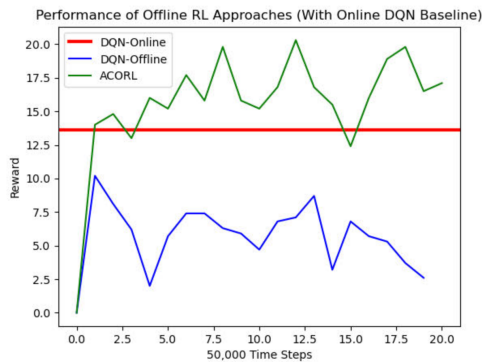# Meeting Notes:

November 29:

- Discussed current results, work to be done over Christmas break, and what the final poster should look like.

Performance of Offline RL Approaches (With Online DQN Baseline)

-

November 15:

- Discussed network architecture and hyperparameter tuning.
- Discussed how to design a fair comparison of approaches.

November 8:

- Discussed neural network parameter sharing between policy learned through imitation learning and the q-network.

November 1:

- Discussed which experiments to run, i.e., which environments to compare against.

October 18:

- Discussed which current state of the art approaches to compare against.
- Discussed current progress in implementing algorithm.

October 11:

- Discussed progress in implementing the algorithm.
- Discussed issues with memory, compute, and time constraints.

September 27:

- Discussed algorithm design and convergence analysis.

Algorithm 1 Action Restricted Offline RL

1: **Input:** D a dataset of trajectories, $\{(s_i, a_i, s_i', r_i)\}$
2: Initialize the cloned policy, a Q-function, and target Q-function neural networks, π_ϕ, Q_θ, and $Q_{\theta'}$
3: **while** not converged **do**
4:      Sample a mini-batch, $\mathcal{B}$, for transitions (s, a, s', r) from D
5:      $\phi \leftarrow \phi - \nabla_\phi[-\sum_{(s,a) \in \mathcal{B}} \log(\pi_\phi(a|s))]$
6: **end while**
7: Define $\mathcal{V}_s = \{a | \pi_\phi(a|s) > \epsilon\}$
8: **while** not converged **do**
9:      Sample a mini-batch, $\mathcal{B}$, for transitions (s, a, s', r) from D
10:      $\theta \leftarrow \theta - \nabla_\theta[\sum_{(s,a,r,s') \in \mathcal{B}} ||r + \gamma \max_{a' \in \mathcal{V}_s} Q_{\theta'}(s', a') - Q_\theta(s,a)||_2^2]$
11:      **if** target update iteration **then**
12:          $\theta' \leftarrow \theta$
13:      **end if**
14: **end while**

-
- Discussed the q-learning convergence proofs.

September 20:

- Discussed my current algorithms design and possible improvements.

September 14:

- Discussed state of the art in Offline RL and current approaches CQL, Cal-QL, Implicit Q-Learning, etc.

September 7:

- Discussed scope of project and what success would look like.

September 1:

- Discussed project ideas: RL of Mario Kart, new approaches for Offline RL