CS6482 Deep Reinforcement Learning

Assignment 2: Sem1 AY 24/25 – DQN for Atari

11/April/25 (Teaching Week 11) – v2.

1. Objectives

To implement a Reinforcement Learning (RL) agent using a Deep Q Network (DQN) applied to the game of Atari in the OpenAI Gym environment.

Please email the lecturer with a request for permission to use an alternative to Atari if committed to a different simulator and/or environment. Provide a link to the alternative and a paragraph stating the rationale. The subject must be "CS6482: ALTERNATIVE"

2. Submission

Submit a **Jupyter Notebook with commented code and the following textual descriptions:**

1. Why Reinforcement Learning is the ML paradigm of choice for this task **(2 marks):** [Text Block]
2. [Text Block] The Environment **(1 mark)**: [Code, Output, and Text blocks]
    a. the Atari game selected,
    b. the inputs received from the Gym environment, and
    c. the control settings for the joystick.
3. VANILLA DQN Implementation **(6 marks)**: [Code, Output, and Text blocks]
    a. Text Block: capture and pre-processing of the data (2 marks),
    b. Text Block: the network structure (2 marks),
    c. Text Block: the Q learning update applied to the weights (2 marks),
4. Results and EVALUATION for VANILLA DQN **(4 marks):** [Code, Output, and Text blocks]
    a. Plots with short accompanying explanations of the information conveyed.

b. How does one evaluate the performance of the RL agent?
c. Is the agent learning?

5. DOUBLE DQN implementation **(4 marks)**: [Code, Output, and Text blocks]
    a. Describe the problem being addressed (2 marks)
    b. Describe the theoretical solution. (2 marks)
6. Results and EVALUATION for DOUBLE DQN **(3 marks)**: [Code, Output, and Text blocks]
7. Comparison of DQN versus DOUBLE DQN **(3 marks)**
8. ADDED VALUE – implemented and evaluated **(4 marks)**: [Code, Output, and Text blocks]
    a. Independently research concepts excluding Double DQN. For example, published hyperparameter settings that appear particularly affective,
    b. Prioritised Experience Replay,
    c. Discussions in published papers or blogs from reputable sources that provide additional insight into the results – implementation may not be required for this one.
9. References (P/F)

The Jupyter notebook must (a) include output i.e. code cell action outputs, and (b) be extensively commented. Failure to comply with one or both of these requests will result in the award of 0 marks for the submission.

Submit a **Jupyter notebook** with the code where:

- The book is named CS6482-Assign2-ID for an individual submission OR CS6482-Assign2-ID1-ID2 for a team-based submission
    - Where ID is the student id number
- The first line in the book is a comment with name(s) and ID number(s) of the student(s).
- The second line in the book should be a comment stating if the code executes to the end without an error.
- The third line in the book should be a comment with a link to the original source where you opted to reuse an existing implementation.

3. Notes and Guidelines

- This assignment **constitutes 27%** of the total marks awarded for this module.
- You may complete this assignment in a team of 2 OR on an individual basis.

- Please email me team details should you opt for a team-based submission.
- Teams must provide a code inspection following submission if requested to do so. Failure to comply results in an F grade for this assignment.
- Only ONE single submission per team please.

- **Submission deadline is 23:59:59 Wednesday 14th May 2025 (Teaching Week 15).**
- Submission is via the Brightspace Assignment tool.
- Programming language is Python.

GRADING RUBRIC

Area	Beginning [0-7]	Developing [8-15]	Accomplished [16-21]	Exemplary [22-27]
Code	Copy and Paste No comments Does not run to completion	Commented satisfactorily, using a few sources Runs to completion	Comments are good Using many sources Runs to Completion	Excellent commentary. Using many sources. Some proprietary implementation. Runs to completion.
Textual Descriptions	Basic, Only 1 or 2 references that are not cited . Communicates an absence of understanding of DQNs and Policy Gradient (PG) approaches. Deviates from spec. Conveys impression of a late start i.e. rushed	Satisfactory A few references cited appropriately Communicated some understanding of DQNs and PG approaches. Adheres to spec	Good Multiple references discussed and cited Communicates solid understanding of DQNs and PG approaches. Adheres to spec	In depth Many relevant references cited Communicated deep understanding of DQNs and PG approaches. Adheres to spec
Results	Trivial Not analysed	Basic plots and basic analysis	Basic plots that are analysed in depth.	Managed to get substantial results within the timeframe Deep analysis
Recent Advances in DQN	Absent	Using 1 reference only Implementation present that does not work, Results not presented	Using a few references Implementation present and runs to completion, Results are trivial.	Using many references Implementation present and runs to completion, Results presented and analysed