

UNIVERSIDAD DEL VALLE DE GUATEMALA

CC3104 – Aprendizaje por Refuerzo

Ing, Javier Josué Fong Guzmán



Excelencia que trasciende

DELVALLE
GRUPO EDUCATIVO

Laboratorio 7

Function Approximation

José Pablo Orellana 21970

Diego Alberto Leiva 21572

Guatemala, 28 de agosto de 2025

Descripción general de la implementación

Se implementaron tres agentes de Aprendizaje por Refuerzo con aproximación de funciones en el entorno CartPole-v1 de Gymnasium, utilizando Monte Carlo (MC) como mecanismo de actualización. Los agentes aproximan la función acción-valor $Q(a)$ a partir de pares (estado, acción) \rightarrow retorno acumulado.

- Árbol de Decisión: modelo de regresión basado en un único árbol, ajustado periódicamente con los datos acumulados.
- Random Forest: ensamble de árboles de decisión, que mejora la estabilidad y generalización respecto a un solo árbol.
- Red Neuronal Feed-Forward: MLP con dos capas ocultas (128 neuronas cada una), entrenado con retropropagación y MSE.

En todos los casos, el vector de entrada combina el estado escalado y la acción codificada one-hot. La política fue ϵ -greedy con decaimiento lineal de ϵ .

Evolución de los agentes

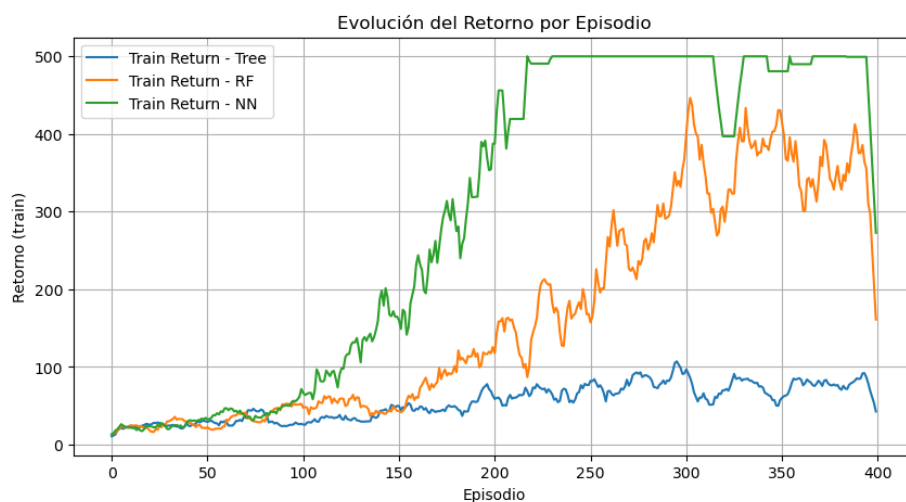


Figura 1: Evolución de los agentes en 500 pasos

Durante el entrenamiento se observan patrones distintos en la progresión de cada aproximador:

- Árbol de Decisión: el retorno aumentó lentamente, pero se estancó en valores bajos (<120). Nunca logró resolver episodios completos (500 pasos).
- Random Forest: mostró un incremento más consistente, alcanzando retornos cercanos a 400 a partir del episodio ~ 300 . Aunque no siempre resolvió los episodios, mantuvo un desempeño medio-alto con tendencia ascendente.

- Red Neuronal: aprendió rápidamente. Desde el episodio ~120 logró retornos máximos de 500, con tramos de éxito sostenido del 100%. Sin embargo, hacia el final del entrenamiento presentó cierta inestabilidad (descensos puntuales).

Resultados

Para comparar los agentes de forma objetiva, se calcularon los promedios de las últimas 3 evaluaciones:

Agente	Retorno Eval Promedio	Tasa de Éxito Promedio
Árbol de Decisión	68.3	0%
Random Forest	370.9	43.3%
Red Neuronal	407.5	53.3%

Tabla 1: Comparación de evaluaciones de los agentes

Esto confirma que:

- El Árbol de Decisión no fue capaz de resolver el entorno.
- El Random Forest generó políticas más estables y con éxito parcial.
- La Red Neuronal alcanzó el mejor desempeño, resolviendo episodios completos y superando en promedio a los otros métodos, aunque con mayor varianza.

Conclusiones

La Red Neuronal, que alcanzó antes el máximo de 500 pasos y resolvió el entorno es el aproximador más eficiente en aprender.

El Random Forest, genero las políticas más estables, con menor varianza y un desempeño alto, pero sin llegar al máximo absoluto de la NN.

Los Árboles (Tree, RF) son simples de entrenar y rápidos, pero limitados en capacidad de generalización (el árbol simple se sobreajusta). El Random Forest mejora la robustez, aunque requiere más memoria y no alcanza el rendimiento máximo de la NN.

La Red Neuronal tiene una mayor capacidad de representación, alcanzando la política óptima, aunque puede presentar inestabilidad en entrenamientos prolongados y requiere más cuidado en hiperparámetros y recursos de cómputo.