# AlphaGo Analysis

The AlphaGo team set out to conquer the game of Go and in doing so they achieved a 99.8% winning rate against other Go programs and defeated the human European Go champion by 5 games to 0, a feat previously thought to be at least a decade away. How was the AlphaGo team able to do this? They implemented a new innovative approach to computer Go that utilized value networks, policy networks and a new search algorithm that combines Monte Carlo simulation with value and policy networks.

**Supervised learning of policy networks**
The first stage of the training pipeline consists of a SL policy network used to predict expert moves in the game. The SL policy network is a 13 layer CNN trained from 30 million positions from the KGS Go server and was capable of predicting expert moves with an accuracy of 57% using all input features and 55.7% using only raw board positions and move history as input. While other research groups achieved an accuracy rate of 44.4%, this small improvement in accuracy led to large improvements in playing strengths.

**Reinforcement learning of policy networks**
The second stage of the training pipeline aims at improving the policy network by policy gradient reinforcement learning (RL). The RL policy network is structured identically to the SL policy network. When played head-to-head, the RL policy network won more than 80% of games against the SL policy network. The RL policy network was also tested against the strongest open-source Go program, where it won 85% of the games. A significant increase over previous state-of-the-art convolutional networks, which won 11% of games.

**Reinforcement learning of value networks**
The final stage of the training pipeline focuses on position evaluation. This neural network has a similar architecture to the policy network, but outputs a single prediction instead of a probability distribution.

**Searching with policy and value networks**
AlphaGo combines the policy and value networks in Monte Carlo Tree Search (MCTS) algorithm that selects actions by lookahead search.

**Evaluating the playing strength of AlphaGo**
The team evaluated the AlphaGo by testing it against some the strongest commercial and open-source program available. These programs were all highly performant MCTS based algorithms. The results were outstanding, with a single machine AlphaGo winning 99.8% of games against other Go programs and the distributed version was significantly stronger. It was the distributed version that beat Fan Hui the professional Go player, 5 games to 0.

In conclusion, the AlphaGo program consists of a combination of deep neural networks and tree search that was capable of accomplishing a feat previously thought to be impossible, beating a professional human Go player, and marked a significant step towards the advancement of artificial intelligence.