# RL Tool Report

Baher Nader

December 29, 2025

## 1 Algorithms Implemented

The tool implements six RL algorithms: Policy Iteration (model-based DP for policy evaluation/improvement), Value Iteration (model-based DP for optimal values), Monte Carlo (model-free episode-based learning with first/every-visit variants), SARSA (on-policy TD control), Q-Learning (off-policy TD control), and n-step TD (multi-step TD generalization).

## 2 Environments Implemented

Four environments are supported: GridWorld (discrete grid navigation), FrozenLake (stochastic grid with holes), CartPole (discretized continuous balance task), and MountainCar (discretized continuous momentum task).

## 3 Parameter Adjustment Capabilities

Users can adjust parameters via UI sliders and API: gamma (discount factor), alpha (learning rate), epsilon (exploration rate), theta (convergence threshold), n (steps for n-step TD).

## 4 Visualization Techniques Used

Visualizations include: Canvas for environment states, agent paths, policy arrows, value colors; Plotly.js charts for rewards, lengths, convergence; Heatmaps for value functions. Real-time updates via API polling.