DATA.ML.100 Introduction to Machine Learning

Exercise week 7: Reinforcement learning in Gymnasium

Be prepared for the exercise sessions (watch the demo lecture if available). 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. **Gymnasium – Taxi (V3) Environment** (50 points)

   In this exercise we will use the Gymnasium environment 'Taxi-v3'. You can install Gymnasium using Anaconda (see the lecture notebook). Launch Python and type the following commands:

   ```
   $> python
   >>> import gymnasium as 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://gymnasium.farama.org/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 find an optimal policy for the Taxi problem. You may use the example codes for Q-learning and evaluation in the lecture notebooks, but it is recommend to implement everything from scratch to maximize your learning.

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

   Return the following items:

   - Python code: `taxi.py`
   - A full desktop screenshot that includes a terminal window executing your code and printing the results:
     `taxi.png`