gradient-free, model-free methods (stochastic search, evolutionary) policy gradient methods (REINFORCE) actor-critic methods (PPO, TRPO, DDPG, etc.) deep value iteration methods (DQN, etc.) shallow value iteration methods (tabular Q-learning, etc.)

fast

wall clock time

slow

model-based RL