

En parejas, implemente y entrene un agente que aprenda a jugar *Blackjack* utilizando el entorno “*Blackjack-v1*” de la librería Gymnasium, aplicando métodos de control por Monte Carlo para encontrar una política óptima.

Instrucciones Generales

1. **Entorno y Código Base:**
 - Utilice el jupyter notebook *montecarlo.ipynb* como base para implementar su agente.
 - Utilice el entorno *Blackjav-v1*
2. **Implementación de Métodos:**
 - Implemente una política inicial aleatoria
 - Escriba una función que simule episodios completos desde el inicio del juego hasta un estado terminal.
 - Estime la el valor $Q(s,a)$
 - En base a los valores obtenidos, elija la mejor política que haya encontrado.
 - Mida el rendimiento promedio de la política final contra el entorno.

Entregables

- Código fuente en un Jupyter Notebook
- Breve informe (máximo 2 páginas) que incluya:
 - Descripción general de implementación.
 - Evolución de la política a lo largo de los episodios.
 - Resultados de la política óptima.
 - Conclusiones