
Instituto Tecnológico de Costa Rica**Escuela de Ingeniería Electrónica****Trabajo Final de Graduación****Proyecto:** Método basado en aprendizaje reforzado para el control automático de una planta no lineal.**Estudiante:** Oscar Andrés Rojas Fonseca

I Semestre 2024

Firma del asesor

Bitácora de trabajo

Fecha	Actividad	Anotaciones	Horas dedicadas
08/04/2024	1. AAA.	a) DDDDDDD	6 horas
09/04/2024	2. Pruebas de implementación <i>CUDA</i> en Windows.	a) SASASASAS.	4 horas
09/04/2024	2. Pruebas de entrenamiento del modelo <i>Pendulum DQN</i> .	a) SASASASAS.	4 horas
10/04/2024	3. Reunión de seguimiento con el asesor del proyecto.	a) Revisión de avance en el código y errores de forma. b) Se acordó realizar entrenamientos con diferentes formatos de indicación del <i>target_angle</i> .	2 horas

11/04/2024	4. Corrección de potenciales errores en el código <i>PendulumDQN</i> señalados por asesor.	a) Replanteo de función de recompensa <i>calculate_reward()</i> para evitar salto. b) Adición de lógica para guardado de <i>checkpoints</i> al entrenamiento y corrección del guardado del modelo.	6 horas
12/04/2024	5. Continuación de corrección de errores potenciales en el código.	a) Replanteo de función <i>select_action()</i> ; cambio de acción aleatoria en exploración a adición de ruido a la opción elegida.	4 horas
12/04/2024	6. Estudio de conceptos <i>MDP</i> [1] y <i>DQN</i> [2] .	a) Revisión de aplicación mediante <i>MDP</i> dada la mención en una fuente en línea donde se utiliza [3]. b) Estudio de teoría <i>DQN</i> para mejor comprensión de la lógica de la función <i>optimize_model()</i> del código original [4] y su adaptación a <i>Pendulum</i> .	4 horas
12/04/2024	5. Pruebas de entrenamiento de modelos <i>CartPole</i> y <i>Pendulum</i> .	a) Se crearon los cuadernos <i>ctrlCartPoleDQN.ipynb</i> y <i>ctrlPendulumDQN.ipynb</i> para pruebas de carga de modelos. b) Entrenamiento del modelo <i>Pendulum_1000eps.pth</i> . c) Se descubrió un error grave en <i>select_action()</i> , corrección en proceso.	4 horas
Total de horas de trabajo:			21 horas

Contenidos de actividades

AAA [4].

Referencias

- [1] J. P. A. Moya, “EL5857 Lección 25: Aprendizaje Reforzado (1/4): MDP,” 2021, [Vídeo de YouTube]. [Online]. Available: https://www.youtube.com/watch?v=FBaoss_Pb5Q
- [2] —, “EL5857 Lección 27: Aprendizaje Reforzado (3/4): DQN y Q-Learning,” 2021, [Vídeo de YouTube]. [Online]. Available: <https://www.youtube.com/watch?v=oXnNRSCe5T4>
- [3] S. Israilov, L. Fu, J. Sánchez-Rodríguez, F. Fusco, G. Allibert, C. Raufaste, and M. Argentina, “Reinforcement learning approach to control an inverted pendulum: A general framework for educational purposes,” *PLoS ONE*, 2023.
- [4] A. Paszke and M. Towers, “Reinforcement learning (dqn) tutorial,” *PyTorch*.