



INTRODUCTION

Our team is interested in applying machine learning techniques to solve the game Connect Four.

Research Question: *How do we apply machine learning techniques to play a game of Connect Four?*

Motivations:

- Research on applying AI to game-playing has surged recently with strong development in many domains.
- Connect Four is a suitable game to experiment with different algorithms, as it is more complex than Tic-Tac-Toe, but has smaller state space than Chess and Go



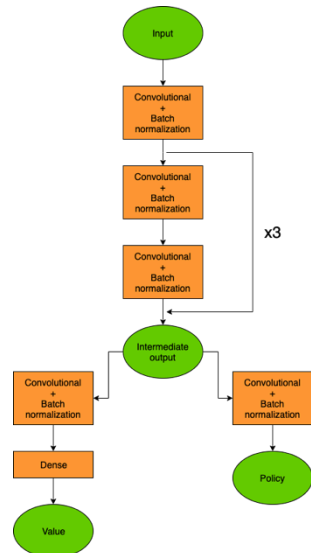
DATA DESCRIPTION

- A dataset from Kaggle. Each row contains end results of the game. Column 1-42 contains position of the board and column 43 contains the winner.
- A dataset from UIC with 67557 board positions after 8 moves. The output is win/loss/draw for the first player.
- A dataset generated by running randomly generated board positions on the solver. The data contains the board positions, policy distribution, the winner and the number of moves left for the game to end.
- A dataset generated with a pre-trained deep reinforcement learning model. It has 2500 full Connect4 games with policy and value output for all board positions.

PRELIMINARY MODELS

- We implemented the Negamax algorithm and the Minimax algorithm with Alpha-Beta pruning to play the game.
- We also ran the following classical machine learning methods:
 - Logistic Regression
 - Support Vector Machines
 - Gradient Boosted Trees
 - Linear Regression

DEEP RL



- We implemented a simple version of the AlphaZero algorithm² that trains the neural network on the left with the Monte Carlo Tree Search algorithm.
- We ran an experiment to compare the playing strength of the model between the no supervised learning implementation and with supervised learning initialization one.

RESULTS

Model	Test Set Accuracy
Logistic Regression	84.44%
SVM with linear kernel	84.44%
GBM	93.34%

- The RL model with supervised initialization converges quickly and hardly sees improvement over its training life, as opposed to the model with no initial supervision.
- With limited training time, the initially supervised model is also significantly stronger.
- Both are weaker than a human player.

DISCUSSION

- Negamax solver without any pruning was very slow to output a solution for an 8-ply board. Minimax AI with alpha-beta pruning ran quickly, only taking a few seconds for the AI to make its move.
- Gradient boosted decision trees does quite well with classifying the outcome of the board
- The RL model clearly underfits and the comparison will be fairer with a more complex model.
- For simpler models, initial supervision has a huge beneficial effect.

REFERENCES

1. <https://www.amazon.com/Connect-Strategy-Board-Amazon-Exclusive/dp/B06XY881H4>
2. "A General Reinforcement Learning Algorithm That Masters Chess, Shogi and Go through Self-Play," *Science*, 7 Dec, 2018. <https://science.sciencemag.org/content/362/6419/1140>