Learning to Play Blackjack using Deep Reinforcement Learning

## Chris Durbin

## Introduction

The purpose of the project is to attempt to use deep reinforcement learning to learn optimal play in the game of Blackjack. By counting cards and properly sized bets players have been able to gain an advantage over casinos and have a positive expected reward. In the project I will attempt to entirely use reinforcement learning algorithms to achieve the same. My hypothesis is that it is possible to have positive expected reward entirely using deep reinforcement learning without any programmed knowledge of optimal play nor using the Kelly criterion to determine bet size.

• Student clearly defines the research topic and project goals. Describes the rough research and/or development approach, hypotheses

• What ML problem solved? What method analysis applied?

## Related Work

Prior research projects have …

1. Original Atari DQN paper (Minh, 2013) – shows how to replace the Q table with a neural network and calculate a target to use for error calculations for backprop. Shows it can learn many Atari games – no game specific knowledge in the DQN agent.
2. Deadly Triad from Sutton – convergence is not guaranteed when using function approximation, off-policy learning, and …
3. DDQN (Dueling DQN paper) – shows how to use two networks freezing the target network during the experience replay learning

• Correct approach or convincing novel approach to ML problem

• Theoretical research has merit and complete

## Methods

In order to…

• Algorithm or method description, math behind it (if applicable) and code are sound

• Sufficient demonstration that student knows his/her applied ML method

### Metrics

## Results

The following…

Show a chart with expected return.

Show learning rate over number of episodes.

• Sound solution with metrics, correctly doing the ML method

• Proper runs and statistics collections

• Proper association between problem and solution shown

## Conclusions

I would recommend further…

• Relevant conclusion with support from the method sections

• Answers "why we have conducted this work "

## References

1. Riedmiller, Martin. "Neural fitted Q iteration–first experiences with a data efficient neural reinforcement learning method." *Machine Learning: ECML 2005: 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005. Proceedings 16*. Springer Berlin Heidelberg, 2005.
2. Sutton, Richard S., and Andrew G. Barto. *Reinforcement Learning, second edition: An Introduction*. MIT Press, 2018.
3. Schaul, Tom, et al. "Prioritized experience replay." *arXiv preprint arXiv:1511.05952* (2015).
4. Wang, Ziyu, et al. "Dueling network architectures for deep reinforcement learning." *International conference on machine learning*. PMLR, 2016.
5. Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." *arXiv preprint arXiv:1312.5602* (2013).