

Algorithm	Description	Model Use	Evaluation	Policy	Method
MBPO	Model-based Policy Optimization (Janner et al. 2019)	Model-based	On-policy	Deterministic	Policy Gradient
MCTS	Monte Carlo Tree Search (Sutton and Barto 2018)	Model-based	Either	Either	Any
A2C	Advantage Actor Critic (Mnih et al. 2016)	Model-free	On-policy	Stochastic	Actor-critic
A3C	Asynchronous A2C (Babaeizadeh et al. 2017)	Model-free	On-policy	Stochastic	Actor-critic
TRPO	Trust Region Policy Optimization (Schulman, Levine, et al. 2017)	Model-free	On-policy	Stochastic	Policy gradient
PPO	Proximal Policy Optimization (Schulman, Wolski, et al. 2017)	Model-free	On-policy	Stochastic	Actor-critic
TDQN	Deep Q Networks (Mnih et al. 2015)	Model-free	Off-policy	Deterministic	Value-based
DDPG	Deep Deterministic Policy Gradient (Lillicrap et al. 2019)	Model-free	Off-policy	Deterministic	Actor-critic
TD3	Twin Delayed DDPG (Fujimoto, Hoof, and Meger 2018)	Model-free	Off-policy	Deterministic	Actor-critic
SAC	Soft Actor Critic (Haarnoja et al. 2018)	Model-free	Off-policy	Stochastic	Actor-critic
IMPALA	Importance Weighted Actor Learner (Espeholt et al. 2018)	Model-free	Off-policy	Stochastic	Actor-critic