## Task 1: MCQS (10 examples for your exam-training)

**Name:** ........................................................        **Duration:** 20 min

| Max. Score | Score |
|:---:|:---:|
| 30 | |

Notes:

- Each question has **only one correct choice.**
- There are **30** questions, each question is worth 1 point.
- Selecting a wrong answer will result in a deduction of 1 point.
  - Correct answer → +1 point
  - Incorrect answer → –1 point
  - Unanswered question → 0 points
- Permitted aids:
  - Two single-sided A4 sheets or one double-sided A4 sheet with your own notes.

---

1. What is the key difference between the "Leaky ReLU" and the "ReLU" activation functions in neural networks?
   a) ReLU sets negative input values to a small, non-zero gradient, while Leaky ReLU sets them to zero
   b) ReLU sets negative input values to zero, while Leaky ReLU allows a small, non-zero gradient for negative inputs
   c) ReLU increases the learning rate, while Leaky ReLU decreases it
   d) ReLU reduces hidden layers, while Leaky ReLU increases them

2. Precision is defined as:
   a) TP / (TP + FP)
   b) TP / (TP + FN)
   c) TN / (TN + FP)
   d) TN / (TN + FN)

3. Data standardization involves:
   a) Transforming features to a range between 0 and 1
   b) Encoding categorical variables into numerical values
   c) Removing outliers from the dataset
   d) Scaling features to have a mean of 0 and a standard deviation of 1

4. How does the One-vs-Rest (OvR) approach handle multi-class classification problems?
   a) It converts the multi-class problem into a single binary classification problem
   b) It converts the multi-class problem into multiple binary classification problems, one for each class.
   c) It merges all classes into a single class and creates a single model to classify data.
   d) It uses a hierarchical structure to merge all classes into a single class and solve it as a binary classification problem.

5. Which of the following statements correctly describes the K-Nearest Neighbors (KNN) algorithm?
   a) KNN constructs a complex model during the training phase.
   b) KNN is a parametric algorithm that assumes a fixed distribution of the data.
   c) KNN stores instances of the training data and performs computation at the time of prediction.
   d) KNN is only applicable to classification tasks, not regression tasks.

6. Which classification metric is suitable when minimizing false positives is the focus:
   a) Recall
   b) Accuracy
   c) Precision
   d) Classification Error

7. Given this confusion matrix, what is the value of the accuracy?
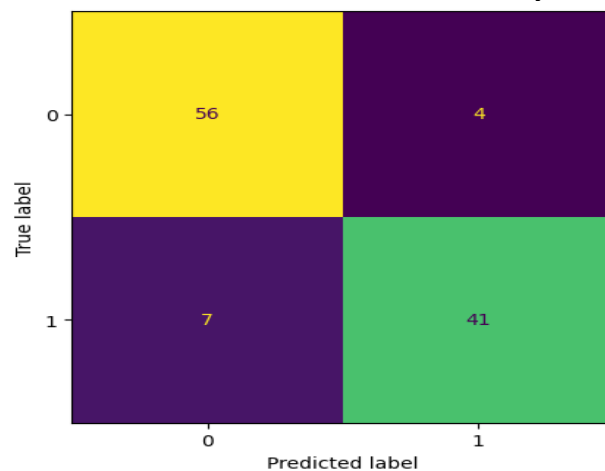


   a) 89.81%
   a) 90.91%
   b) 90.61%
   c) 85.51%

8. Given the following graph, which line does represent gradient decent optimization and which one does represent the Stochastic Gradient Descent (SGD)?
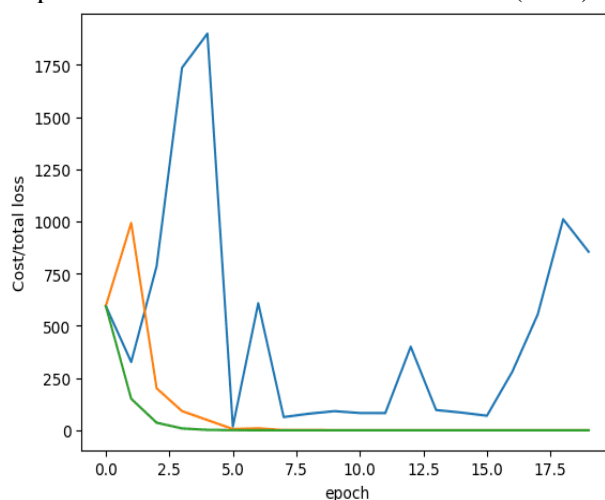


   b) Stochastic Gradient Descent (blue) and Gradient Descent (green)
   a) Stochastic Gradient Descent (orange) and Gradient Descent (green)
   b) Stochastic Gradient Descent (blue) and Gradient Descent (orange)

9. For a standardized dataset with three features, if the eigenvalues of the covariance matrix are $\lambda_1 = 1.5$, $\lambda_2 = 1$, and $\lambda_3 = 0.5$, what is the total variance in the dataset?

   a) 1.5
   b) 2
   c) 3
   d) 3.5

10. Gradient Descent: Weight Update

   If the learning rate ($\eta$) is 0.01, the gradient of the overall loss function with respect to weight $w_1$ ($\partial L/\partial w_1$) is -0.5, and the current weight value is 0.2, what is the updated weight $w_1$ after one iteration?

   a) 0.195
   b) 0.205
   c) 0.215
   d) 0.225