Universidade do Minho
Departamento de Informática
Mestrado Integrado em Engenharia Informática
Mestrado em Engenharia Informática

Intelligent Systems
Soft Computing
4°/1° Year, 2° Semester
2020/2021 Edition

Group Practical Work

Theme

Deep Reinforcement Learning

Learning Objectives

With the realization of this practical work, it is intended that groups learn the following procedures:

- Preparation of a Deep Q-Learning Model;
- Analysis and validation of the Reinforcement Learning model performance;

Problem Statement

This practical work intends to be the starting point for the development of a Deep Reinforcement Learning (DRL) algorithm using the Python development environment together with Tensorflow and OpenAI Gym libraries. For this, it will be necessary to develop a solution to the following problem:

"Preparation of a DRL algorithm capable of learning to play the game Breakout from Atari 2600"

In this project, it is intended to apply the knowledge taught during the curricular unit of Soft Computing for the creation of an agent capable of learning to play the game Breakout, through the implementation of DRL algorithm. For this, each group will use the OpenAI Gym library to have access into the respective game environment (named BreakoutDeterministic-v4). On this environment, a layer of bricks lines the top third of the screen and the goal is to destroy them all by repeatedly bouncing a ball off a paddle into them. Each observation obtained from this environment produces an RGB image of the screen, with an array shape of (210, 160, 3).

More information can be found in the OpenAI documentation (for Windows a starting guideline is also available).

Figure 1: Breakout - Atari 2600

Considering the project objectives, the work can be divided into two main phases:

(1) A first phase dedicated to the training of a prototype model, where the agent will train a Deep Q-Network. This model will receive as input a sequence of pre-processed frames (e.g., 3-5 frames) of the current state of the game (state), and output the predicted Q-values for each action;
(2) A second phase dedicated to the testing and optimization of the Deep Q-Learning network.

	The practical work includes the delivery of the developed code and corresponding digital report, describing all the procedures applied and respective justification for its use, based on the demonstration of the results obtained. Additionally, each group can use any open-source libraries that are considered useful for solving the problem at hand (considering the defined development environments).
Delivery	The code resulting from this practical work and respective report (in digital format - PDF) must be sent via elearning platform to the respective evaluation item (found in: **Conteúdo – Instrumentos de Avaliação – Group Practical Work**), in compressed file (ZIP format), **until 4th of June 2021 – 10:00h AM**. The file must be identified in the form "**[CN:TPGXX]**", in which **[GXX]** designates the identification number of the group. The presentation session of the practical work will take place on **7th of June 2021**, in a format to be announced in due course.
Bibliographic References	Saito, S., Wenzhuo, Y., & Shanmugamani, R. (2018). Python reinforcement learning projects: eight hands-on projects exploring reinforcement learning algorithms using TensorFlow. Packt Publishing Ltd. Fan, J., Wang, Z., Xie, Y., & Yang, Z. (2020, July). A theoretical analysis of deep Q-learning. In Learning for Dynamics and Control (pp. 486-489). PMLR. Dong, X., Shen, J., Wang, W., Liu, Y., Shao, L., & Porikli, F. (2018). Hyperparameter optimization for tracking with continuous deep q-learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 518-527).