

Modelos de Decisão para Testagem de Casos Suspeitos de Arboviroses por Aprendizagem por Reforço

Zuilho R. C. Segundo¹, Flávio C. Coelho²

FGV EMap, Rio de Janeiro, RJ

Este trabalho teve como objetivo aplicar a teoria da aprendizagem por reforço para resolver o seguinte problema: maximizar a acurácia na confirmação dos casos, minimizando ao mesmo tempo o custo, ou seja, o número de testes laboratoriais realizados. O modelo leva em conta a distribuição espaço-temporal dos casos notificados, as características operacionais dos testes (como sensibilidade e especificidade) e cenários simulados de epidemias de dengue e chikungunya com diferentes números básicos de reprodução (R_0).

O ambiente de simulação foi desenvolvido em Python/Gymnasium [2], utilizando um modelo SIR (Suscetível-Infetado-Recuperado) [1] para representar a propagação de doenças. A distribuição espacial dos casos foi modelada como uma dispersão gaussiana em torno de pontos de foco pré-definidos, simulando surtos localizados. O agente de aprendizagem por reforço interage com esse ambiente por meio de ações que incluem: (1) testagem específica para cada arbovírus (dengue ou chikungunya); (2) confirmação ou descarte epidemiológico baseado em critérios operacionais; e (3) inação estratégica para evitar custos desnecessários. A cada ação, o agente recebe uma recompensa que busca equilibrar a acurácia diagnóstica e os custos operacionais, conforme a função:

$$R(s, a) = \alpha \cdot \text{Acurácia}(s, a) - \beta \cdot \text{Custo}(a). \quad (1)$$

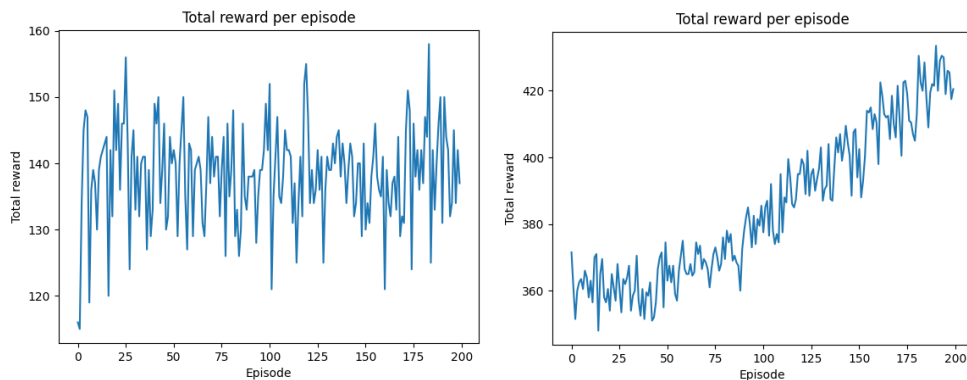


Figura 1: Desempenho comparativo dos agentes ao longo do treinamento. Fonte: Autoria própria.

Na primeira etapa, comparou-se o desempenho de dois agentes: um agente aleatório (baseline) e outro utilizando Q-Learning [4]. O agente baseline seleciona ações de forma aleatória, enquanto o

¹zuilho.segundo@fgv.edu.br

²flavio.codeco.coelho@fgv.br

Q-Learning atualiza iterativamente uma tabela de valores esperados de recompensa. Os resultados (Figura 1) demonstraram que o Q-Learning superou significativamente o agente aleatório, com um aumento progressivo da recompensa acumulada ao longo das épocas de treinamento, indicando a aprendizagem de uma política eficiente de testagem.

Estes resultados preliminares validam o potencial da aprendizagem por reforço para otimizar estratégias de testagem em arboviroses, combinando eficiência operacional e precisão diagnóstica. Como próximos passos, planeja-se: (1) implementar algoritmos de Deep Reinforcement Learning [3] (como Deep Q-Learning), capazes de generalizar em espaços de parâmetros complexos; (2) incorporar dados reais de vigilância epidemiológica para calibrar as simulações; e (3) analisar a sensibilidade do modelo a parâmetros como a especificidade dos testes e atrasos na liberação de resultados. A expectativa é que esta abordagem forneça insights valiosos para a otimização de políticas de testagem em situações reais, onde os recursos são limitados e a agilidade na confirmação de casos é crítica.

Agradecimentos

Os autores agradecem ao Wellcome Trust pelo financiamento ao projeto.
Grant nº 226088/Z/22/Z.

Referências

- [1] A.-L. Barabási e M. Pósfai. **Network Science**. Cambridge: Cambridge University Press, 2016. ISBN: 9781107076266.
- [2] Farama Foundation. **Site oficial do Gymnasium**. Online. Acessado em 15/03/2025, <https://gymnasium.farama.org/>.
- [3] M. Lapan. **Deep Reinforcement Learning Hands-on: Apply Modern RL Methods to Practical Problems of Chatbots, Robotics, Discrete Optimization, Web Automation, and More**. 2a. ed. Expert Insight. Packt Publishing, 2020. ISBN: 9781838826994.
- [4] C. J. C. H. Watkins e P. Dayan. “Technical Note: Q-Learning”. Em: **Machine Learning** 8 (1992), pp. 279–292. DOI: 10.1007/BF00992698.