

Introduction to Data Science and Artificial Intelligence - Data Science Part,

Mock exam 2,

Department of Data Science and Knowledge Engineering.

Read each question thoroughly and motivate all of your answers with arguments and/or computations.

This is still a closed book exam. The use of a simple calculator is allowed. The use of graphical calculators, phones, watches and other electrical devices is prohibited. You can earn up to 50 points in this part.

1. **[5 points]** Given the following situation:

A person's blood pressure is measured and the person is given medication in case of hypertension.

- Explain what is data and what is information in the example above.
- Explain how a series of blood pressure values measured over several days could be transformed from data into knowledge.
- Express the knowledge generated at the previous point into a rule which could be implemented into a knowledge based system.

2. **[10 points]** Let $z = 1 + 3i$ and $w = 2 - 3i$.

- Find \bar{z}
- Find $|z|$
- Express z and w in trigonometric form and plot them in the complex plane.

Compute:

- $\frac{\bar{z}}{w}$
- Find $z \in \mathbb{C}$ such that: $z^2 = -1$.

3. **[10 points]** Given the three time series:

$$x = \{2, 1, 2, 0\}, \quad y = \{0, 1, 2, 5\}, \quad z = \{-1, 2, 3, 0\}$$

- Compute the Euclidean distances $d(x, y)$, $d(x, z)$, and $d(y, z)$. What are the most similar time series?
- Given the following distance between two time series z and w :

$$d(z, w) = \sum_{i=1}^N z_i w_i$$

Compute the distances $d(x, y)$, $d(x, z)$, and $d(y, z)$. What are the most similar time series?

- Which one(s) of the first three properties of a distance measure are not fulfilled by the distance given at the previous point?

4. **[10 points]** In a physical experiment four different measurements y_k are obtained over four different time instants

t_k	1	2	3	4
y_k	-1	1	0	1

The physicist who collected those is interested in fitting these measurements to a model of the form: $y_k = at_k + b$

Compute the optimal values for a and b according to the chosen model and the data collected during the experiment.

5. [5 points] Consider the experiment of rolling a dice.

- a. Complete the following table summarizing the probability distribution of this experiment:

Outcome value	Probability	Cumulative probability

- b. Write a rule that would allow you to simulate this experiment with a computer.

6. [10 points] Given the following game:

		Player 2		
		1	2	3
Player 1	Strategy 1	5,1	0,4	1,0
	2	3,1	0,2	3,5
	3	3,1	4,4	2,5

Use both iterated elimination of strictly dominated strategies and the concept of Nash equilibrium to predict the outcome of this game.

Success!