Agent Jackie RL: DQN-Based Atari Agent Report

Project Name: Agent Jackie RL
Student Name: Srinithya Devraj
Roll Number: 24B0997
Mentor: Ananya Kunisetty, Jalajakshi, Deepthika

1 What Worked Well

Our DQN agent, Agent Jackie, was successfully trained on the `KungFuMasterDeterministic-v4` Atari environment using experience replay and an ϵ-greedy strategy. The final policy showed strong performance, with the agent frequently achieving high rewards and demonstrating reactive, learned behavior. Key aspects that contributed to success:

- Proper pre-processing (grayscale conversion, resizing, and frame stacking).
- Gradual ϵ decay with a factor of 0.998 encouraged exploration early and exploitation later.
- Experience replay started after 300 steps, allowing the buffer to diversify before learning.
- Replay occurred every 4 steps, balancing speed and learning frequency.

2 What Didn't Work

Several technical challenges were encountered during the project:

- TensorFlow 2.19 required downgrading `numpy` to 1.24.3 due to compatibility issues.
- Environment rendering errors and 'np.bool8' deprecation in NumPy 2.x led to runtime crashes.
- Missing environment `KungFuMasterDeterministic-v4` prompted reinstallation using `AutoROM`.
- The training process was extremely slow due to CPU-only execution: 1000 episodes took 48 hours.

3 Agent Jackie's Improvement

The agent initially performed poorly, showing random or ineffective actions. Key improvements:

- After about 200–300 episodes, the agent began showing signs of pattern learning.
- Switching to delayed replay (every 4 steps) and using a buffer warm-up threshold helped learning.
- Final recorded performance (episode 900) achieved a reward of **17,200**, a clear improvement over early runs with rewards as low as 200–500.

4 Training Summary

- Total Episodes: 1000
- Total Training Time: ∼48 hours
- Final Model Saved: `model_episode_900.h5`
- Final Video: `videos/agent_play.mp4`

5 Screenshot

Agent Jackie executing learned policy mid-episode.