Krish Bhanderi - Report

WiDS: Exploring Reinforcement Learning

Introduction

Attending WiDS 2024 at IIT Bombay's Analytics Club was an incredible opportunity. Under the mentorship of Balaji Karedla and Nirav Bhattad, I worked on my project titled "Exploring Reinforcement Learning." This experience deepened my understanding of AI, particularly how machines learn from their environment. Through this project, I transitioned from an introductory grasp of Reinforcement Learning (RL) to an advanced understanding, acquiring both theoretical knowledge and practical implementation skills. One of my major achievements was developing a Sliding Puzzle Solver using RL techniques, applying all the RL concepts I learned throughout the program. This report encapsulates my learning journey, key insights, and how I intend to utilize this knowledge moving forward.

Fundamentals (Week 0)

Python and Core Libraries

Having a background in C++ from the CS101 course, shifting to Python was an intriguing challenge. I quickly realized Python’s strength in problem-solving through its extensive libraries. The most essential ones I explored included:

NumPy – Essential for numerical calculations
Pandas – Used for organizing and analyzing datasets
Matplotlib – Facilitated clear and effective data visualization

Gaining proficiency in these tools helped me approach problem-solving in new and efficient ways, laying a strong foundation for understanding RL.

Version Control with Git

Mastering Git was another key milestone. Though initially intimidating, I soon grasped its importance in collaborative programming. It allowed me to:

Track modifications in my codebase
Maintain different versions effectively

Collaborate seamlessly with peers

Multi-Armed Bandit Problem (Week 1)

My first deep dive into RL was the Multi-Armed Bandit problem, as outlined in Sutton and Barto's work. This problem helped me comprehend how intelligent agents make optimal choices in uncertain environments.

Key Implementations

Developed a greedy algorithm to explore basic decision-making
Implemented an **ε-greedy** algorithm for a more balanced approach
Compared and evaluated both methods to analyze their effectiveness

Lessons Learned

This hands-on exercise illustrated the importance of balancing exploration and exploitation in decision-making. The **ε-greedy** strategy significantly enhanced performance, an insight that proved valuable throughout my RL studies.

Markov Decision Processes (Week 2)

This week introduced me to the mathematical principles of RL, particularly Markov Decision Processes (MDPs). Initially, the abstract nature of MDPs was challenging, but implementing models for simple environments made them much clearer.

Key Concepts Acquired

Understanding state transitions and reward systems in MDPs
Learning about Markov chains and their probabilistic properties
Using expected values and probability models for decision-making

I found it fascinating that MDPs form the backbone of various RL applications, from gaming AI to robotics.

Dynamic Programming and MDP Applications (Week 3)

This phase bridged theory with practice, allowing me to apply Dynamic Programming (DP) techniques to RL problems.

Key Implementations

Developed policy evaluation and iteration techniques
Solved DP-related problems using optimization strategies
Assessed how slight modifications in policy impact long-term outcomes

This was a pivotal stage in my learning process, as I could now see the connection between RL, MDPs, and practical applications.

Final Project: 15 Puzzle Solver

After weeks of preparation, I was ready to tackle a more complex problem—solving a 15-puzzle using reinforcement learning techniques.

Implementation Highlights

State Representation
Designed an efficient board representation for puzzle states
Created transition functions to define valid moves
Implemented a solvability checker to detect unsolvable configurations
Reinforcement Learning Approach
Applied RL techniques to optimize puzzle-solving strategies
Experimented with value iteration and policy improvement
Used a reward-based system to improve decision-making efficiency
Key Features
Automated detection of unsolvable puzzle states
Optimized solution paths using RL algorithms
Performance enhancements to reduce computation time

Developing this solver was a major accomplishment, demonstrating how RL principles could be effectively applied to real-world problem-solving.

Future Prospects and Applications

Beyond this project, RL techniques have widespread applications in AI and decision-making fields. Some potential areas include:

Gaming AI – Enhancing in-game decision-making and strategy
Robotics – Enabling autonomous navigation and path planning
Optimization Problems – Streamlining logistics and resource allocation

Financial Modeling – Using RL for stock market predictions and trading strategies

This project has fueled my curiosity to further explore these domains and contribute to impactful AI advancements.

Personal Growth

Reflecting on this journey, WiDS 2024 has been a transformative experience. My biggest takeaways include:

Gaining in-depth knowledge of RL principles and their applications
Developing stronger problem-solving abilities, learning to break down complex challenges
Sharpening my programming skills, particularly in writing efficient and structured code
Becoming more adept at technical documentation and explaining complex topics concisely

Progressing from learning Python basics to building a sophisticated AI-based puzzle solver was a demanding yet fulfilling experience. Each obstacle provided an opportunity to refine my thought process and grow as a learner.

Acknowledgments

I am immensely grateful to my mentors, Balaji Karedla and Nirav Bhattad, for their invaluable support and guidance throughout this project. Their expertise and encouragement made my exploration of Reinforcement Learning both enjoyable and enriching. Without their insights and motivation, this learning experience would not have been as impactful.