DATA.ML.100 Introduction to Pattern Recognition and Machine Learning
TAU Computing Sciences
Exercise - Week 7
Reinforcement learning (Gymnasium)

Be prepared for the exercise sessions (watch the demo lecture). You may ask TAs to help if you cannot make your program to work, but don't expect them to show you how to start from the scratch.

1. **OpenAI Gym – Taxi (V3) Environment** (50 points)

   In this exercise we will use the Gym environment (https://www.gymlibrary.dev/). You may install it via Anaconda (see the course Moodle page). Launch Python and type the following commands:

   ```
   $> python
   >>> import gym
   >>> env = gym.make("Taxi-v3",render_mode='ansi')
   >>> env.reset()
   >>> print(env.render())
   ```

   You should see a map with four locations. Read the description of the map from the source: https://www.gymlibrary.dev/environments/toy_text/taxi/ Use the commands 0-5 and solve the problem manually (render after each step):

   ```
   >>> state, reward, done, truncated, info = env.step(1)
   >>> print(env.render())
   ```

   Your task is to implement Q-learning to solve the Taxi problem with optimal policy. For this you need to fill in the missing parts in the *openai_taxi_skeleton.py* available in the course Moodle page. By default the skeleton creates a random Q-table, you may try:

   ```
   $> python openai_taxi_skeleton.py
   ```

   When you have implemented the training part, then run your method ten times and compute the average *total reward* and the average *number of actions*.

   Return the following items:

   - Python code: **<surname>_taxi.py**
   - A full desktop screenshot that includes a terminal window executing your code and printing the results:
     <surname>_taxi.png