Value-Based

- SARSA
- DQN: Deep Q-Networks
- HER: Hindsight Experience Replay
- QT-OPT

Policy-Based

- REINFORCE

Model-Based

- iLQR: Iterative Linear Quadratic Regulator
- MPC: Model Predictive Control
- MCTS: Monte Carlo Tree Search

Combined Methods: Value & Policy

- Actor-Critic: A2C, A3C
- TRPO: Trust Region Policy Optimization
- PPO: Proximal Policy Optimization
- SAC: Soft Actor-Critic
- ACER: Actor-Critic with Experience Replay
- ACKTR: Actor-Critic Using Kroenecker-Factored Trust Region

Combined Methods Model + Value and/or Policy

- Dyna-Q / Dyna-AC
- AlphaZero
- I2A: Imagination Augmented Agents
- VPN: Value Prediction Network