Reinforcement Learning Lab Assignment:

In this assignment, you will work with the PC-Gym framework for Reinforcement Learning (RL) in chemical process control. The **learning goal** is to understand the training metrics and how to tune the hyperparameters of an RL algorithm for improving the final performance of the trained policy.

Resources

- Python 3.14 (recommended uv environment). Use the *project.toml* to install (*pip install .*) the dependencies required to work in the laboratory.
- GitHub repository with the preconfigured Jupyter notebook and imported dependencies.

Task 1: ----------

Goal: Train PPO, SAC, and DDPG policies on the CSTR environment using the default hyperparameters from the Stable-Baselines3 (SB3) implementation and compare its performance against the NMPC oracle available in PC-Gym.

The instructions are specified in the *PCGym_lab.ipynp* Jupyter notebook.

Evaluation:

- Visualize the performance plots. (5 pts.)
- Discuss the performance of the three algorithms in the CSTR environment. (5 pts.)

Optional task: If you like, you can experiment using a different reward function already implemented in the Jupyter notebook and therefore analyze the changes in performance and the possible improvements that should be applied to achieve similar or even better performance than using the previous reward function.

Task 2: ----------

Goal: Add realistic disturbances and constraints to approximate the training environment to a real scenario, closing the sim-to-real gap.

The Jupyter notebook has an environment defined, including disturbances in the temperature and constraints for the inlet temperature. In this task, you must train three new policies (PPO, SAC, and DDPG) using the default hyperparameters for each algorithm, then, after analyzing the training metrics logged in TensorBoard, pick the algorithm with the worst performance and propose a new set of hyperparameters, train the agent again, and compare the final performances as in Task 1.

Evaluation:

- Visualize the performance plots of the algorithm with the worst performance, before and after fine-tuning the hyperparameters. (5 pts.)
- Discuss the hyperparameters tuning process based on the Tensor Board logs. (5 pts.)

Task 3 - after the lab session: ----------

Goal: Study the impact of the reward function on the performance of the RL algorithm.

Using the environment defined in task 1.1, train a policy with one of the studied algorithms, fine-tune it if needed, and compare the policy performance against the NMPC Oracle.

Evaluation:

- Visualize the performance plots. (5 pts.)
- Discuss the selected algorithm. (5 pts.)