LAHORE UNIVERSITY OF MANAGEMENT SCIENCES
Department of Electrical Engineering

EE 514 (CS 535) Machine Learning – Spring 2025
Quiz 1 Solutions

Name: _____

Campus ID: _____

Total Marks: 10
Time Duration: 15 minutes

Question 1 (5 marks)

1. What is the primary goal of supervised learning?
   a) To find patterns in unlabeled data
   b) To predict outputs based on labeled training data
   c) To cluster data into groups
   d) To reduce the dimensionality of data

   Solution: b) To predict outputs based on labeled training data

2. The validation set is used to assess the performance of a model after it has been fully trained.
   ☐ True
   ☐ False

   Solution: b) False

3. What is the main purpose of the train-test split in machine learning?
   a) To ensure the model memorizes the training data
   b) To reduce the size of the dataset
   c) To evaluate the model's performance on unseen data
   d) To increase the complexity of the model

   Solution: c) To evaluate the model's performance on unseen data

4. Which of the following is a key assumption in supervised learning?
   a) The data is linearly separable.
   b) The labels are always binary.
   c) The model must be a neural network.
   d) The training and test data are drawn from the same distribution.

   Solution: d) The training and test data are drawn from the same distribution.

5. Cross-validation is a technique used to simulate multiple train-test splits on the training data.
   ☐ True
   ☐ False

   Solution: True

Question 2 (2 marks)

Explain the difference between supervised and unsupervised learning. Provide one example of each. Use the back of the sheet to provide your answer.

Question 3 (3 marks)

What is the purpose of a loss function? For data formulated as $\mathcal{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)\}$ and $h(\mathbf{x})$ is the hypothesis function (model) that predicts the output for a given input $\mathbf{x}$, formulate RMSE and MAE loss functions used in the regression.

Solution: A loss function quantifies the error between the predicted values $h(\mathbf{x}_i)$ and the true values y_i for a given dataset. Its primary purposes are (any of the following will be considered as correct):

- To measure how well a model is performing by evaluating the difference between predictions and actual values.
- To guide the optimization process during model training by providing a metric to minimize (e.g., gradient descent).
- To help in model selection and hyperparameter tuning by comparing the performance of different models or configurations.

Root Mean Squared Error (RMSE):

$$\text{RMSE}(h) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(h(\mathbf{x}_i) - y_i\right)^2}$$

Mean Absolute Error (MAE):

$$\text{MAE}(h) = \frac{1}{n}\sum_{i=1}^{n}\left|h(\mathbf{x}_i) - y_i\right|$$