SELF-DRIVING TAXI: REINFORCEMENT LEARNING ALGORITHM COMPARISON

Team Members:

Sapna Baniya

Bipin Puri

Course: Artificial Intelligence

Project Objectives

- Compare 3 fundamental RL algorithms on identical navigation task.
- Implement: Policy Iteration, Value Iteration, and Q-learning.
- Analyze performance differences: speed, efficiency, optimality
- Visualize algorithm decision-making through animations

Environment Design

Grid World: 6×6 environment with obstacles

States: 144 states (position × passenger status)

Actions: 4 movements (Up, Down, Left, Right)

Rewards:

Success delivery: +20

Invalid moves: −1

Each step: −0.1

Wrong drop-off: −10

Algorithms Implemented

Three RL Methods:

Policy Iteration

- Model-based, guaranteed convergence
- Policy evaluation + improvement cycles

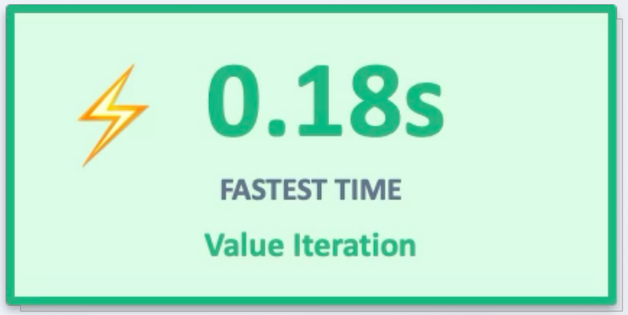
Value Iteration

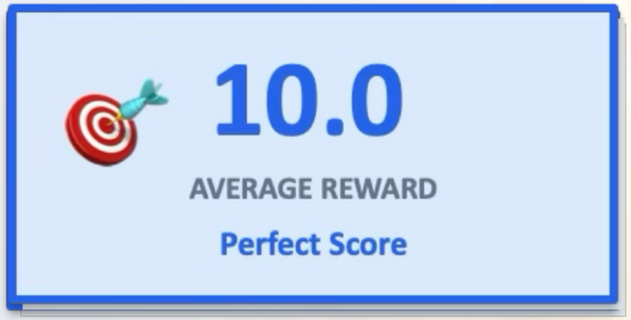
- Model-based, Bellman optimality updates
- Direct value function optimization

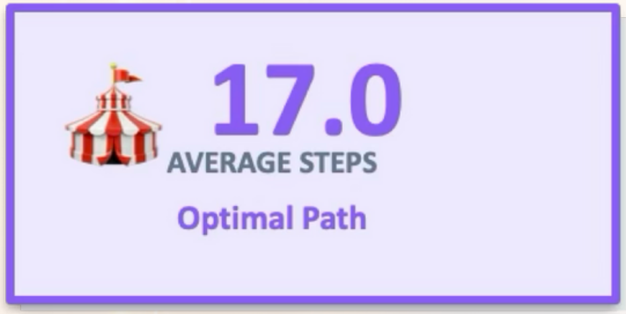
Q-Learning

- Model-free, learns from experience
- Exploration-exploitation tradeoff

Results – Performance Comparison



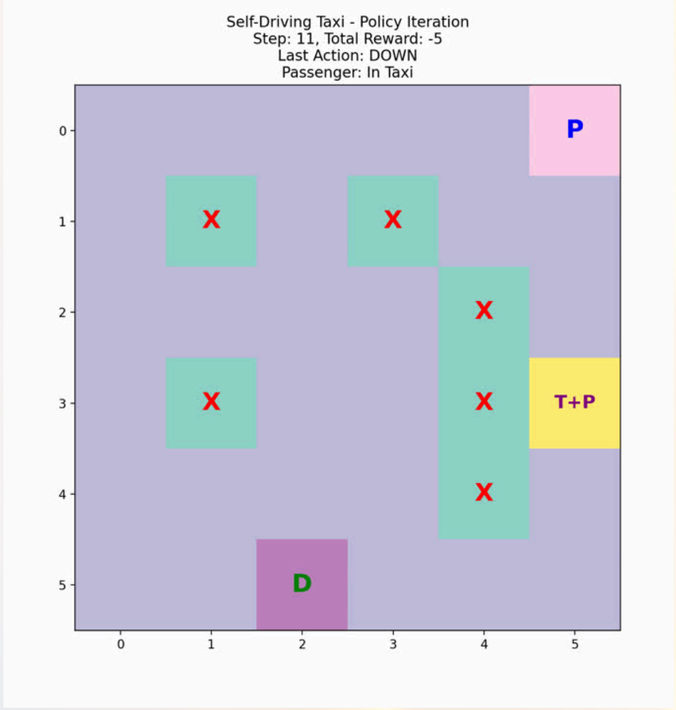


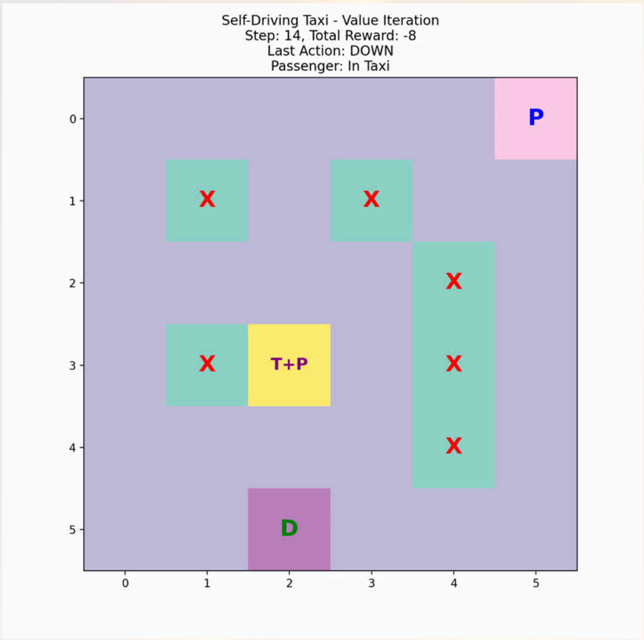


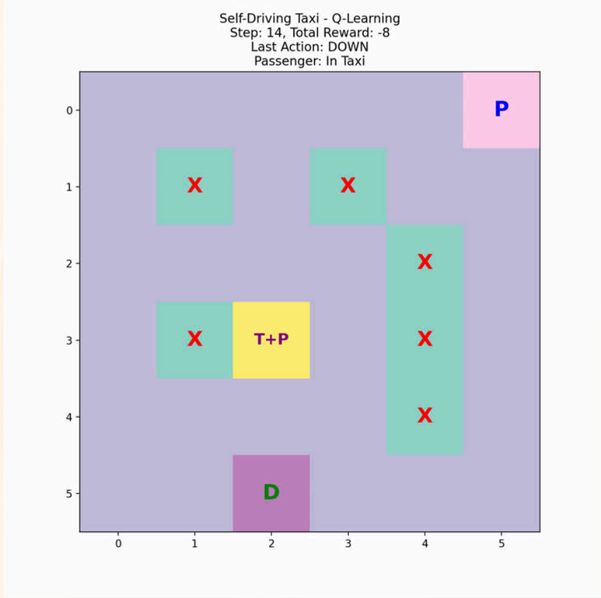
Algorithm	Time (s)	Avg Reward	Avg Steps	Model-Based	Convergence
Policy Iteration	0.69	10	17	✓ Yes	Guaranteed
Value Iteration	**0.18**	10	17	✓ Yes	Guaranteed
Q-Learning	0.44	10	17	✗ No	Probabilistic

Visualization Results

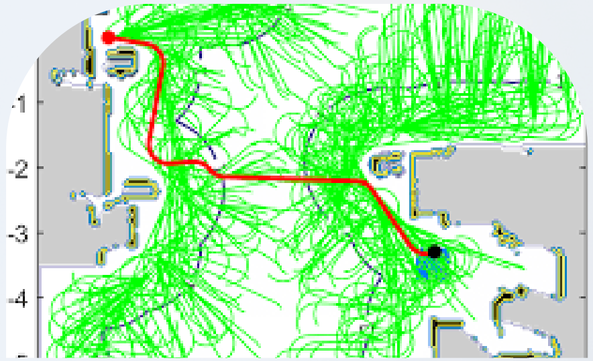
Animated Outputs Created:







Visual Features:



Real-time path visualization

Step counter and reward tracker

Passenger status display

Key Findings

01 Speed: Value Iteration fastest (0.18s)

02 Consistency: Policies are nearly identical

03 Accuracy: All achieve ~10 average reward

04 Trade-off: Q-Learning slower but model-free

05 Efficiency: Average 17 steps to complete task

Successfully compared three RL approaches

Code available on GitHub for educational use

Value Iteration recommended for similar deterministic problems

Conclusion & Future Work

Thank You!