Theory

1. **What is different about reinforcement learning and supervised learning? What is the same?**
   Differences: the targets (rewards) are sparse rather than every input having a target output. This means that we need techniques for valuing an action e.g. how does the total rewards for this episode (or time step/s) compare to total rewards on average or that we expected (policy gradients). Or how does what happened compare to what we expected to happen (Q-learning).
   Similarities: the underlying architectures used to model the patterns in the data are the same, e.g. a CNN for predicting a move from an image in a game, it's just the target and loss that are different.
2. **What is the difference between Q-Learning and Policy Gradients?**
   Q-Learning aims to learn the Q value function to then infer the optimal policy. Policy Gradients aims to learn the optimal policy directly.

Discussion

3. **Do you think human level AGI can be created from our current DL algorithms or improvements to them? Do you think they can become self aware?**
   Note there is no correct answer here, but something to keep in mind is that all the formulas that we use in DL are also used for other purposes e.g. calculus is used in engineering, building etc. the value formulas in RL are identical to those used in finance for valuing future payments, calculating loan repayments etc. So if we're thinking that these formulas are going to become self-aware, why are we not concerned about our engineering diagrams and loan contracts becoming self aware?