

Deep Learning: Navegación Autónoma con DQN

Diego Valdez 21328 • Joaquín Puente 22296 • Nelson García 22434

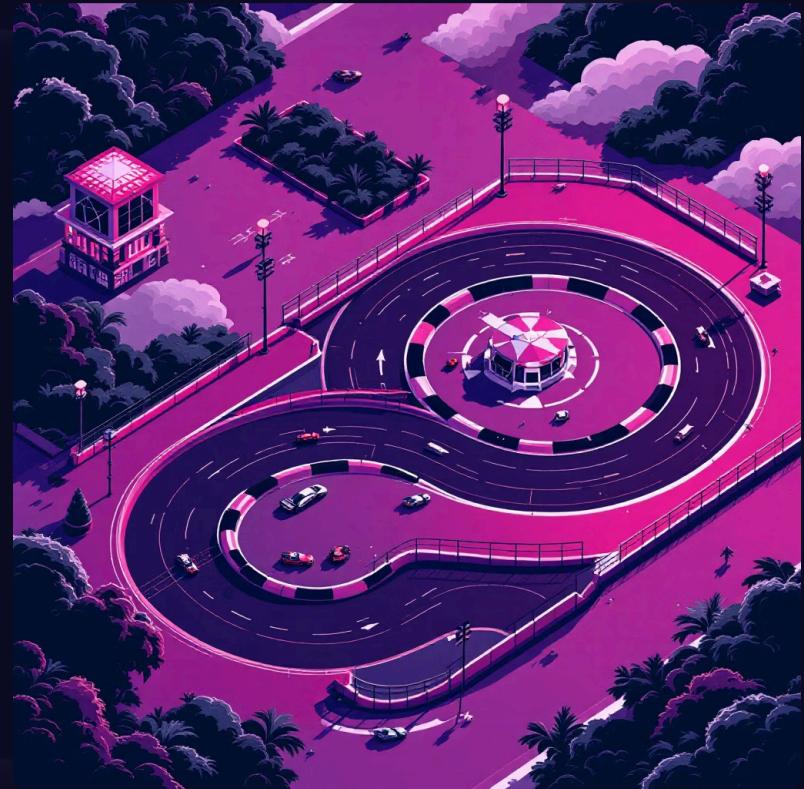


El Desafío

Un vehículo autónomo debe navegar por pistas cargadas desde archivos CSV utilizando aprendizaje por refuerzo con DQN. El entorno simula carreras donde el carro avanza horizontalmente, controlando dirección lateral y aceleración.

Elementos de la Pista

- Pavimento estándar
- Muros (colisión terminal)
- Aceite (reduce velocidad 10%)
- Terracería (reduce velocidad 5%)
- Boost (aumenta velocidad 5%)



Sistema de Recompensas



Progreso

Incentiva avance hacia la derecha y alcanzar la meta



Precisión

Mantener el vehículo cerca del centro del carril



Velocidad

Minimizar tiempo de carrera con penalización constante



Seguridad

Penalización severa por colisiones con muros

Análisis del Problema

El agente enfrenta el desafío de aprender una política que balancee velocidad, seguridad y precisión en un entorno parcialmente observable. Debe anticipar obstáculos con información local, similar a conducir con niebla.

1

Presión Temporal

Penalización por tiempo impulsa velocidad, pero reduce margen de maniobra

2

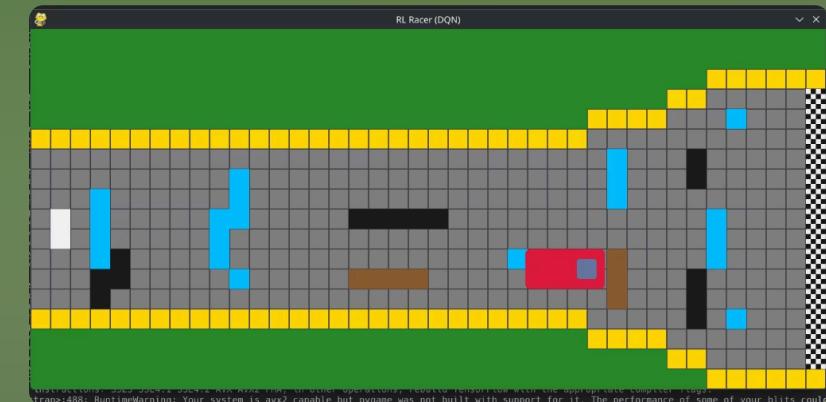
Balance Espacial

Desviación del centro compite con necesidad de esquivar obstáculos

3

Dinámicas Complejas

Superficies especiales alteran comportamiento del vehículo



Propuesta de Mejora

01

Espacio de Observación Mejorado

Añadir canales para velocidad normalizada y offset lateral respecto al centro del carril

02

Ajuste de Recompensas

Incrementar coeficiente de progreso, hacer penalización por desviación más tolerante

03

Límite de Pasos

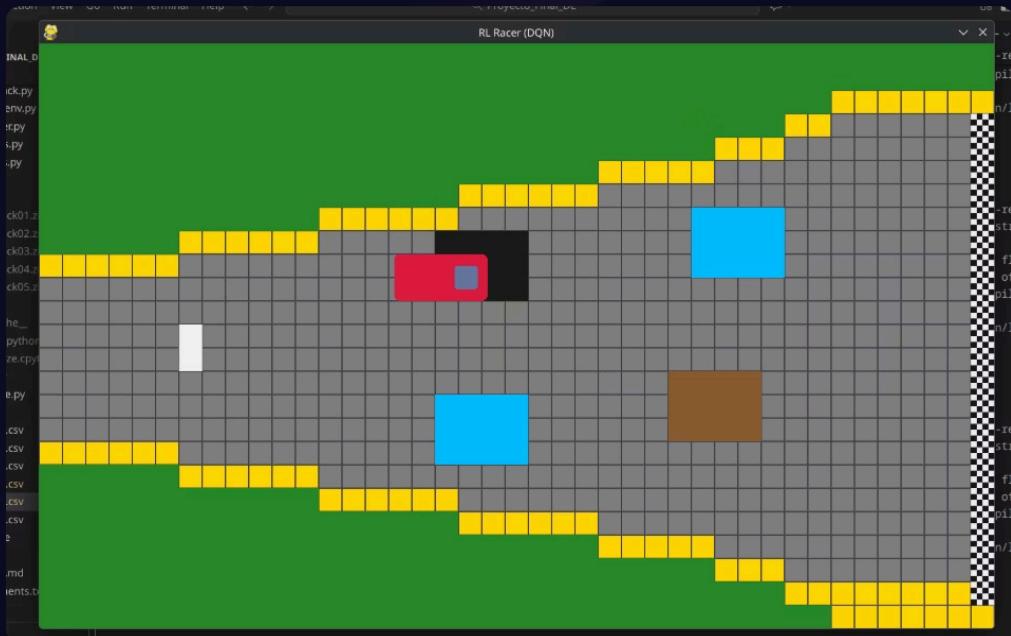
Implementar truncamiento después de 500 pasos para prevenir episodios infinitos

04

Entrenamiento Gradual

Comenzar con pistas simples antes de introducir complejidad avanzada

Arquitectura de la Solución



Sistema de aprendizaje por refuerzo donde el agente aprende mediante interacción directa. Observaciones capturan contexto local en formato matricial multicanal.

Componentes Clave

- Decisiones discretas sobre dirección y aceleración
- Retroalimentación mediante recompensas
- Simulación realista de efectos de superficie
- Convergencia gradual hacia políticas efectivas

Entrenamiento/Resultados

```
== Resumen de entrenamiento ==
```

```
episodios: 322
retorno_prom: 14.055919866459627
retorno_std: 11.223144917018663
retorno_mejor: 20.751355
largo_prom: 31.024844720496894
exitos: 240
choques: 83
tasa_exito: 0.7453416149068323
tiempo_total: 54.738423109054565
```

```
Modelo guardado en: models/dqn_track04.zip
```

Track	Episodios	Tasa de éxito	Tiempo Total (s)
1	214	45.76%	4.8
2	294	55.78%	67
3	821	57.91%	2812
4	1919	59.67%	3069

Track	Episodios	Tasa de éxito	Tiempo Total (s)
1	5345	55.21%	4392
2	5987	58.43%	4459
3	5363	60.77%	4731
4	6423	57.92%	4932

Conclusiones Clave

Efectividad del DQN

La implementación demostró ser efectiva para navegación autónoma, logrando políticas que balancean velocidad, seguridad y precisión mediante CNNs que procesan observaciones egocéntricas.

Diseño de Recompensas

El sistema de recompensas resultó fundamental, combinando progreso incremental, penalizaciones por tiempo y desviación, junto con recompensas terminales.

Arquitectura de Mapas

La simulación de efectos de superficie como aceite, terracería y boost añade realismo y complejidad al entorno.

Complejidad Realista

Los efectos de superficie añaden realismo, requiriendo que el agente aprenda modelos internos de dinámicas complejas mediante experiencia directa.



Referencias Bibliográficas

1. Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533.
2. Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep Reinforcement Learning with Double Q-learning. *AAAI Conference*, 30(1), 2094-2100.
3. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
4. Raffin, A., et al. (2021). Stable-Baselines3: Reliable Reinforcement Learning Implementations. *JMLR*, 22(268), 1-8.
5. Brockman, G., et al. (2016). OpenAI Gym. *arXiv preprint arXiv:1606.01540*.