RL Training: Cumulative Reward vs Episode

Break-even

Exploration Learning Convergence

Random (baseline)
Q-Learning
Deep Q-Network

Y-axis: Average Reward (per episode)
X-axis: Episode

github.com/joerg-osterrieder/Methods_and_Algorithms