



Matemáticas y Datos in a nutshell.

VIII Big Data Reinforcement Learning (Offline RL)



El aprendizaje por refuerzo (RL) es un método para construir un modelo matemático llamado política que tome en cuenta datos secuenciales en un problema de AI. Tradicionalmente ha sido utilizado con éxito para diversas aplicaciones como:

- a. Vencer a campeones de ajedrez o Go.
- b. Optimizar los recursos energéticos de un Data Center.
- c. Controlar robots que realicen una tarea de AI.
- d. Sistemas de conversación parecidos a Chat GPT.





Big Data Reinforcement Learning (Offline RL)



La mayor parte de los algoritmos de RL son del tipo Off-Policy y On-Policy y tienen una desventaja respecto a métodos supervisados en machine learning. Durante el periodo de exploración, las secuencias de estados que visitarán están determinados por la iteración de la política que el algoritmo haya entrenado hasta ese momento lo cual significa que podrían no ser suficientemente realistas.

Recientemente se han desarrollado técnicas lo que se conoce como Offline Reinforcement Learning cuya fase de entrenamiento no depende de la política actual sino que aprovecha una base de datos suficientemente grande como las que se encontraría uno en los LLM. Es por eso que en esta presentación nos referimos a estas técnicas como Big Data RL.

