Intelligent Systems, 2022/23, written exam, 01 February 2023
All questions count equally. Literature, electronic and communication devices are not allowed. It is allowed to use one sheet of A4 format paper with notes. You can write your answers in English or Slovene. Duration: 90 minutes.
A possibility to see grading of your exams will take place on Monday, 06 February 2023, at 11:00 in the office of Prof Robnik Šikonja (2nd floor, room 2.06).

1. We have tested several prediction models on a given dataset. While training and validation set errors look promising, the evaluation in the actual industrial setting is far from satisfactory and we suspect that we are overfitting the data. We consider the following approaches to reduce overfitting:
    i) use a smaller training and larger validation set,
    b) switch to a model with more parameters,
    c) for the neural network model, reduce the learning rate during gradient descent,
    d) add a regularization term to the cost function.
    For each of the four proposals describe if and how would they reduce overfitting.

2. A company is working on a multi-label image classification task, and they have a dataset of 10,000 partially labelled relevant images. Due to the high cost of manual annotation, only 2,000 of the images are fully labelled.
    a) Describe two approaches the company could use to make use of the partially labelled data to improve the classification performance.
    b) Provide a pseudocode of each approach.
    c) Explain the benefits and drawbacks of each approach.

3. A student is studying reinforcement learning and is trying to understand the Bellman equations. He is given the following Markov decision process, where S is the set of states, S' the set of next states, and A the set of actions: S = {s1, s2, s3, s4}, A = {a1, a2}, P(S'|S,A) and R(S,A,S').
    a) What are the Bellman equations for the value function and the Q-function in this Markov decision process?
    b) How do these equations relate to each other and how do they help in finding the optimal policy?
    c) Give an example of how the student could use the Bellman equations to update the Q-value for a specific state-action pair in this MDP?

4. A team of data scientists is working on a speech recognition task, and they have encountered a problem with background noise in the audio recordings. They have decided to use autoencoders to remove the noise from the audio before feeding it to the speech recognition model.
    a) Propose an architecture of the noise removal system.
    b) Give a pseudocode for training the system.
    c) Give a pseudocode for using the system.

5. You are tasked with building an object detection system for a self-driving car. Using the provided dataset of images taken from the car's cameras, you decide to train an ensemble of object detection models.
    a) Compare and contrast the performance of boosting and stacking.
    b) Explain the reasoning behind your selection of the applied method.
    c) Discuss the potential challenges and limitations of deploying the ensemble model in a real-world self-driving car system.