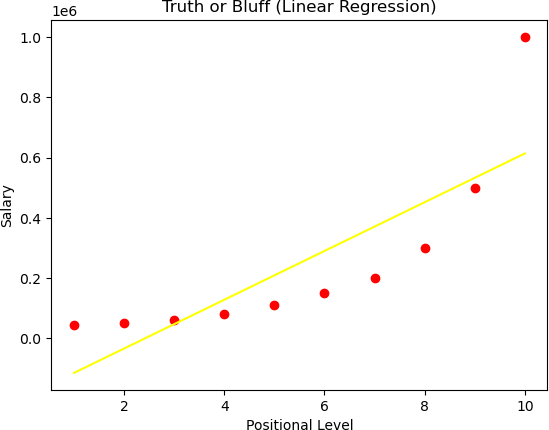
1. : LinearRegression()

[35]:

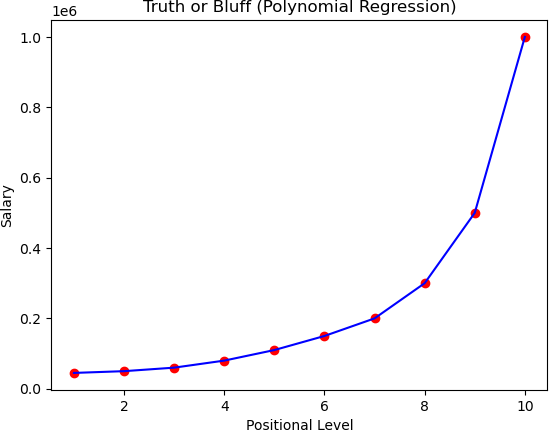
1. : LinearRegression()

[36]:

[37]:



[38]:



[39]:

[39]: array([330378.79])