GEORGETOWN UNIVERSITY

Medication Reminder System Using Deep Q-Networks and Q-Learning

TECHNICAL WRITEUP

Prepared By

Parsa Keyvani

DSAN 6650 - Spring 2024

Abstract

In this report, we present the development and evaluation of a medication reminder system designed to optimize reminder timing using reinforcement learning techniques. Inspired by a sophisticated model from existing literature which was not fully replicable due to data and complexity constraints, this project implements a more accessible solution. We employed a Deep Q-Network (DQN) model for the primary system and used a basic Q-Learning algorithm as a baseline for performance comparison. The results showed relatively good performance for the DQN model as well as the baseline Q-Learning model. Our results suggest that Reinforcement learning may be a promising approach to improve patients' medication adherence.

1. Introduction

There have been many advancements in healthcare that have significantly impacted patient health outcomes. One of the most critical factors affecting both patient health and the overall effectiveness of healthcare systems is medication adherence. Ensuring consistent medication adherence, especially for individuals with chronic conditions, remains a critical challenge to this day.

This project was inspired by a work documented in a recent study, "The impact of using reinforcement learning to personalize communication on medication adherence" published in *NPJ Digital Medicine*[1]. This study explored the use of reinforcement learning (RL) to tailor communication strategies, specifically text messaging, to improve medication adherence among diabetes patients. The authors developed an RL model that adapted its strategies based on the responsiveness of individuals to different types of message content, which significantly improved medication adherence. Despite the compelling findings, the full replication of the study was constrained by the unavailability of its complete dataset to the public and the complexity of its model implementation.

Motivated by the potential shown in the referenced paper but faced with these limitations, our project aimed to implement a more accessible yet robust approach using RL techniques. Specifically, we developed a medication reminder system using a Deep Q-Network (DQN) to optimize the timing and frequency of medication reminders. Additionally, we employed a basic Q-Learning model as a baseline to juxtapose its performance against the more sophisticated

[1] Lauffenburger, J.C., Yom-Tov, E., Keller, P.A. *et al.* The impact of using reinforcement learning to personalize communication on medication adherence: findings from the REINFORCE trial. *npj Digit. Med.* **7**, 39 (2024). https://doi.org/10.1038/s41746-024-01028-5

DQN model. This approach not only simplified the system's complexity but also made it feasible to implement with the available resources and within the project's constraints.

2. Goals

The primary goal of this project was to develop a reinforcement learning-based system capable of effectively reminding patients to take their medication at the right times. Our secondary goal was to compare the effectiveness of a DQN model against a traditional Q-Learning model in this context.

3. Methods

3.1 System Design

The medication reminder system was conceptualized as a reinforcement learning (RL) agent to determine optimal reminder times for patients based on their prescribed medication schedules. To simulate real-world scenarios, an environment was designed wherein each state was defined by a specific time of day coupled with boolean indicators—whether the medication had been taken, and whether a reminder had already been sent.

To achieve this, two distinct reinforcement learning models, a Deep Q-Network (DQN) and a Q-Learning model, were developed to assess their efficacy in a controlled learning environment.

Deep Q-Network (DQN):

The DQN model was designed using a neural network architecture to approximate the Q-value function. This network comprised three layers: an input layer sized to match the number of state features, two hidden layers each with 24 neurons employing ReLU activation functions to introduce non-linearity, and an output layer tailored to the number of possible actions, using linear activation to output estimated Q-values for each action. The learning process also featured a replay buffer to mitigate the correlation between sequential updates and prevent overfitting. This buffer stored experiences as tuples and allowed for batch updates where experiences were randomly sampled and used to update the network, with losses computed based on the Bellman equation.

Q-Learning (Baseline Model):

In contrast, the Q-Learning model employed a simpler, tabular approach where a Q-table was used. The Q-table dimensions correspond to the states and actions defined in the environment and mentioned earlier, with each entry representing the Q-value for a specific state-action pair. The Q-values within this table were updated iteratively following a defined learning update rule, which adjusted the values based on the learning rate, received rewards, a discount factor, and the maximum predicted Q-values for subsequent states. This model is more straightforward, though less flexible, approach to learning compared to the neural network-based DQN model.

3.3 Data Handling

In the simulation of our medication reminder system, synthetic data was generated to represent a typical day broken down into distinct time slots, each corresponding to a specific state within the environment. As the agent navigates through these states, its decisions at each time slot result in transitions to subsequent states. These transitions are directly influenced by the agent's actions, specifically whether a medication reminder was sent at the appropriate times and whether the medication was subsequently taken by the patient.

The reward structure within this simulated environment is designed to reinforce the accuracy and timeliness of the agent's actions. Positive rewards are given for sending reminders that successfully prompt medication adherence, which ensurs patient compliance at the designated times. Conversely, the agent incurs negative rewards for either failing to send reminders or sending unnecessary reminders; therefore, penalizing poor decision-making that could potentially disrupt the patient's medication schedule. Additionally, the magnitude of both positive and negative rewards is calibrated to reflect the severity of the agent’s correct or incorrect actions, which improves the learning impact of each decision.

Moreover, both the Deep Q-Network (DQN) and Q-Learning models were evaluated over multiple episodes to assess their learning efficacy and generalization capabilities across different scenarios within the simulated environment. Key performance metrics tracked during these evaluations included cumulative reward, average reward, sample efficiency, and convergence rate. These metrics provide us with a quantitative basis for analyzing the models' performance, which gives us insights into their operational effectiveness and adaptive learning behaviors in response to the dynamic environmental conditions.

4. Results

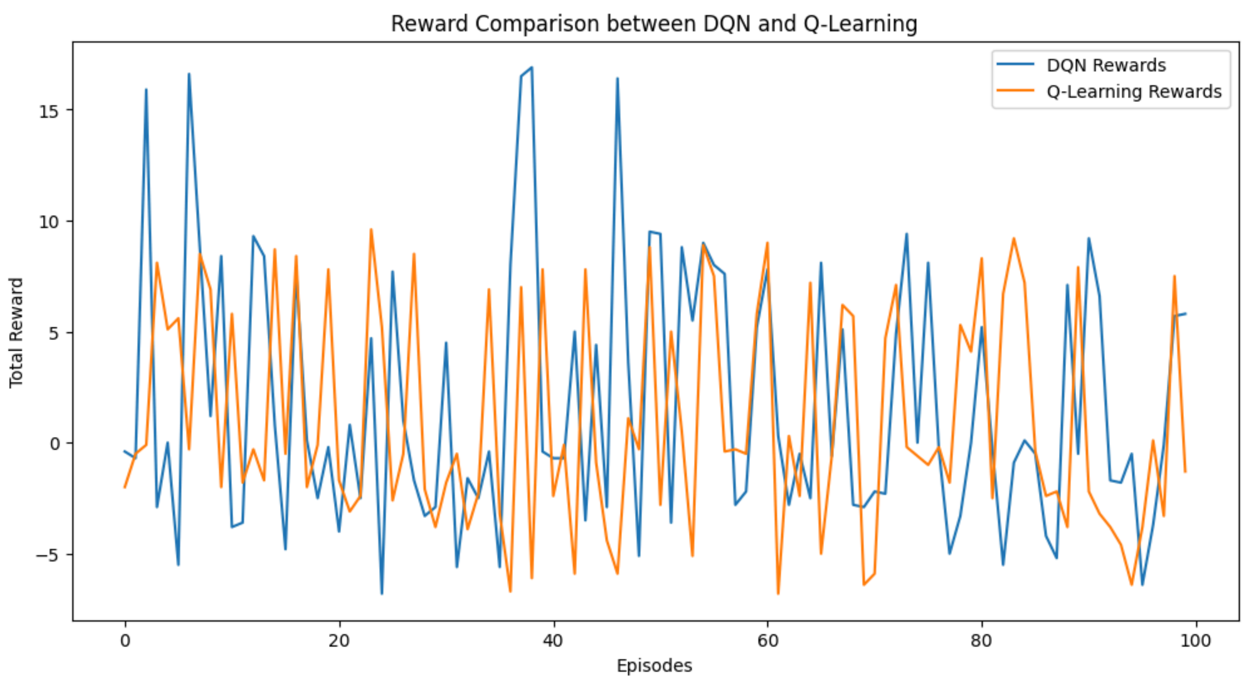
The evaluation of the Deep Q-Network (DQN) and Q-Learning models on the medication reminder system revealed distinct performance characteristics across 100 episodes. The DQN model showed a notably superior performance with a cumulative reward of 94.40 and an average reward per episode of 0.944, reflecting a more consistent and effective learning strategy. Its sample efficiency was recorded at 12.94 steps per episode, and it achieved a positive convergence rate of 0.67, indicating a steady improvement in learning over time.

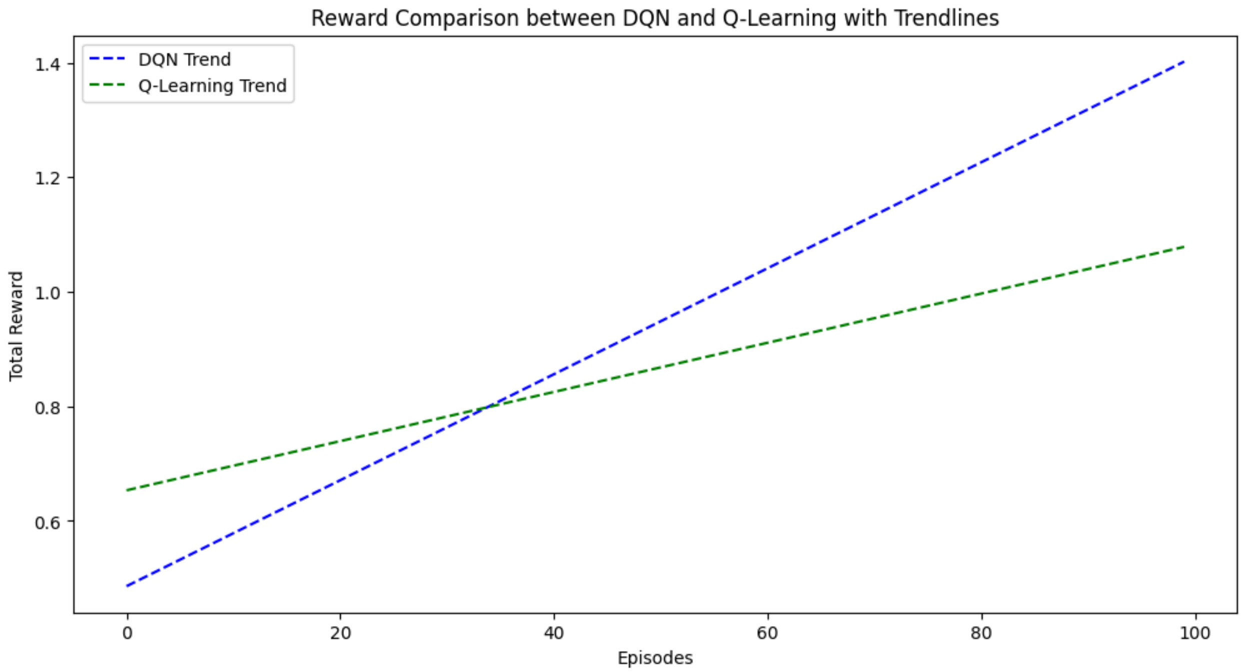
In contrast, the Q-Learning model demonstrated a lower cumulative reward of 86.60 and an average reward per episode of 0.866, alongside a sample efficiency of 13.28 steps per episode.

Remarkably, the Q-Learning model presented a convergence rate of 3.64, suggesting a substantial increase in performance towards the end of the training episodes. This positive convergence rate indicates a potential late-stage recovery in learning efficiency and model adaptability compared to the DQN model.

These results show that the DQN model within the simulated environment, outperforms the simpler Q-Learning approach in terms of reward robustness and learning progression. The data also suggest that while both models navigate the task environment, the DQN's neural network architecture provides a more powerful mechanism for optimizing decision-making processes in complex scenarios, such as medication adherence.

Model Type	Cumulative Reward	Average Reward	Sample Efficiency	Convergence Rate
DQN Model	94.40	0.944	12.94	0.67
Q-Learning Model	86.60	0.866	13.28	3.64





5. Conclusions

This project demonstrated the application of reinforcement learning (RL) to enhance medication adherence through a medication reminder system. By employing both a Deep Q-Network (DQN) and a simpler Q-Learning model, we explored the potential of RL to optimize the timing and delivery of medication reminders. The evaluation results indicated a performance advantage of the DQN model over the Q-Learning model, with the DQN model achieving better cumulative and average rewards, a lower step count per episode, and a positive convergence rate that suggests a steady improvement in learning over time.

Conversely, the Q-Learning model, while initially underperforming in cumulative and average rewards compared to the DQN, exhibited a remarkable late-stage performance improvement as indicated by its higher convergence rate of 3.64. This suggests that the Q-Learning model, despite its simplicity, may have a capacity for significant adaptation and efficiency improvements towards the later stages of training, potentially catching up to or surpassing the DQN's performance.

In this study, we demonstrated that both sophisticated and simpler RL models have their unique strengths and potential in real-world healthcare settings. Our next steps would be to refine the reward structure, explore different state representations, and integrate more complex RL algorithms that could further improve the performance.

6. Code Availability

The source code for both models has been made available in my GitHub repository, accessible at https://github.com/keyvanip/med_reminder . This repository includes all codes used for implementing the DQN and Q-Learning models, as well as the environment setup.

References

Lauffenburger, J.C., Yom-Tov, E., Keller, P.A. *et al.* The impact of using reinforcement learning to personalize communication on medication adherence: findings from the REINFORCE trial. *npj Digit. Med.* **7**, 39 (2024). https://doi.org/10.1038/s41746-024-01028-5