Optimizing Sepsis Treatment Decisions with Reinforcement Learning

Punit Kumar; Vaibhav Saran; Divyesh Patel

Summary

This project aims to address optimal decision-making in sepsis treatment using Reinforcement Learning (RL). Specifically, we will explore how to improve patient outcomes by recommending timely and safe medical interventions based on historical ICU patient data. Our research will focus on creating intelligent RL agents that learn from clinical data to suggest optimal treatments, potentially reducing mortality and improving patient outcomes.

Problem Statement: Sepsis is a life-threatening condition with significant mortality rates in intensive care units (ICUs). Current treatment protocols often rely on generalized guidelines, which may not be optimal for individual patient conditions. We aim to address this limitation by developing personalized treatment recommendations using reinforcement learning, thereby enhancing clinical decision-making processes.

Objectives

- Develop personalized RL-driven treatment strategies for sepsis patients.
- Integrate innovative RL methodologies to enhance decision-making accuracy and safety.
- Demonstrate improved patient outcomes through intelligent agent recommendations.
- Compare performance with existing standard treatment guidelines and baseline RL models.

Methodology

Baseline RL techniques, including Deep Q-Networks (DQN) and Advantage Actor-Critic (A2C), will be initially employed (Hunt). To enhance these models, the following innovative methodologies will be introduced:

1. **Safety-aware Reinforcement Learning:** Implementing explicit safety constraints within the RL framework to ensure recommendations remain within clinically acceptable and safe parameters.
2. **Attention-based Policy Networks:** Incorporating attention mechanisms in neural networks to dynamically prioritize critical clinical variables, allowing agents to focus on relevant patient data. (Sartoretti)
3. **Hybrid Reward Shaping:** Integrating expert clinical feedback directly into the reward structure, guiding the RL agent towards medically sound and practically valuable decisions.

Evaluation

Qualitative evaluation:

- Visualization of reward convergence plots.
- Policy evolution graphs demonstrating improved decision-making.
- Attention heatmaps indicating the agent's focus on critical patient features.

Quantitative evaluation:

- Metrics including cumulative rewards and policy stability.
- Clinical performance indicators: mortality rates, length of ICU stays, and accuracy of treatment recommendations compared with baseline guidelines.
- Statistical significance testing (e.g., t-tests, ANOVA) to validate performance improvements over baseline approaches.

Expected Results:

- Clear visual evidence showing enhanced agent performance through reward convergence and stabilized policy behavior.
- Significant improvement in clinical outcomes and decision-making quality compared to baseline RL approaches without advanced modifications.

Environment

Clinical Data Foundation

- **MIMIC-III Database**: Curated ICU dataset with 40,000+ patient records, including vitals, lab results, medications, and outcomes. (Johnson et al.)
- **Preprocessing**: Time-series alignment of heterogeneous data (e.g., EHRs, medications) with ethical handling of missing values and biases.

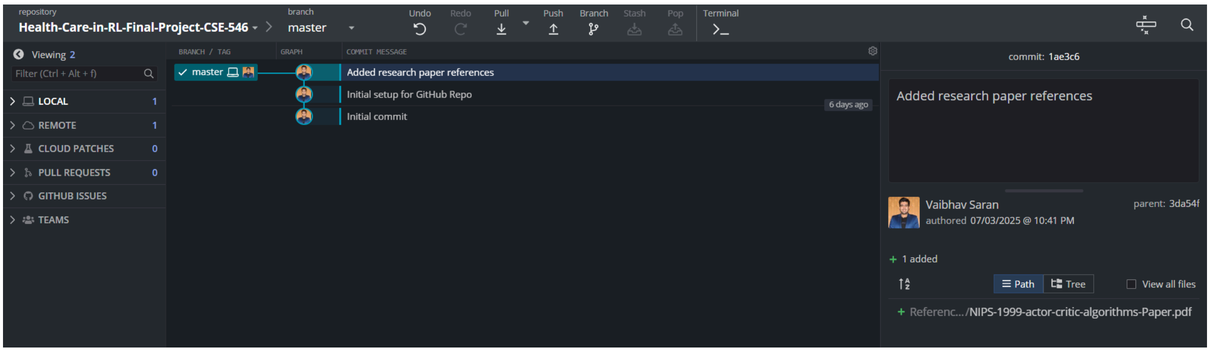
Simulation Platform

- **Modified Sepsis Simulator**: Built on the work of Komorowski (2018) and Raghu (2017), this environment models sepsis physiology using MIMIC-III data. (Komorowski) (Raghu)
- **Key Features**:
    - Realistic patient state transitions (e.g., response to fluids/antibiotics).
    - Risk penalties for unsafe actions (e.g., hypotension from excessive vasodilation).
    - Customizable scenarios (e.g., septic shock, comorbidities) to test generalization.
- **Ethical Safeguards**: Actions are validated against clinical guidelines and reviewed by ICU physicians to prevent harmful recommendations.

References

1. Achiam, Joshua. “[1705.10528] Constrained Policy Optimization.” *arXiv*, 30 May 2017, https://arxiv.org/abs/1705.10528. Accessed 7 March 2025.
2. Hunt, Jonathan. “[1509.02971] Continuous control with deep reinforcement learning.” *arXiv*, 9 September 2015, https://arxiv.org/abs/1509.02971. Accessed 7 March 2025.
3. Johnson, Alistair, et al. “MIMIC-III Clinical Database v1.4.” *PhysioNet*, 4 September 2016, https://physionet.org/content/mimiciii/1.4/. Accessed 7 March 2025.
4. Jones, Llion. “[1706.03762] Attention Is All You Need.” *arXiv*, 12 June 2017, https://arxiv.org/abs/1706.03762. Accessed 7 March 2025.
5. Kiani, Amir. “GYMIC.” OpenAI, https://github.com/akiani/gym-sepsis.
6. Komorowski, Matthieu. “The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care.” *PubMed*, https://doi.org/10.1038/s41591-018-0213-5. Accessed 7 March 2025.
7. Konda, Vijay R., and John N. Tsitsiklis. “Actor-Critic Algorithms.” https://proceedings.neurips.cc/paper/1999/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf. Accessed 7 March 2025.
8. Mnih, Volodymyr. “Playing Atari with Deep Reinforcement Learning.” *arxiv*, https://arxiv.org/abs/1312.5602v1.
9. Raghu, Aniruddh. “sepsisrl.” https://github.com/aniruddhraghu/sepsisrl/tree/master.
10. Sartoretti, Guillaume. “[2301.11575] ARiADNE: A Reinforcement learning approach using Attention-based Deep Networks for Exploration.” *arXiv*, 27 January 2023, https://arxiv.org/abs/2301.11575. Accessed 7 March 2025.
11. Schulman, John. “[1707.06347] Proximal Policy Optimization Algorithms.” *arXiv*, 20 July 2017, https://arxiv.org/abs/1707.06347. Accessed 7 March 2025.

Github Commits



Github Projects