

### Exercício Extra - Documentário AlphaGo

O documentário mostra o desenvolvimento e a aplicação de um algoritmo de aprendizado por reforço, construído pela empresa britânica DeepMind, que foi construído para jogar Go, um jogo chinês de tabuleiro.

Antes de enfrentar o melhor jogador de Go de todos, Lee Sedol, a equipe queria melhorar o seu algoritmo, e para isso pediu para um outro jogador profissional jogar diversas vezes contra a máquina a procura de erros. Com essa análise, olhando os resultados, foi possível encontrar problemas e possíveis melhorias para o algoritmo.

O algoritmo foi construído tendo 3 componentes principais:

1. Policy Network: onde foi treinado em diversos jogos de profissionais, e tenta imitar as jogadas que foram feitas.
2. Value Network: olhando a posição do tabuleiro, ele calcula a probabilidade de vitória de cada jogada.
3. Tree search: olha diferentes variações do jogo para tentar prever o que pode acontecer no futuro.

É possível perceber que, nesse caso, a aplicação de um algoritmo de aprendizado por reforço faz muito mais sentido do que a utilização de outras técnicas de aprendizado de máquina, pois o ser humano ainda não chegou em um nível tão alto de Go para podermos nos basear em exemplos, dessa forma precisamos que o algoritmo consiga aprender sozinho, e decidir por ele mesmo a melhor tomada de decisão, sem se basear 100% em casos que já aconteceram.

No final, o algoritmo é bem sucedido em derrotar Lee Sedol no match quatro partidas a um, mostrando o potencial desse método.