NATIONAL UNIVERSITY OF COMPUTER AND EMERGING SCIENCES
(KARACHI CAMPUS)

Department of Software Engineering

QUIZ NO 01

Subject: Machine Learning Date: 17th Sep 2025
Total Marks: 100
Course Instructor: Muhammad Minhal Raza

Instructions to be strictly followed. Attempt All Question's.

Question no 1 : Briefly answer the following questions.

A. Suppose we increased the size of the testing set. Would this likely improve or deteriorate the performance of the model on new data? Why?

Answer: Increasing the test set size does not improve the model itself because the model is already trained. But it makes the performance evaluation more reliable, since more test data gives a better idea of how the model behaves on unseen data. So, the performance number (accuracy, error, etc.) becomes more stable and trustworthy, but the actual model performance does not improve. **In short: It neither improves nor worsens the model but gives a more accurate estimate of its performance.**

B. A data scientist is building a regression model to predict house prices (in $1000s) based on features like size, location, and number of rooms. After testing the model on 5 houses, the predicted prices and actual prices are recorded:

House	Actual Price ($1000s)	Predicted Price ($1000s)
1	200	180
2	150	160
3	300	310
4	250	240
5	400	420

i. Compute the error for each house
ii. Find the Mean Squared Error (MSE).
iii. Calculate the Root Mean Squared Error (RMSE).
iv. Interpret the RMSE in the context of house price prediction (what does it mean in $1000s?).

Solution :

We are given **Actual Price (in $1000s)** and **Predicted Price (in $1000s)** for 5 houses:

House	Actual ($1000s)	Predicted ($1000s)	Error = Actual – Predicted
1	200	180	200 – 180 = 20
2	150	160	150 – 160 = -10
3	300	310	300 – 310 = -10
4	250	240	250 – 240 = 10
5	400	420	400 – 420 = -20

(i) Errors

So, the errors are:

- House 1: **20**
- House 2: **-10**
- House 3: **-10**
- House 4: **10**
- House 5: **-20**

(ii) Mean Squared Error (MSE)

Formula:

$$MSE = \frac{\sum(\text{Error})^2}{n}$$

Square each error:

- $20^2 = 400$
- $(-10)^2 = 100$
- $(-10)^2 = 100$
- $10^2 = 100$
- $(-20)^2 = 400$

Sum = $400 + 100 + 100 + 100 + 400 = 1100$

Now divide by number of houses $n = 5$:

$$MSE = \frac{1100}{5} = 220$$

(iii) Root Mean Squared Error (RMSE)

Formula:

$$RMSE = \sqrt{MSE}$$

$$RMSE = \sqrt{220} \approx 14.83$$

(iv) Interpretation of RMSE

Since prices are measured in **$1000s**,

RMSE = 14.83 means the model's predictions are off by about $14,830 on average.
