# Machine Learning Exam

(2 hours, 20 points, **2 pages**)

**Instructions:** Answer all questions concisely. No course materials are allowed.

## Exercise 1: Machine Learning Concepts & Applications (4 pts)

1. **(1 pt)** Machine learning can be classified into **three main types**. Name these three types and briefly explain how they differ.
2. **(1 pt)** A company wants to build a **spam detection model** using labeled email data (spam or not spam). What type of machine learning task does this correspond to? Justify your answer.
3. **(2 pts)** Define data leakage in the context of machine learning and explain how it can lead to over-optimistic model performance. Provide an example of a situation where data leakage might occur.

## Exercise 2: Data Preprocessing & Feature Engineering (4 pts)

4. **(1 pt)** A dataset contains missing values in the **"age"** column. Provide one method to handle these missing values and explain its advantages.
5. **(1 pt)** Some machine learning models require numerical inputs. How should we transform a **categorical feature** such as "City" (e.g., Paris, New York, Tokyo) so that it can be used in a model?
6. **(2 pts)** Given the dataset $X = [[5, 20, 5], [20, 5, 20], [10, 30, 10]]$, apply Min-Max scaling.

## Exercise 3: Decision Trees, Ensemble Methods & Model Evaluation (5 pts)

7. **(2 pts)** Decision trees use **splitting criteria** to decide how to divide the data at each node. Name and briefly explain two commonly used criteria.
8. **(2 pts)** After training a decision tree model, you notice that it **perfectly classifies** all training data but performs **poorly on new data**. What issue is occurring? Name **one method** to mitigate this problem.
9. **(1 pts)** A **Random Forest** model is generally more robust than a single decision tree. Explain why, in terms of **variance and bias**.

**Turn the page!**

## Exercise 4: Neural Networks (7 pts)

A **deep learning team** is building a neural network with multiple layers. The team has collected **two input data points**, represented as:

$$X = [[1, -1], [0, 2]]$$

The network has several layers, each represented by the following **weight matrices** and **bias vectors**:

- $W_1 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$, $b_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$
- $W_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$, $b_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $W_3 = \begin{bmatrix} 1 & -2 \end{bmatrix}$, $b_3 = \begin{bmatrix} 0 \end{bmatrix}$

**Activation functions**:

- **Identity function** for the output layer.
- **ReLU** otherwise.

**Questions:**

10. **(1 pt)** Based on the weight and bias matrices given below, determine the total number of neurons in the network.
11. **(2 pts)** Compute $\hat{y}$.
12. **(1 pt)** Explain the usefulness of the learning rate in deep learning.
13. **(1 pt)** Name a method to adapt the learning rate dynamically.
14. **(1 pt)** What is the saturation problem, and what neural network element can cause it?
15. **(1 pt)** Briefly describe how dropout works.