Research Paper Review

Artificial Intelligence and Specializations Nanodegree Program

Yoon Namkung

Article Title: Mastering the game of Go with deep neural networks and tree search

1. Brief summary and goal of this paper

This article suggests very powerful winning algorithm in terms of playing the game of Go by using neural networks with a new approach: 'value network' and 'policy network'. These networks use both supervised learning from data of human player moves and reinforcement learning from self playing game. Finally Alpha go uses Monte Carlo tree search (MCTS) to combine policy and value network.

First, Alpha Go uses supervised learning (SL) policy neural network for predicting best possible move in the game. Through this step, it updates learning with fast feedback and good gradient values. Fast policy can be used for quick sampling of actions during rollouts.

Second, Alpha Go train a reinforcement learning (RL) policy network to improve SL policy network. This helps the network to wind the game instead of increasing prediction accuracy.

Lastly, Alpha Go uses RL policy network by self-playing to train a value network which predict the winner of game. Also, it combines the policy and value network by using MCTS.

2. Brief summary of results

Research Paper Review

Alpha Go achieved remarkable result in the game of Go. Two versions of Alpha Go were used , single machine and distributed version with multiple machines. The single machine Alpha Go won 494 out of 495 games against opponent Go program. The distributed version won 77% against single machine version of Alpha Go and won 100% against all other competent programs. Also, distributed version of Alpha Go has achieved 5 games to 0 win against human professional Go player and European Go champion, Fan Hui.