REPORT

Name: Harshit Singh

Roll No: 2001cs29

The code implements and compares two reinforcement learning agents, Double DQN and Dueling DQN, within the CartPole-v1 environment using TensorFlow and OpenAI Gym.

Libraries and Frameworks

The code utilizes various libraries and frameworks:

1. **TensorFlow**: Deep learning library for creating neural network architectures.
2. **OpenAI Gym**: Provides reinforcement learning environments like CartPole.
3. **RL (Reinforcement Learning) Library**: Specifically, the agents and memory modules (DQNAgent, SequentialMemory) are used for reinforcement learning tasks.

Model Architectures

- Double DQN Model:
    - Architecture: Utilizes a standard feedforward neural network via a Sequential model.
    - Architecture Details: Consists of multiple fully connected layers with a final output layer predicting actions.
- Dueling DQN Model:
    - Architecture: Utilizes a custom neural network model created using the Keras functional API.
    - Architecture Details: Implements a dueling network, separating value and advantage functions via two

parallel branches, ultimately combining them in the output layer.

Hyperparameters

- Training was conducted across **100,000 episodes**.
- Exploration rate: 0.01
- Learning rate: 1e-3
- Discount factor: 0.99

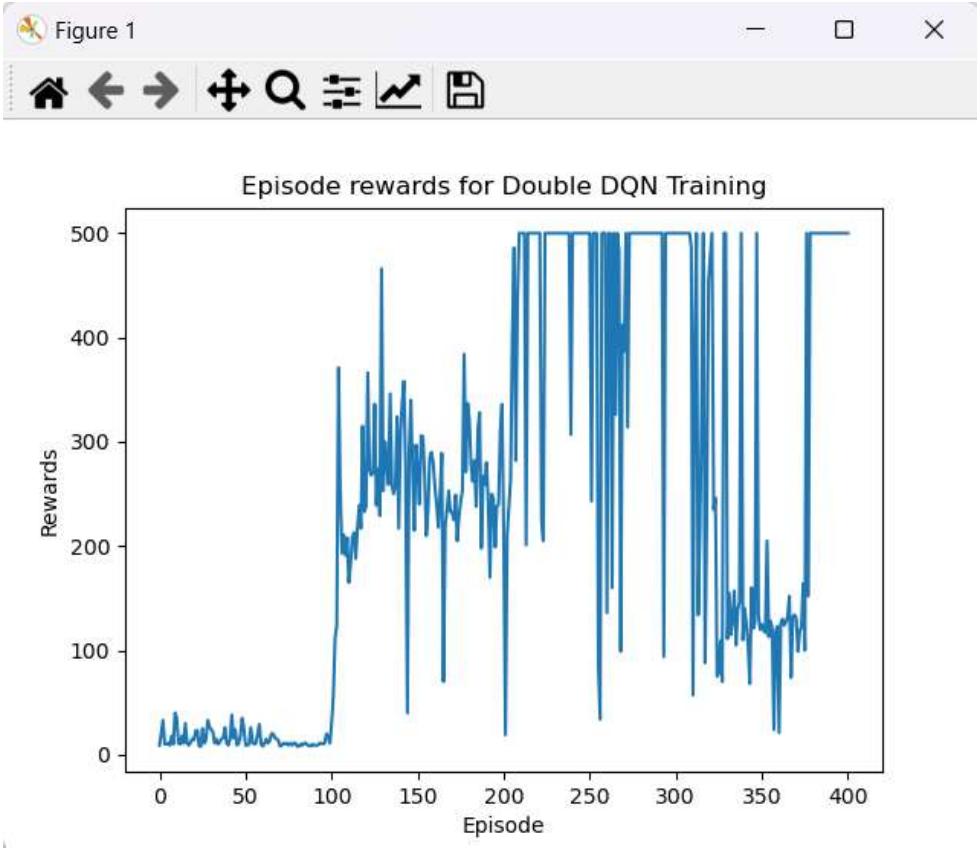
Training and Testing

- Double DQN Agent:
    - Trains a Double DQN agent in the CartPole environment for 100,000 steps.
    - Visualizes the training and tests the agent in 10 episodes.
    - The rewards obtained in testing are plotted.

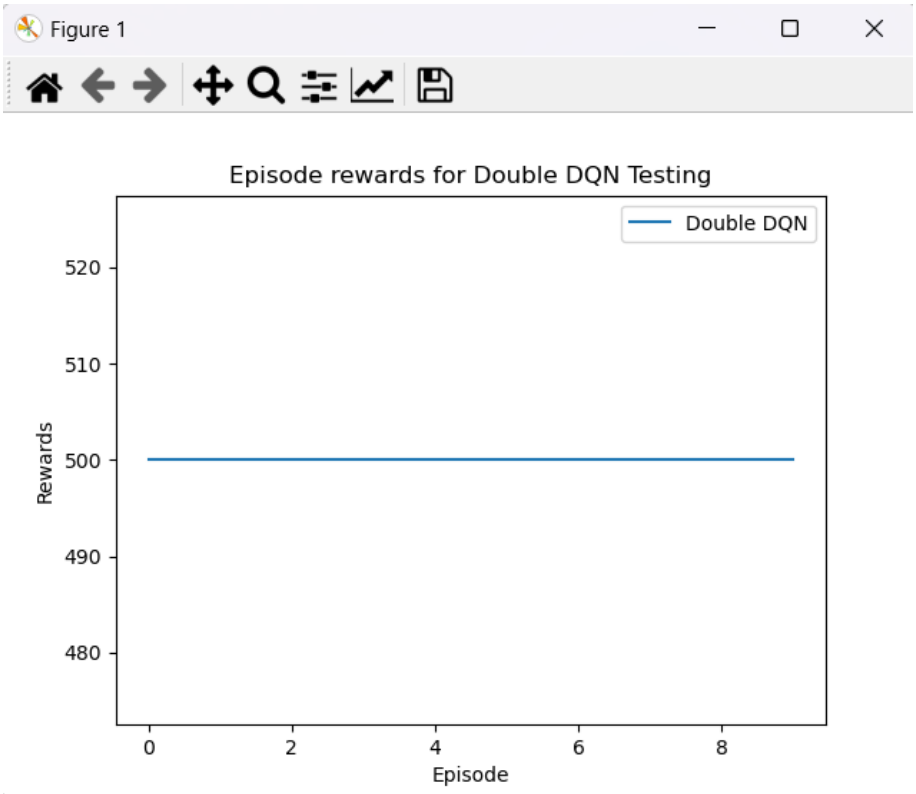
- Dueling DQN Agent (Commented Out):
    - The code for the Dueling DQN agent is provided but is currently commented out.
    - Similar training, testing, and plotting procedures are outlined but not executed.

Results Visualization

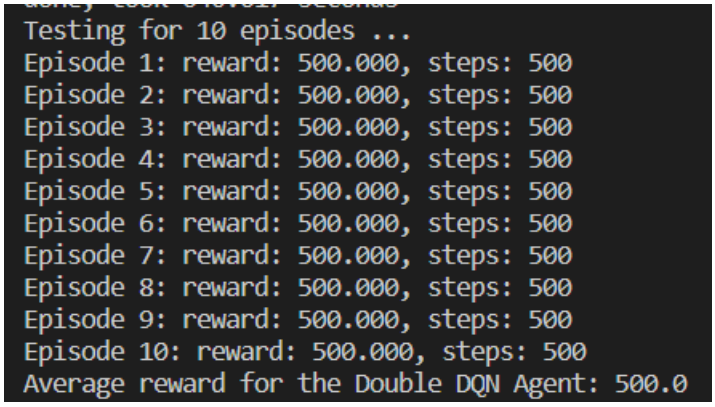
- Episode Rewards Plot:
    - Plots the rewards obtained during the testing phase for the Double DQN agent.
    - We can see in the following graph that initially there is no reward and after learning the agent is able to mostly balance the pole.



- This graph shows the reward of the trained agent, an episode is of length 500 moves and we can see that it is able to survive the entire episode.



- The following is the rewards of the trained agent playing the cartpole game.



Conclusion

The implemented code trains and tests two types of DQN agents in the CartPole environment. The Double DQN agent's

training and testing are executed and visualized, while the Dueling DQN agent's code is provided but not actively

executed. Results from the Double DQN agent testing are displayed via reward plots.