



NATIONAL UNIVERSITY - MANILA  
COLLEGE OF COMPUTING AND INFORMATION TECHNOLOGY

# Introduction to Machine Learning

# LINEAR

# REGRESSION

(PREDICTING EXAM SCORES)

# ASSIGNMENT NO. 2

Prepared by:

**Rae S. Paulos**

## TRAINING DATA (5 STUDENTS)

Student	Hours Studied (x)	Exam Score (y)
1	1	52
2	2	57
3	3	61
4	4	65
5	5	70

### 1. Fill in the table (14 points)

- Compute for  $x^2$  for each student
- Compute for  $xy$  for each student
- Find the total for x, y,  $x^2$ , and  $xy$

Student	Hours Studied (x)	Exam Score (y)	xy	$x^2$
1	1	52	52	1
2	2	57	114	4
3	3	61	183	9
4	4	65	260	16
5	5	70	350	25
<b>Total:</b>	<b>15</b>	<b>305</b>	<b>959</b>	<b>55</b>

$$1. \quad x = 1, \quad y = 52$$

$$xy = 1 \times 52$$

$$= \underline{\underline{52}}$$

$$2. \quad x = 2, \quad y = 57$$

$$xy = 2 \times 57$$

$$= \underline{\underline{114}}$$

$$3. \quad x = 3, \quad y = 61$$

$$xy = 3 \times 61$$

$$= \underline{\underline{183}}$$

$$4. \quad x = 4, \quad y = 65$$

$$xy = 4 \times 65$$

$$= \underline{\underline{260}}$$

$$5. \quad x = 5, \quad y = 70$$

$$xy = 5 \times 70$$

$$= \underline{\underline{350}}$$

$$1. \quad x = 1$$

$$1^2 = \underline{\underline{1}}$$

$$2. \quad x = 2$$

$$2^2 = \underline{\underline{4}}$$

$$3. \quad x = 3$$

$$3^2 = \underline{\underline{9}}$$

$$4. \quad x = 4$$

$$4^2 = \underline{\underline{16}}$$

$$5. \quad x = 5$$

$$5^2 = \underline{\underline{25}}$$

$$\sum x = 1 + 2 + 3 + 4 + 5$$

$$= \underline{\underline{15}}$$

$$\sum y = 52 + 57 + 61 + 65 + 70$$

$$= \underline{\underline{305}}$$

$$\sum xy = 52 + 114 + 183 + 260 + 350$$

$$= \underline{\underline{959}}$$

$$\sum x^2 = 1 + 4 + 9 + 16 + 25$$

$$= \underline{\underline{55}}$$

## 2. Compute the Slope $m$ (5 points)

Slope	4.4
-------	-----

FORMULA =  $m = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$

GIVEN =

$$n = 5$$

$$\sum xy = 959$$

$$\sum x = 15$$

$$\sum y = 305$$

$$\sum x^2 = 55$$

SOLUTION

$$m = \frac{5(959) - (15)(305)}{5(55) - (15)^2}$$

$$= \frac{4,795 - 4,575}{275 - 225}$$

$$= \frac{220}{50} = \underline{\underline{4.4}}$$

### 3. Compute the intercept $b$ (5 points)

Intercept	47.8
-----------	------

FORMULA =  $b = \frac{\sum y - m \sum x}{n}$

GIVEN =

$$n = 5$$

$$\sum y = 305$$

$$\sum x = 15$$

$$m = 4.4$$

SOLUTION

$$b = \frac{305 - (4.4)(15)}{5}$$

$$= \frac{305 - 66}{5}$$

$$= \frac{239}{5} = \underline{\underline{47.8}}$$

1. REGRESSION EQUATION

### 4. Regression Equation (5 points)

Regression Line	$Y = 4.4(x) + 47.8$
-----------------	---------------------

### 5. Draw the Regression Line using a scatter plot (10 points)

1.  $x = 1$

$$\begin{aligned} y &= 4.4(1) + 47.8 \\ &= 4.4 + 47.8 \\ &= \underline{\underline{52.2}} \end{aligned}$$

3.  $x = 3$

$$\begin{aligned} y &= 4.4(3) + 47.8 \\ &= 13.2 + 47.8 \\ &= \underline{\underline{61.0}} \end{aligned}$$

$x = 5$

$$\begin{aligned} y &= 4.4(5) + 47.8 \\ &= 22 + 47.8 \\ &= \underline{\underline{69.8}} \end{aligned}$$

2.  $x = 2$

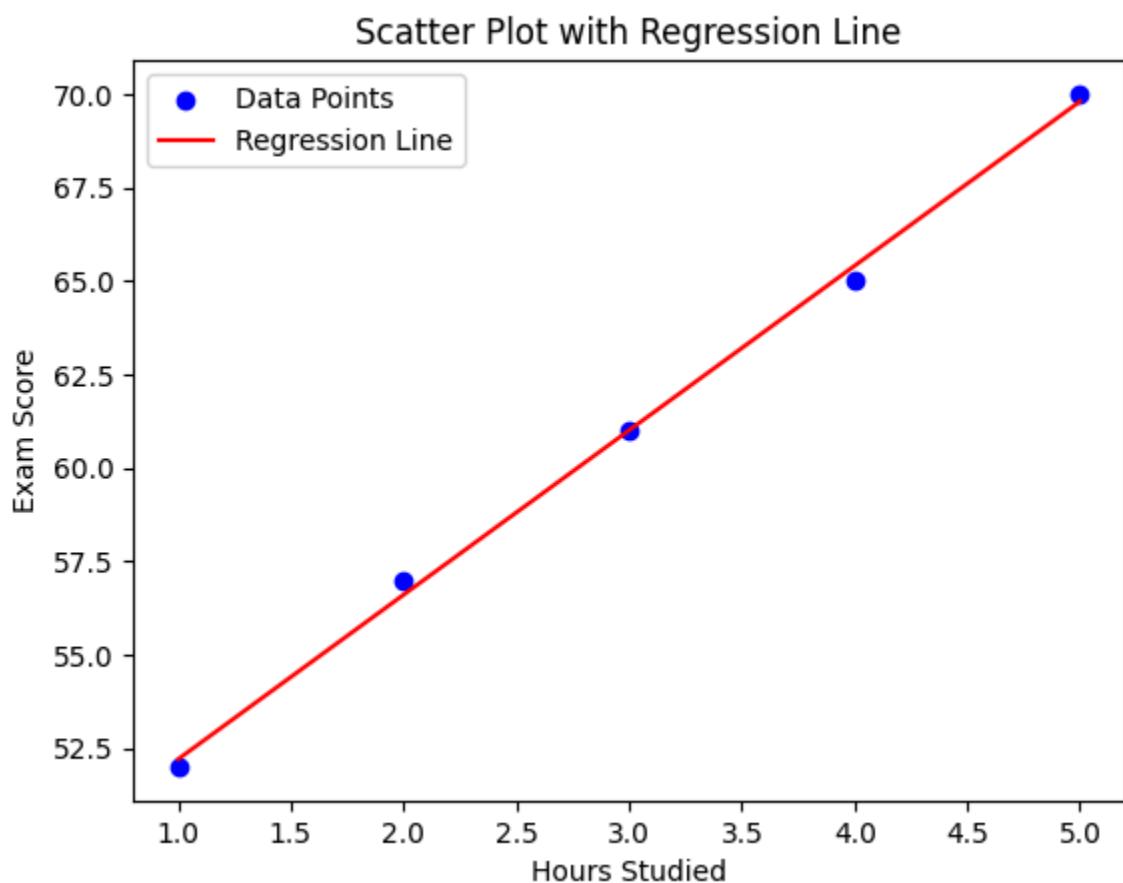
$$\begin{aligned} y &= 4.4(2) + 47.8 \\ &= 8.8 + 47.8 \\ &= \underline{\underline{56.6}} \end{aligned}$$

4.  $x = 4$

$$\begin{aligned} y &= 4.4(4) + 47.8 \\ &= 17.6 + 47.8 \\ &= \underline{\underline{65.4}} \end{aligned}$$

Student	Hours Studied (x)	Exam Score (y)	Predicted Exam Score
1	1	52	52.2

2	2	57	<b>56.6</b>
3	3	61	<b>61.0</b>
4	4	65	<b>65.4</b>
5	5	70	<b>69.8</b>



## 6. Calculate the Sum of Squared Errors (20 points)

SSE	0.4
-----	-----

$$(y_i - y_{\text{PREDICT}}) \ni (y_i - y_{\text{PREDICT}})^2$$

$$\begin{aligned} 1. \quad y_1 &= 52 \\ y_p &= 52.2 \\ &= 52 - 52.2 \\ &= \underline{-0.2}, \underline{0.04} \end{aligned}$$

$$\begin{aligned} 3. \quad y_3 &= 61 \\ y_p &= 61 \\ &= 61 - 61 \\ &= \underline{0}, \underline{0} \end{aligned}$$

$$\begin{aligned} 5. \quad y_5 &= 70 \\ y_p &= 69.8 \\ &= 70 - 69.8 \\ &= \underline{0.2}, \underline{0.04} \end{aligned}$$

$$\begin{aligned} 2. \quad y_2 &= 57 \\ y_p &= 56.6 \\ &= 57 - 56.6 \\ &= \underline{0.4}, \underline{0.16} \end{aligned}$$

$$\begin{aligned} 4. \quad y_4 &= 65 \\ y_p &= 65.4 \\ &= 65 - 65.4 \\ &= \underline{-0.4}, \underline{0.16} \end{aligned}$$

$$SSE = \underline{\underline{0.4}}$$

## 7. Calculate the Sum of Squared Total (20 points)

SST	194
-----	-----

$$\begin{aligned} \bar{y} &= \frac{\sum y_i}{n} \ni SST = \sum (y_i - \bar{y})^2 \\ &= \frac{305}{5} \\ &= \underline{\underline{61}} \end{aligned}$$

$$\begin{aligned} y_i - \bar{y} &\quad (y_i - \bar{y})^2 \\ 1. \quad 52 - 61 &= -9, \quad 81 \\ 2. \quad 57 - 61 &= -4, \quad 16 \\ 3. \quad 61 - 61 &= 0, \quad 0 \\ 4. \quad 65 - 61 &= 4, \quad 16 \\ 5. \quad 70 - 61 &= 9, \quad \underline{\underline{81}} \end{aligned}$$

$$SST = \underline{\underline{194}}$$

**8. Compute R^2 (20 points)**

R^2	0.997 / 0.998
-----	---------------

$$\begin{aligned}\text{FORMULA} = \quad R^2 &= 1 - \frac{\text{SSE}}{\text{SST}} \\ &= 1 - \frac{0.41}{194} \\ &= \underline{\underline{0.997}} \approx \underline{\underline{0.998}}\end{aligned}$$

**9. Prediction (1 point)**

Prediction	74.2
------------	------

$$\begin{aligned}y &= m(6) + b \\ &= 4.4(6) + 47.8 \\ &= \underline{\underline{74.2}}\end{aligned}$$

