# Mastering the game of Go with deep neural networks and tree search
# Research Review

Abhi Ojha

March 16, 2017

## 1 Goals

The game of Go when played on a $19 \times 19$ board has 35 legal moves per position, $b = 35$, and a game depth of 80, $d = 80$. Due to such complexity, an exhaustive search is infeasible. The goal of this paper is to introduce a new approach that uses deep learning in to evaluate board positions and select moves. The deep neural networks are trained by supervised learning from human expert games, and reinforcement learning from games of self play. A new search algorithm has also been introduced, which combines Monte Carlo simulations with value and policy networks.

The neural networks are trained using a pipeline consisting of several stages of machine learning. At the first stage, a supervised learning policy network is trained directly from human moves. Next, a reinforcement learning policy network is used to optimize the final outcomes of games of self play. Finally, a value network is trained that predicts the winner of games played by the RL policy network against itself. AlphaGo combines policy networks and value networks with Monte Carlo tree search.

## 2 Results

With the help of AlphaGo, DeepMind's team has proven that Artificial Intelligence agents can achieve human level performance even in games with massive search space. This research further opens up possibilities for using AI in fields such as personalized medicine.

Some of the prominent results from this paper are as follows:

- AlphaGo dominates other computer programs by a 99.8 % winning rate.
- It can also be implemented as a distributed version, which uses 40 search threads, 1202 CPUs and 176 GPUs.
- All the previous models work on Monte Carlo search tree schemes. AlphaGo gives us a completely novel approach to tackle these challenges.