

— Background:

Q-learning

need to read

MCTS

DQN

— Policy network:

- Flatten pics to vectors
- Stochastic policy

















