Course: Artificial Neural Networks and Deep Learning

Practical assignment of Unit 3
"Deep Reinforcement Learning"
Project Description

In this project, students will use the implementation of the DQN algorithm published on Moodle (file "DQN_Cartpole.ipynb") which is applied to the environment "CartPole-v1". Students will work in groups of four people to perform the following task:

- **Apply the DQN algorithm to an environment of "Gymnasium"**. Students will modify the code of the file "DQN_Cartpole.ipynb" to apply the DQN algorithm to an environment of Gymnasium[1] (different from the Cart Pole environment). For this purpose, students will select one of the following environments: Acrobot, Mountain Car, Lunar Lander, Black Jack, Taxi, Cliff Walking, or Frozen Lake.

As a result of this work, students will submit a Jupyter notebook with the best implementation achieved (only one implementation). The neural model will be implemented using TensorFlow/Keras (specialized software libraries for reinforcement learning, such as Stable Baselines, are not allowed in this project).

Evaluation criteria. The assessment of the work done by students in this project will take into account the complexity of the chosen environment and the code changes to <u>minimize the training time</u> and <u>maximize performance</u>. It is expected that the students will explain how they experimented with different hyperparameter values and will explain how they changed or extended the algorithm in order to minimize the training time and maximize performance.

Content of the notebook. The Jupyter notebook will include the following content:
- A text section will identify the authors at the beginning of the notebook (with name and email address).
- The notebook will include the code of the programs for training and testing. Text descriptions will explain clearly the changes and extensions performed by the students to the original file.
- The notebook will show the result of the last execution with the following information:
    - Total training time,
    - A graphical view of the evolution of the performance (e.g., score) during the training process,
    - After training, the average and maximum performance of n episodes ($n \geq 30$),
    - Characteristics of the computer used for the execution.
- Optionally, the notebook may include a text section describing the list of changes tried by students that produced intermediate unsatisfactory results.

Submission. Students will submit the result of the project as a notebook using the Moodle task **Jupiter notebook (file format .ipynb)**. Only one member of the group should submit the file using Moodle.

---

[1] https://gymnasium.farama.org