LEARNING OPTIMAL DECISIONS

Harnessing AI for real-time market adaptation

REINFORCEMENT LEARNING COURSE

Abdulmueez Emiola
Ozioma Okonicha
Pavel Tishkin

PROBLEM AREA

- The goal of this project is to provide some help to drivers and develop an RL agent that can learn to maximize the daily profit, based on the historical data.

PROJECT GOALS

To create a system that strategically shows the driver the best course of action to take, reflecting real-time demand and maximizing long-term revenue.

RL ALGORITHMS

We used algorithms like **Q-Learning, SARSA** and **Deep Q-Learning** to find the best option for the driver to take.

RL ENVIRONMENT

Agent

A taxi driver working in a New York city

States

Include:

- Current hour
- Current minute
- Current area zip code

Actions

Pick an order from the list, or wait another hour

Rewards

- Revenue from fullfilling the order

DATASET

NYC Taxi Trips

PREPROCESSING

Subsampling

Selection of just 1 month data for use

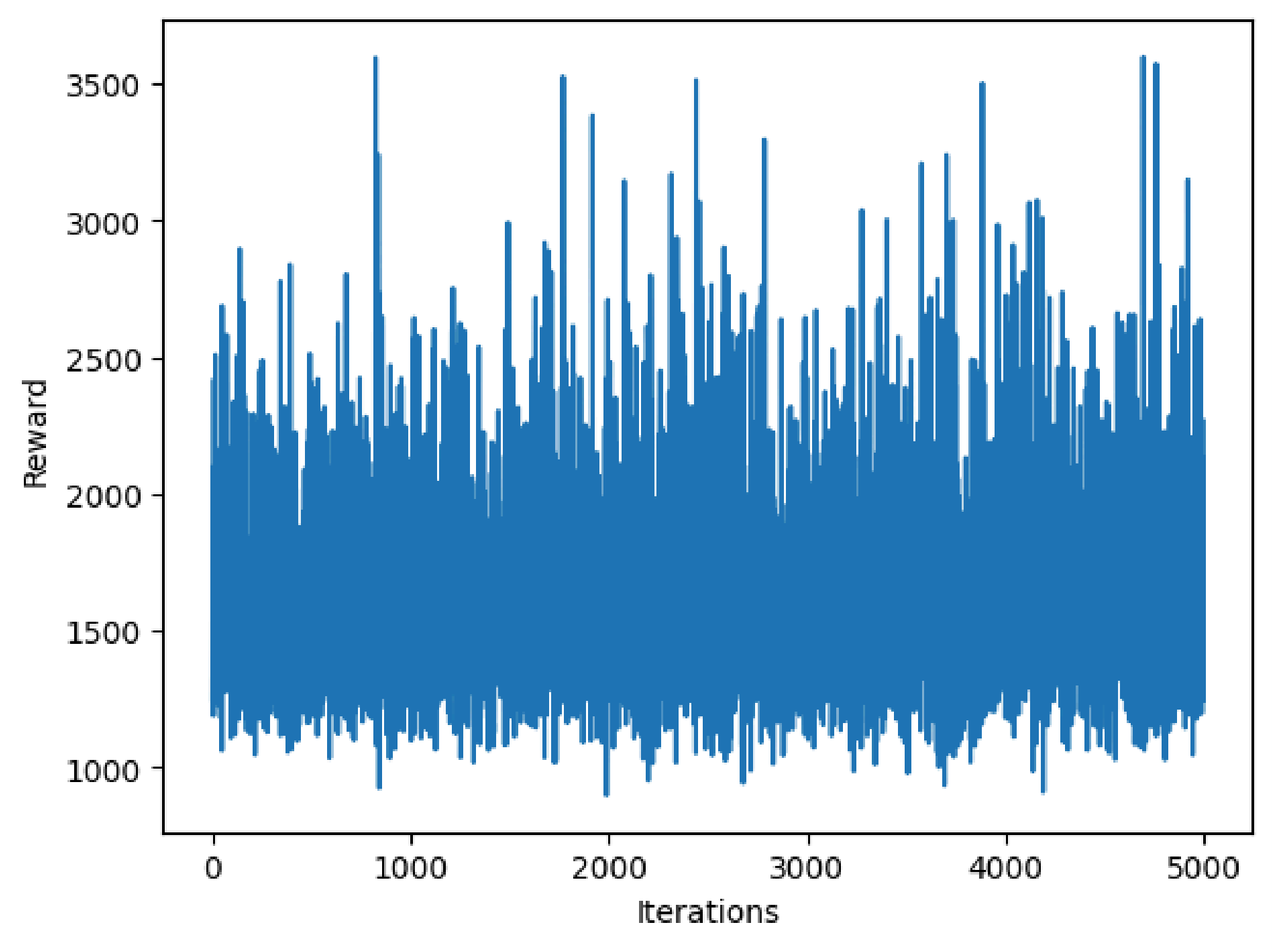
Location Transformation

Conversion of latitude and longitude to zipcode.

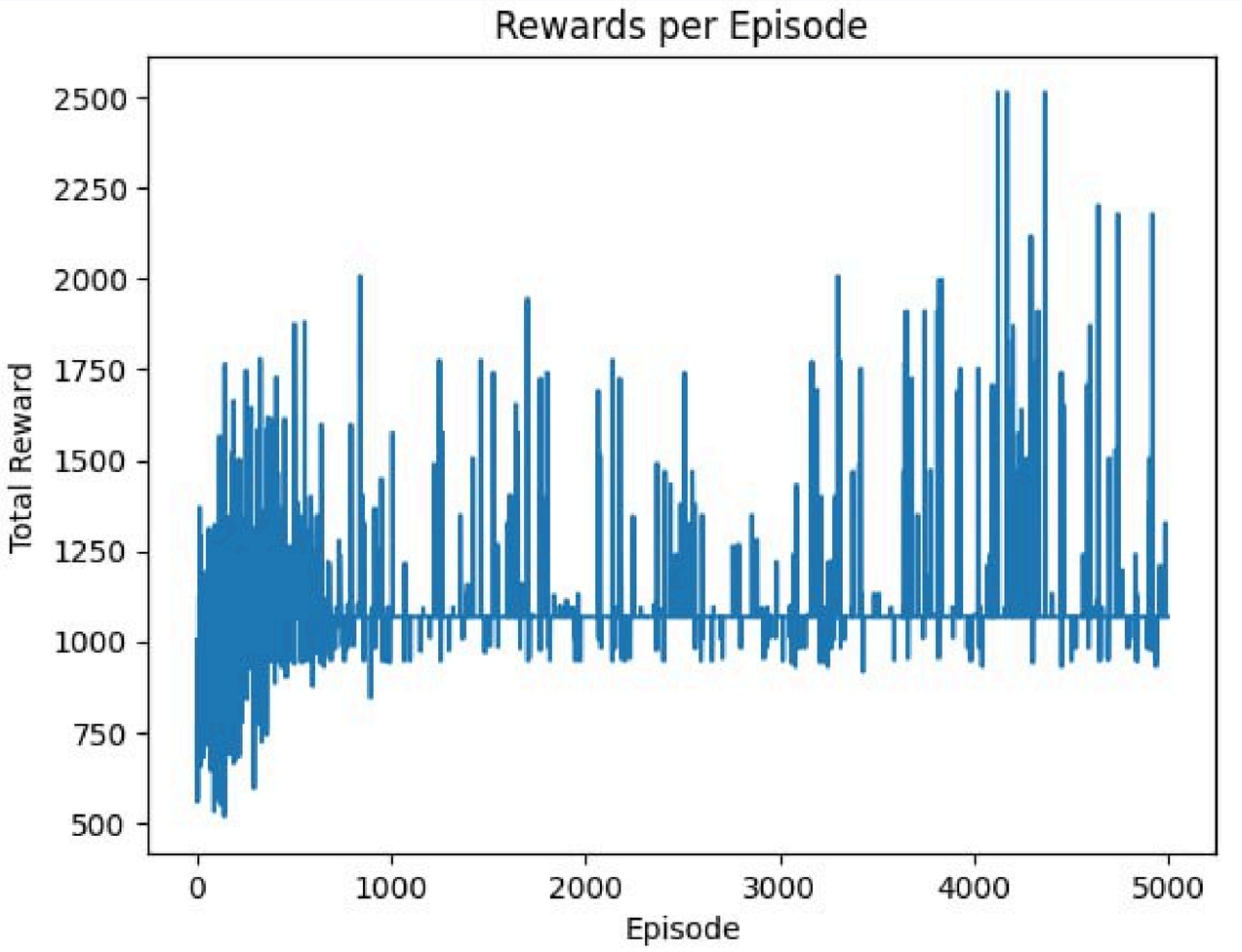
Summary Statistics

Generate summary statistics based on grouping by the state.

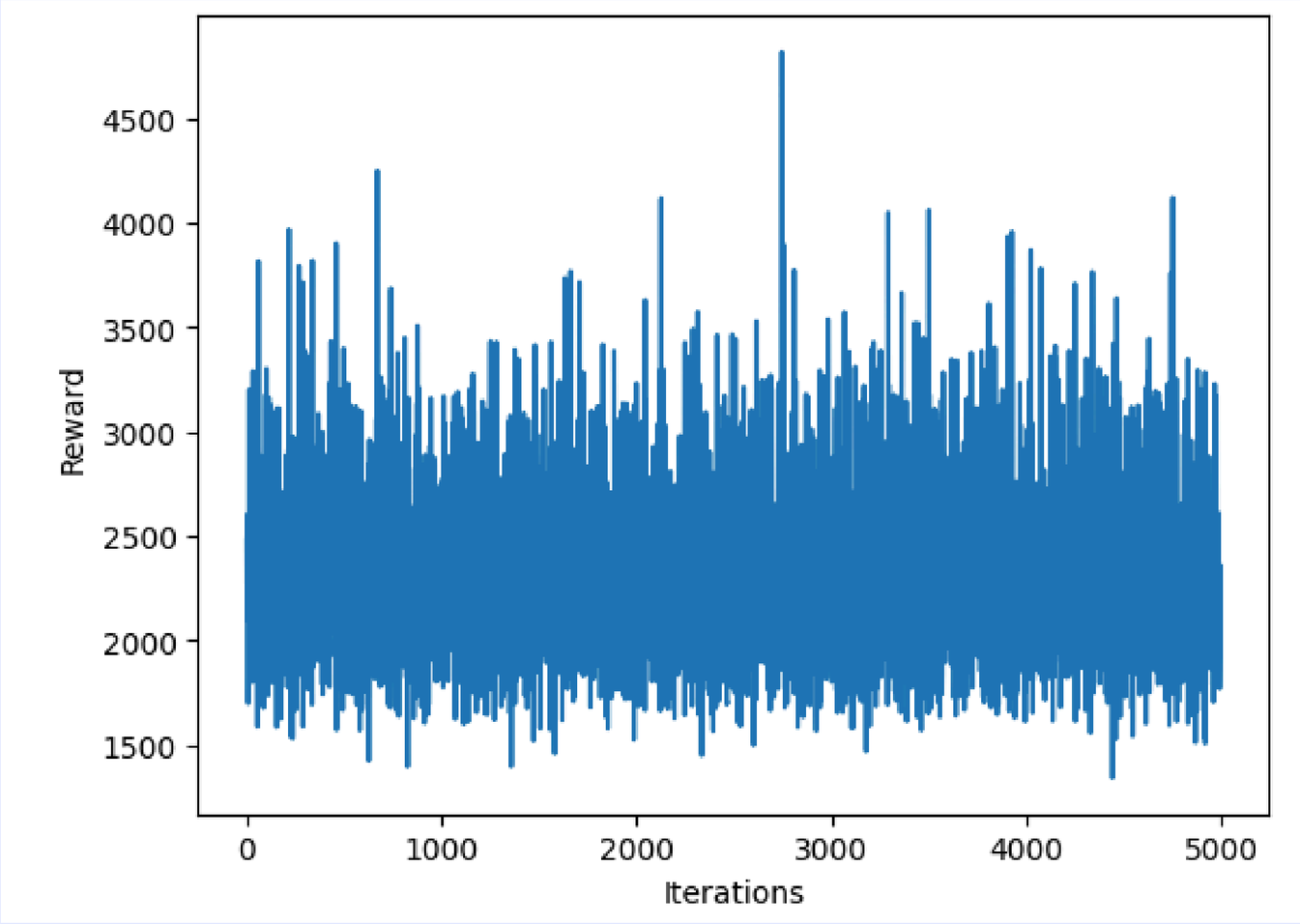
RANDOM REWARDS



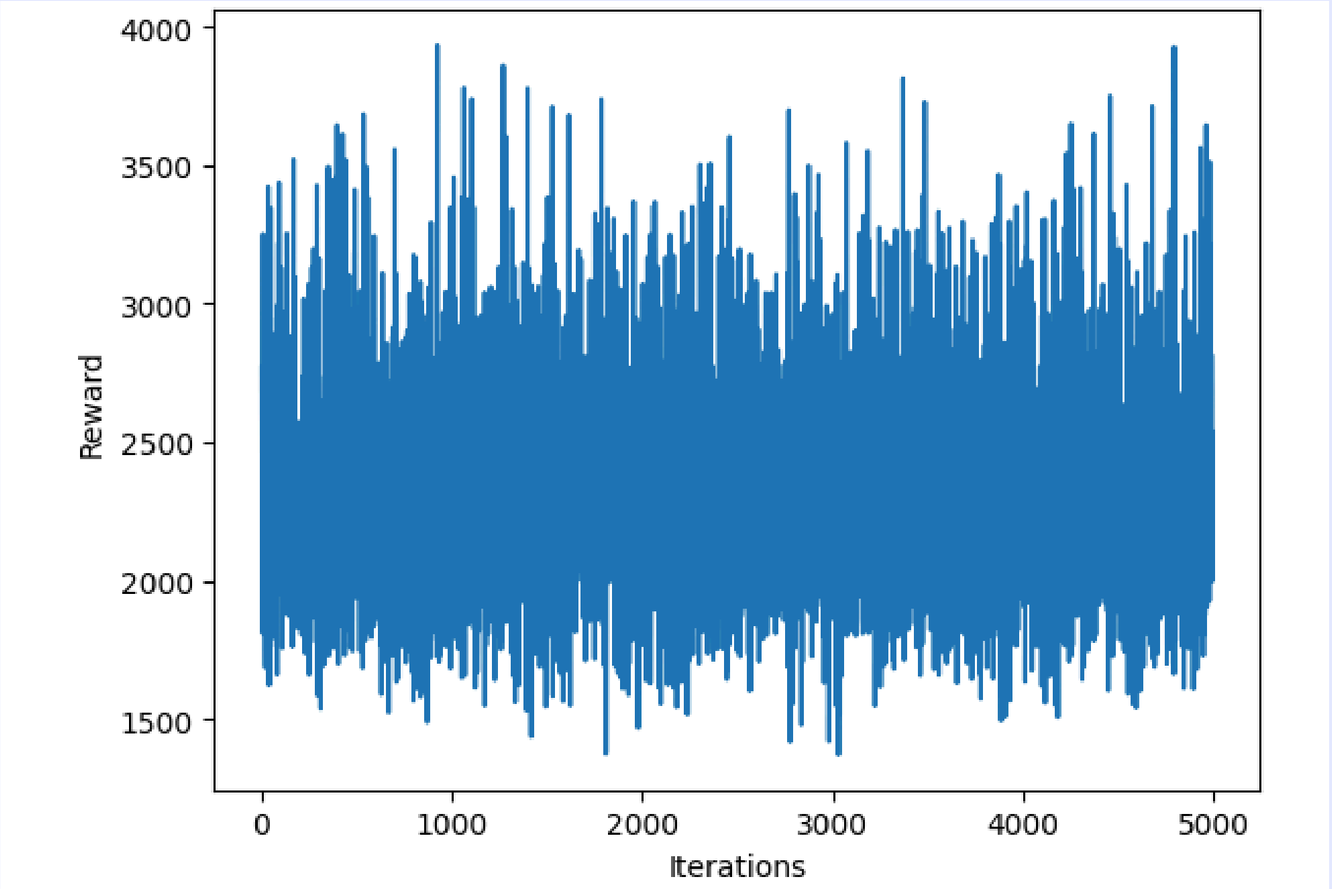
Q-LEARNING


Rewards per Episode
Total Reward
2500
2250
2000
1750
1500
1250
1000
750
500
0
1000
2000
3000
4000
5000
Episode

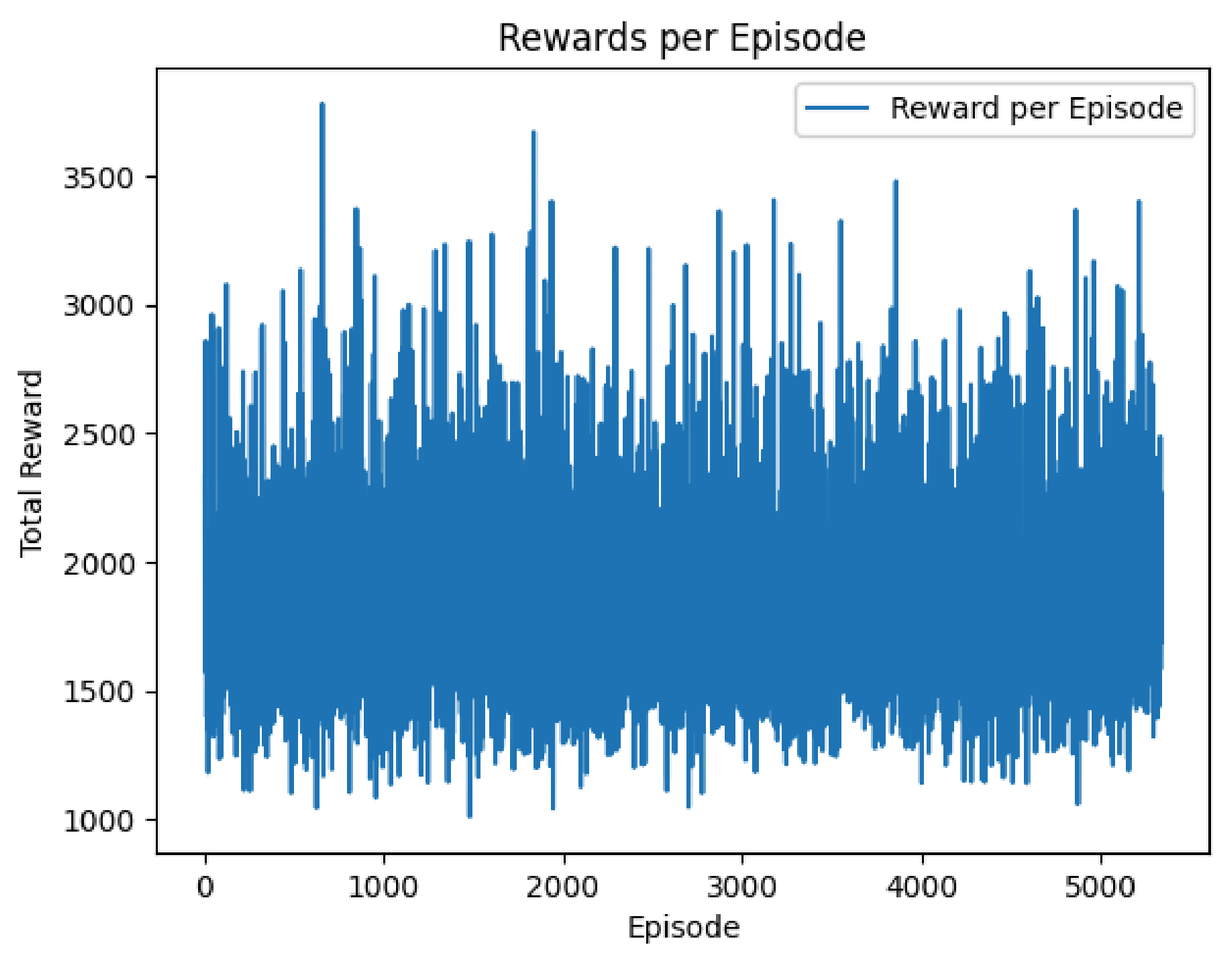
DEEP Q-NET


Reward
4500
4000
3500
3000
2500
2000
1500
0
1000
2000
3000
4000
5000
Iterations

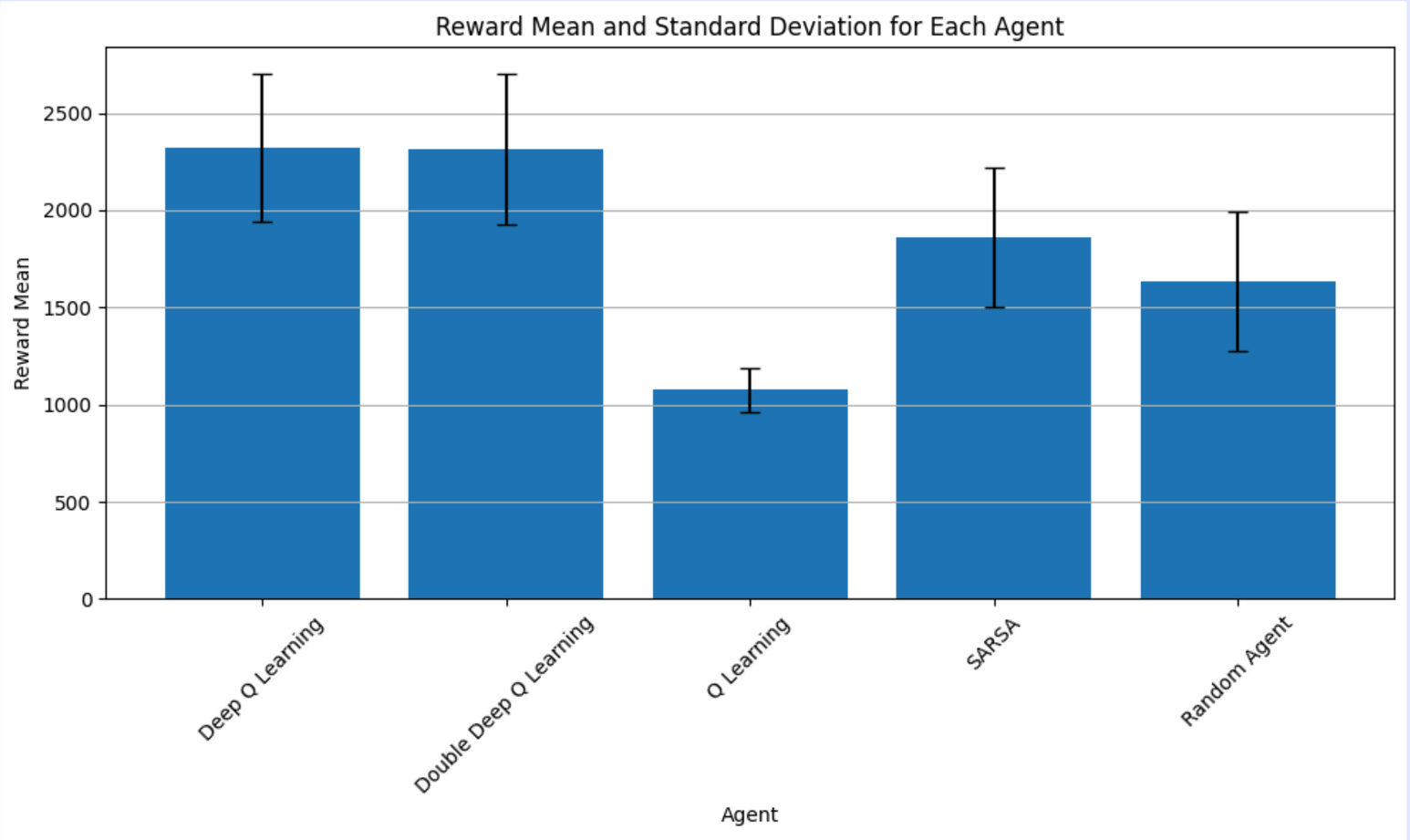
DOUBLE DEEP Q-NET



SARSA


Rewards per Episode
Reward per Episode
Total Reward
3500
3000
2500
2000
1500
1000
0
1000
2000
3000
4000
5000
Episode

AGENT COMPARISON



REWARDS

Agent	Reward Mean	Reward STD
Deep Q Learning	2320.93	381.50
Double Deep Q Learning	2314.11	386.19
Q Learning	1074.54	112.92
SARSA	1861.225	358.469
Random Agent	1632.196	356.37

CHALLENGES

- Environment Design
- State and action size
- Model finetuning

Q&A

Thank you!