



AGENTE DE ENTREGAS CON APRENDIZAJE POR REFUERZO

VERDUGUEZ CHOQUEVILLCA SHARIEL AYLIN

¿QUE ES LO
QUE SE
QUIERE
MOSTRAR?

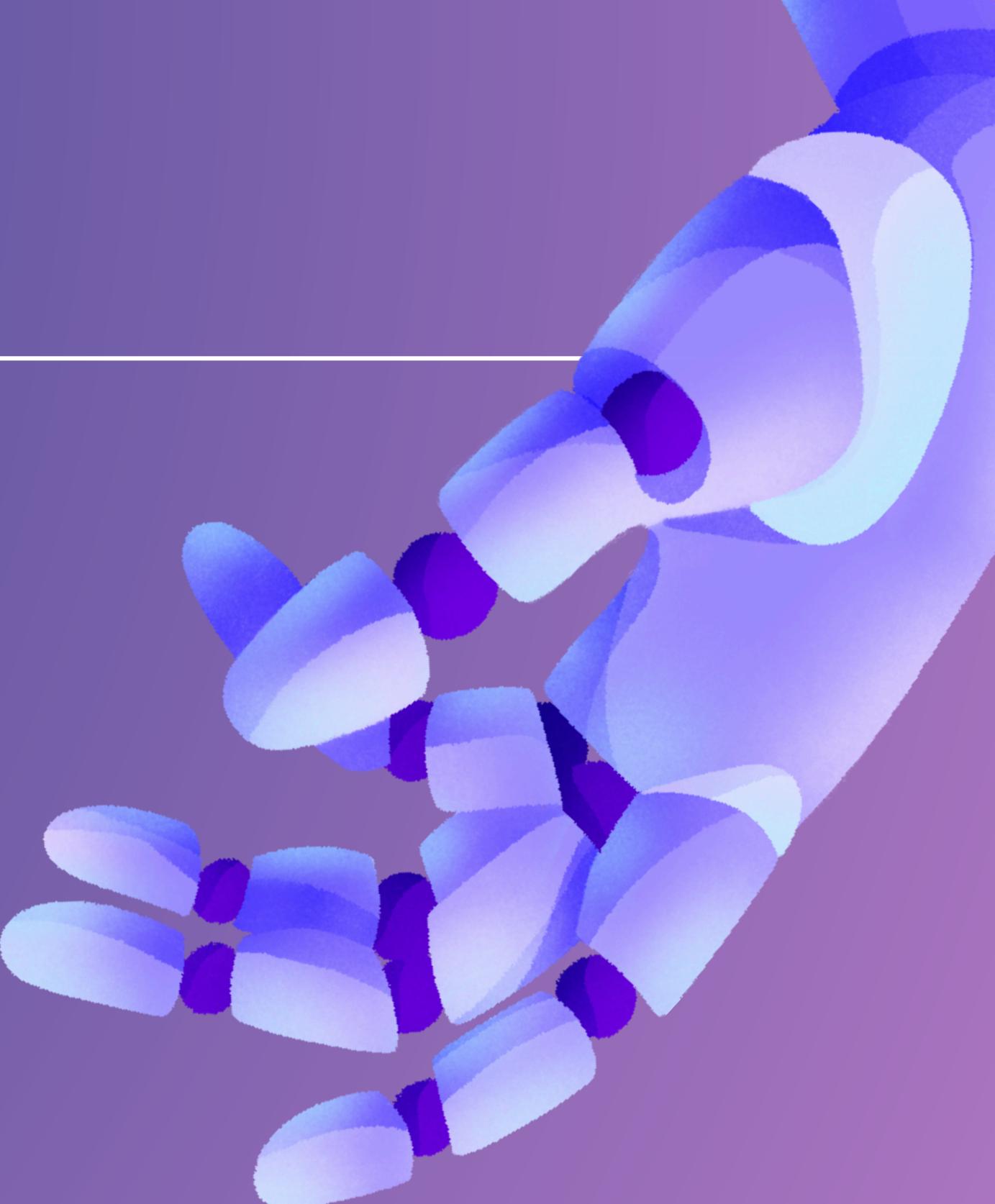


PLANTEAMIENTO DEL PROBLEMA

Tengo muchos clientes y un depósito. ¿En qué orden visito a los clientes para entregar a todos, recorriendo lo menos posible?

OBJETIVO

Busco desarrollar un agente de entregas que aprenda a planificar el orden en que visita a los clientes usando aprendizaje por refuerzo, para minimizar tiempos.



DISEÑO DEL ENTORNO

El estado de mi agente es un vector de la forma x,y,p_1,\dots,p_n

La acción es simplemente elegir un cliente por su índice, entre 0 y $N-1$.

FUNCION DE RECOMPENSA

1

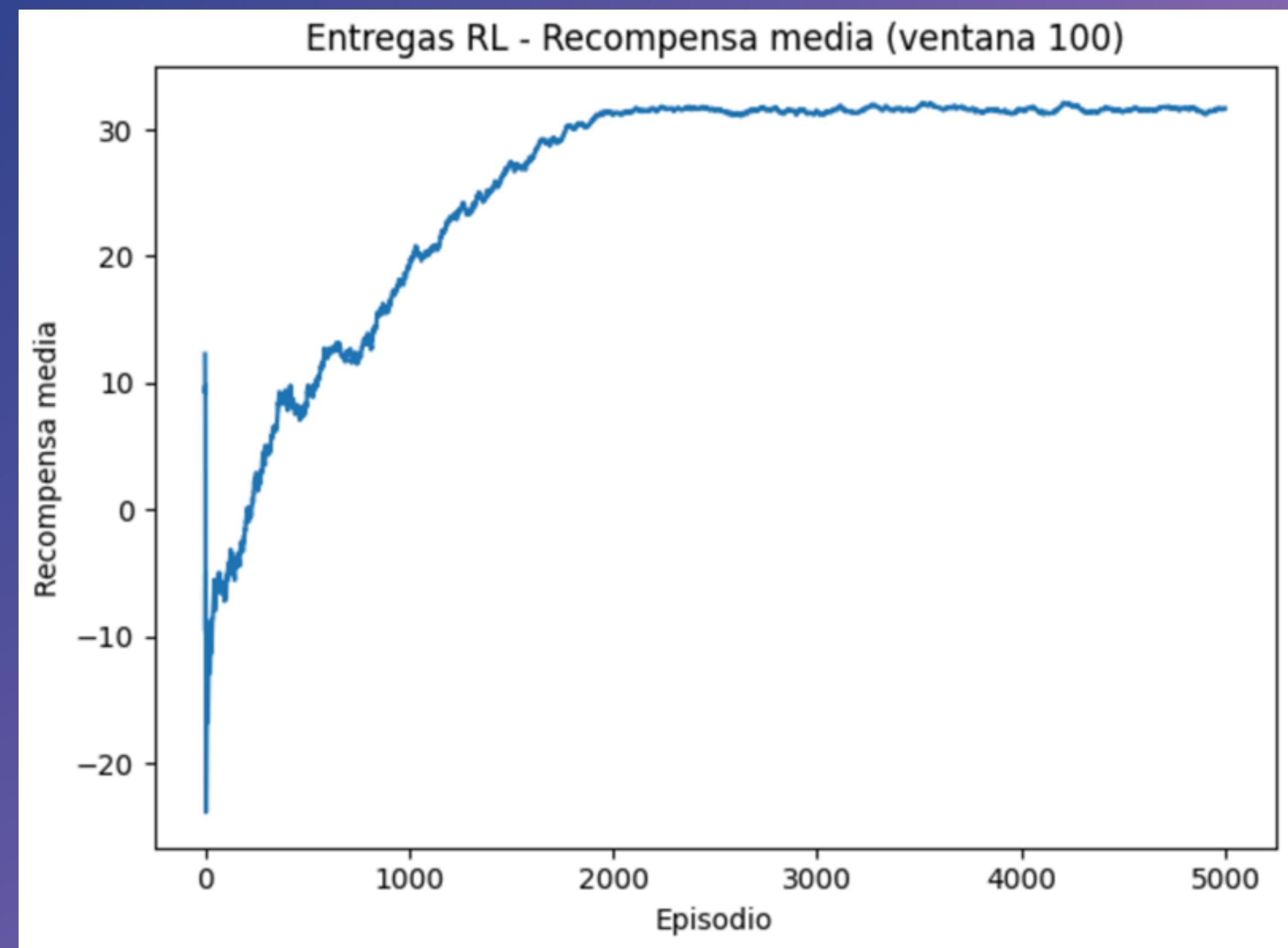
+10 por entregar a un cliente
nuevo -distancia recorrida.

2

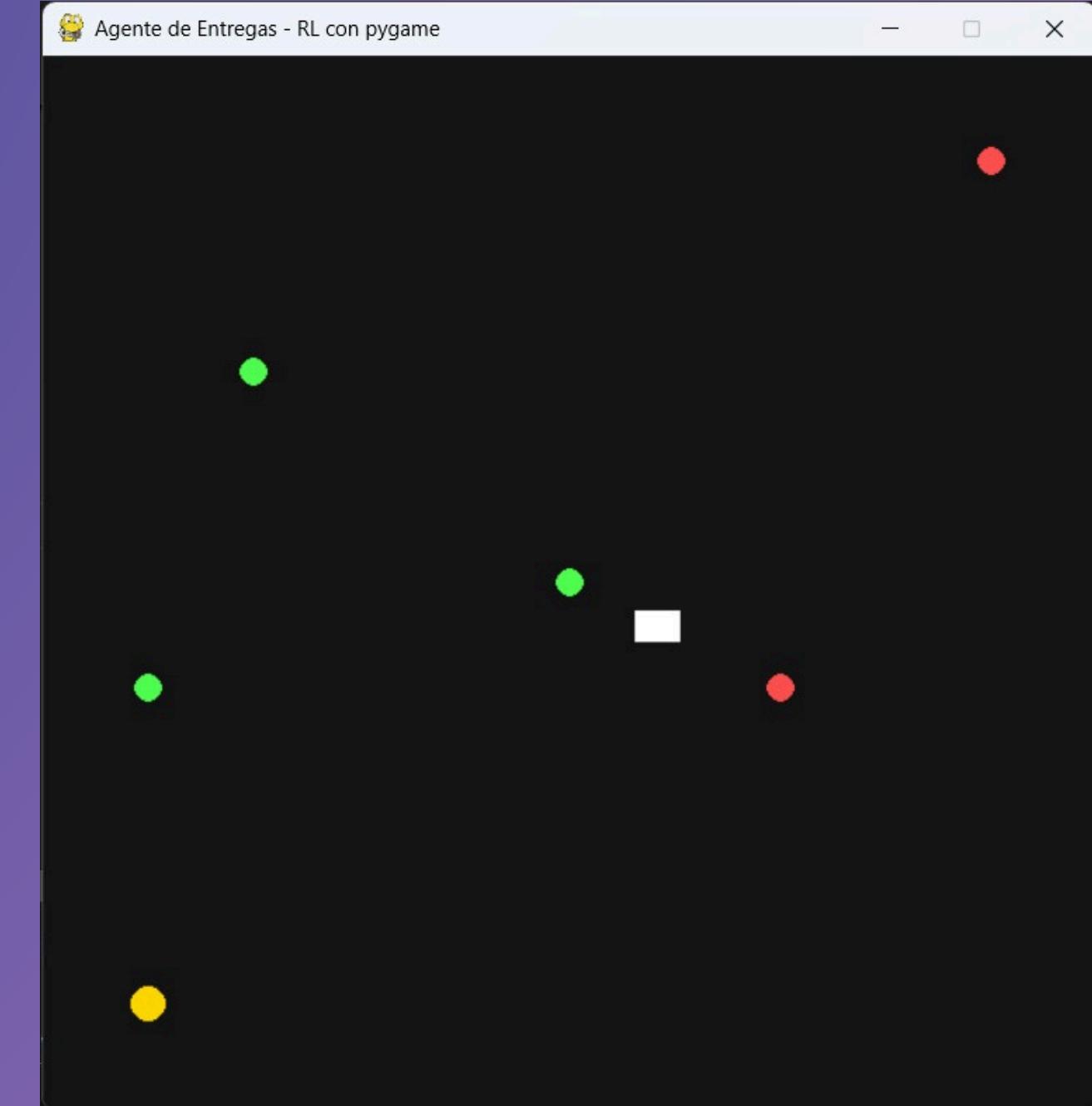
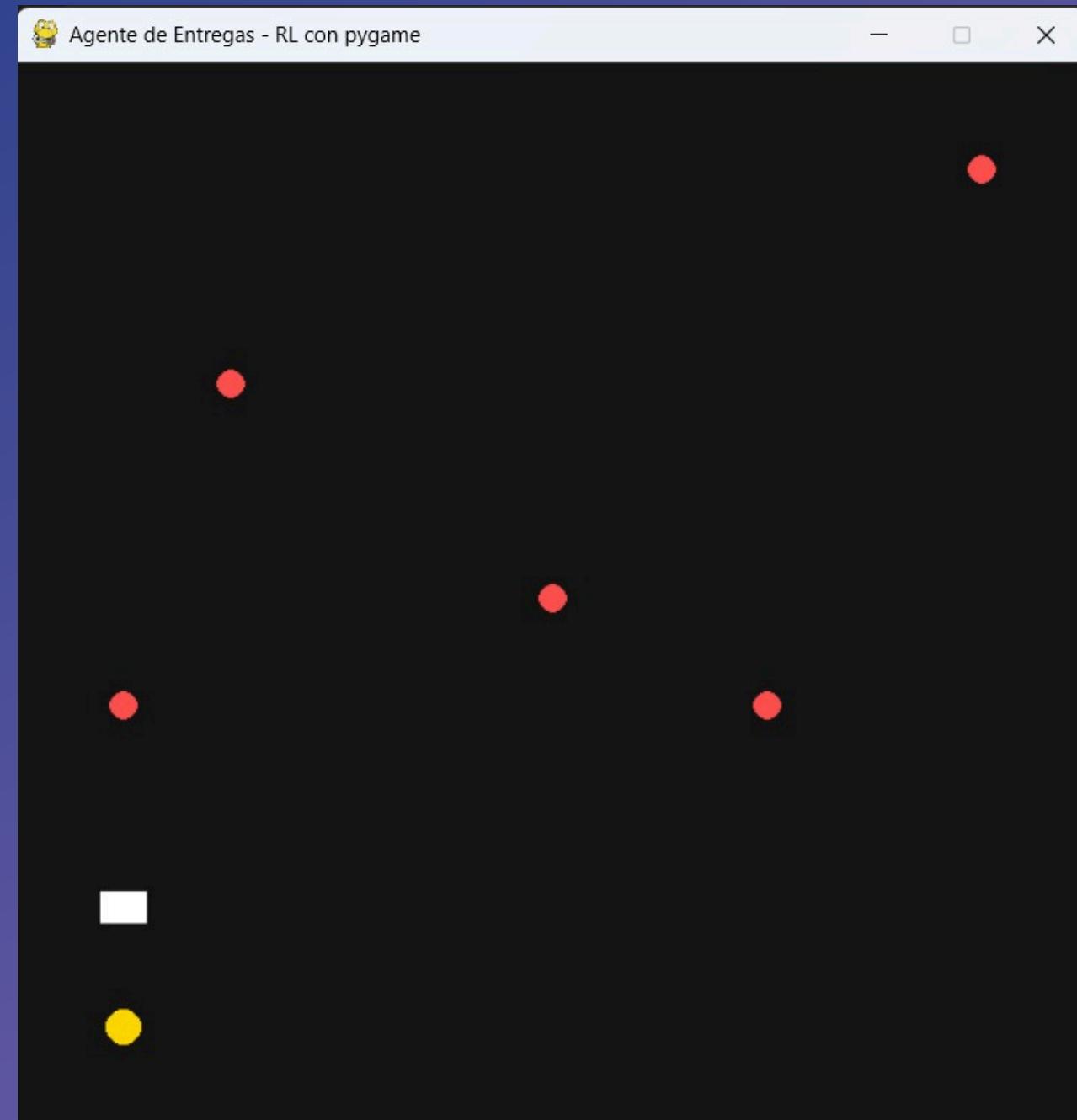
-5 si intenta ir a un cliente ya
atendido.



CURVA DE APRENDIZAJE



VISUALIZACIÓN CON PYGAME



CONCLUSIONES

- Modelamos un problema de entregas y entrenamos un agente con Q-Learning.
- La política aprendida reduce distancia y evita entregas repetidas.
- La curva de aprendizaje se estabiliza (~30-32 de recompensa media), señal de convergencia.

GRACIAS POR
SU ATENCIÓN

