

Implemented Double DQN and a simplified prioritized replay (prioritizing positive reward experiences).

Experimented with both Polyak averaging and target full copy.

Final Score - ~40

Learning Curves (Score vs Episode No) :-

