



MODELOS DE APRENDIZAJE

APAZA DURAN JOSE JEREMY

MAROCHO CAPA GABRIELA IRMA

QUISPE CCOPA EVELYN

RIVAS ARANIBAR JEREMIAS ABEL

SILVA TRESIERRA EFRAM ABDEEL SAMUEL

TURNER MINAYA CESAR ROGELIO



PRESENTACIÓN



Hoy vamos a explorar dos poderosos enfoques en el mundo del aprendizaje automático y la inteligencia artificial: los Modelos de Mezcla Gaussiana (GMM) y el Q-Learning. Aunque ambos modelos se aplican en contextos diferentes, comparten el objetivo de ayudarnos a extraer patrones y tomar decisiones inteligentes a partir de datos.

A lo largo de esta clase, profundizaremos en estos dos modelos, discutiremos sus fundamentos teóricos, y los veremos en acción con ejemplos prácticos. Esto nos permitirá apreciar su potencial en la resolución de problemas complejos en campos tan variados como el análisis de datos y la inteligencia artificial.



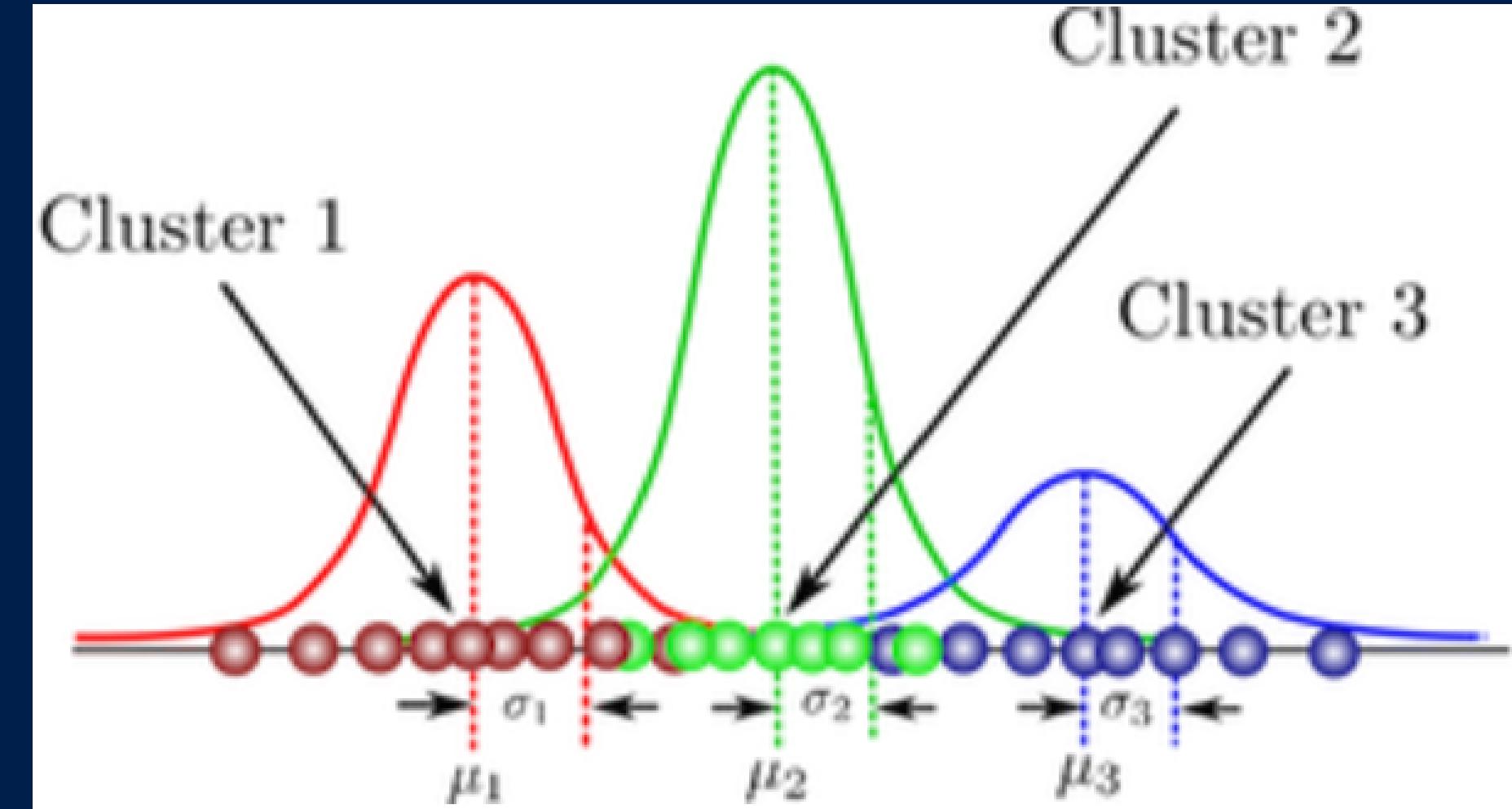


REPOSITORIO DE GITHUB

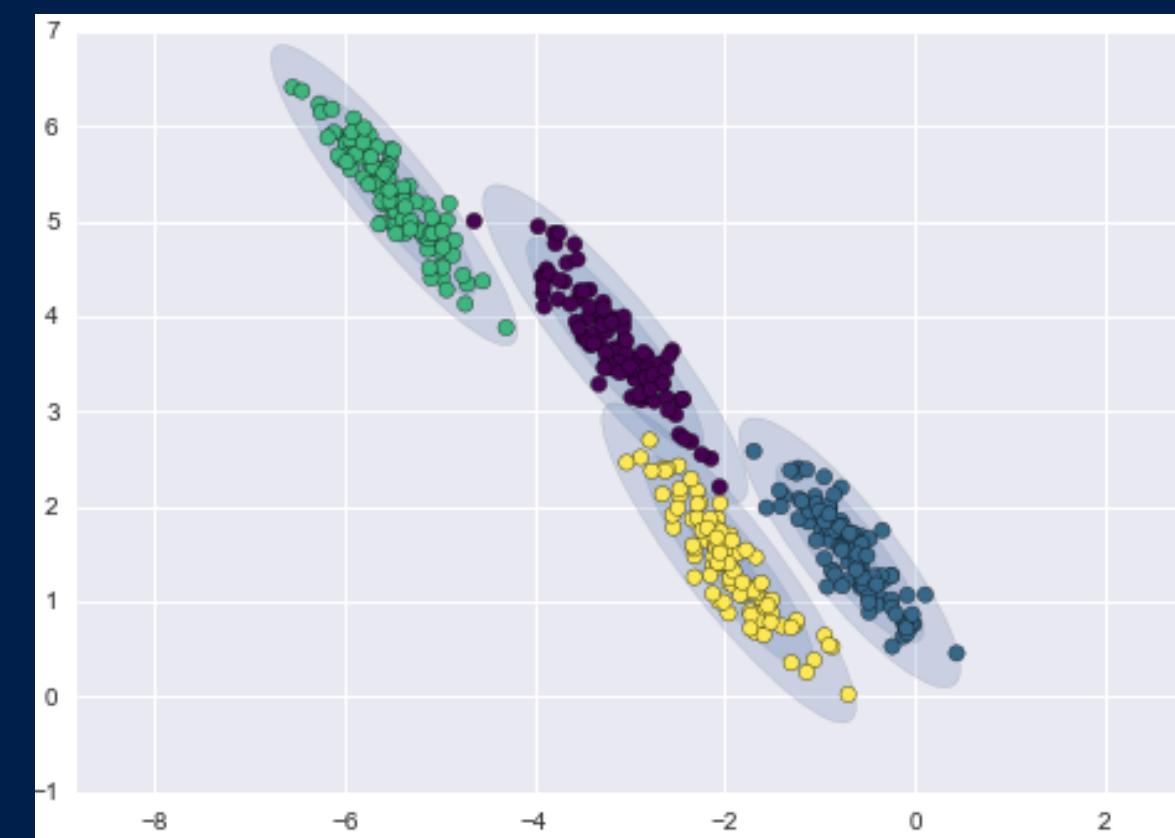
DESARROLLO

Clustering GMM (Gaussian Mixture Models)

Es un modelo probabilístico que asume que los datos provienen de una mezcla de distribuciones gaussianas. Cada cluster se representa mediante una distribución normal, y el GMM modela la probabilidad de que un dato pertenezca a cada grupo.



La técnica implica mezclar múltiples distribuciones gaussianas. En lugar de identificar los conglomerados por los centroides más cercanos, se ajustan k gaussianas a los datos y se estiman parámetros como la media y la varianza para cada cluster, así como su peso. Con estos parámetros, se pueden calcular las probabilidades de pertenencia de cada punto de datos a los conglomerados.





ETAPAS (MAXIMIZACIÓN DE EXPECTATIVAS)

Para estimar los parámetros del GMM (medias, varianzas, y pesos), se utiliza el Algoritmo Expectation-Maximization (EM)



Etapa de Expectativa (E)

Calcula la probabilidad de que cada punto de datos pertenezca a cada uno de los clusters, basándose en los parámetros actuales del modelo



Etapa de Maximización (M)

Actualiza los parámetro del modelo (medias, varianzas y pesos de los clusters) utilizando las probabilidades calculadas en la etapa de expectativa.

DESVENTAJAS

- Computacionalmente intensivo: Requiere más tiempo de cálculo en comparación con métodos más simples como K-means.
- Selección de K: La elección del número de componentes (clusters) no siempre es evidente y puede requerir validación adicional



APLICACIÓN

- Segmentación de imágenes: Agrupando píxeles similares en una imagen para facilitar el procesamiento y la comprensión.
- Reconocimiento de patrones: Identificando y clasificando patrones en datos complejos.
- Análisis de datos financieros: Agrupando transacciones o comportamientos de clientes para detectar fraudes o segmentar mercados.



IMPLEMENTACIÓN DEL MODELO GMM

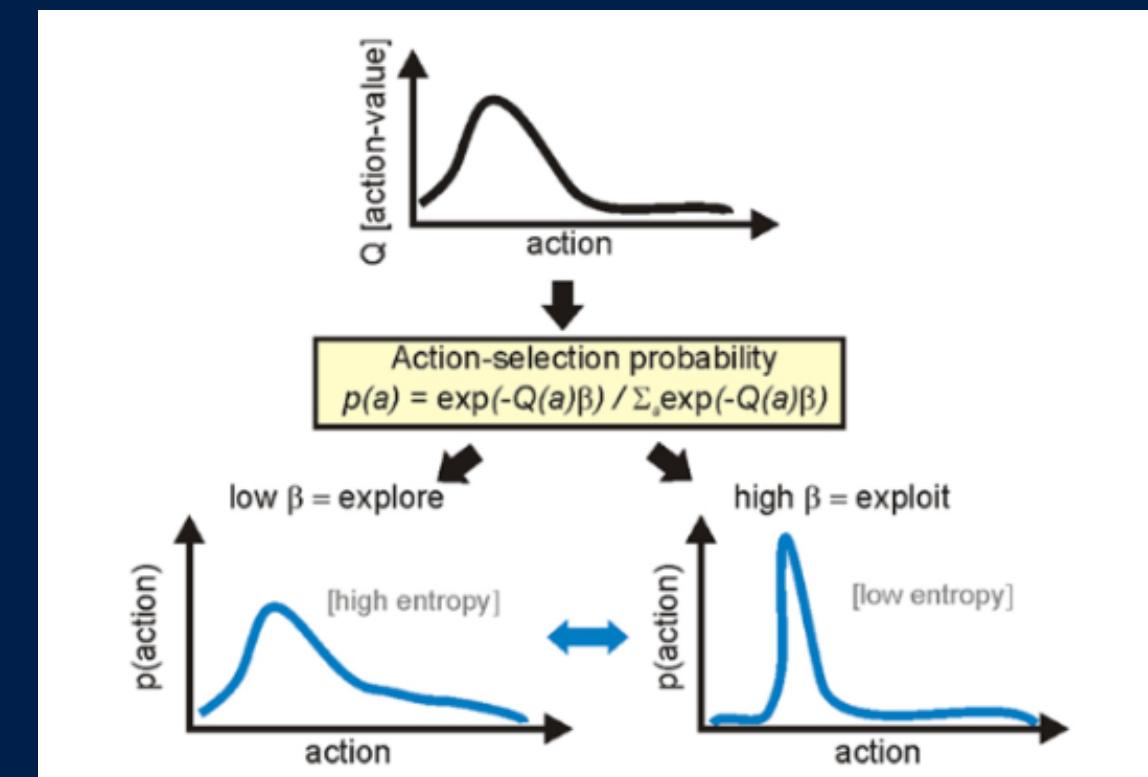
IMPLEMENTACIÓN DEL EJEMPLO EN EL COLAB.



Aprendizaje por refuerzo

Q-Learning

El algoritmo Q-Learning es un enfoque específico dentro de la Inteligencia Artificial y el aprendizaje por refuerzo. Está diseñado para encontrar la política de selección de acciones óptima para cualquier proceso de decisión de Markov (MDP) finito. El objetivo del Q-learning es aprender la calidad de las acciones, que está representada por los valores Q. Estos valores Q se utilizan para determinar la mejor acción a realizar en un estado determinado. El aprendizaje por refuerzo no requiere datos. En su lugar, aprende del entorno y del sistema de recompensas para tomar mejores decisiones.





TERMINOLOGÍAS CLAVE EN Q-LEARNING

- **Estado(s)**: Posición actual del agente en el entorno.
- **Acción(a)**: Paso dado por el agente en un estado determinado.
- **Recompensas**: Por cada acción, el agente recibe una recompensa y una penalización.
- **Episodios**: El final de la etapa, donde los agentes no pueden emprender nuevas acciones. Se produce cuando el agente ha alcanzado el objetivo o ha fracasado.
- **$Q(St+1, a)$** : Valor Q óptimo esperado de realizar la acción en un estado concreto.
- **$Q(St, At)$** : Es la estimación actual de $Q(St+1, a)$.
- **Tabla Q**: El agente mantiene la tabla Q de conjuntos de estados y acciones.
- **Diferencias temporales (TD)**: Se utiliza para estimar el valor esperado de $Q(St+1, a)$ utilizando el estado y la acción actuales y el estado y la acción anteriores.

¿CÓMO FUNCIONA Q-LEARNING?

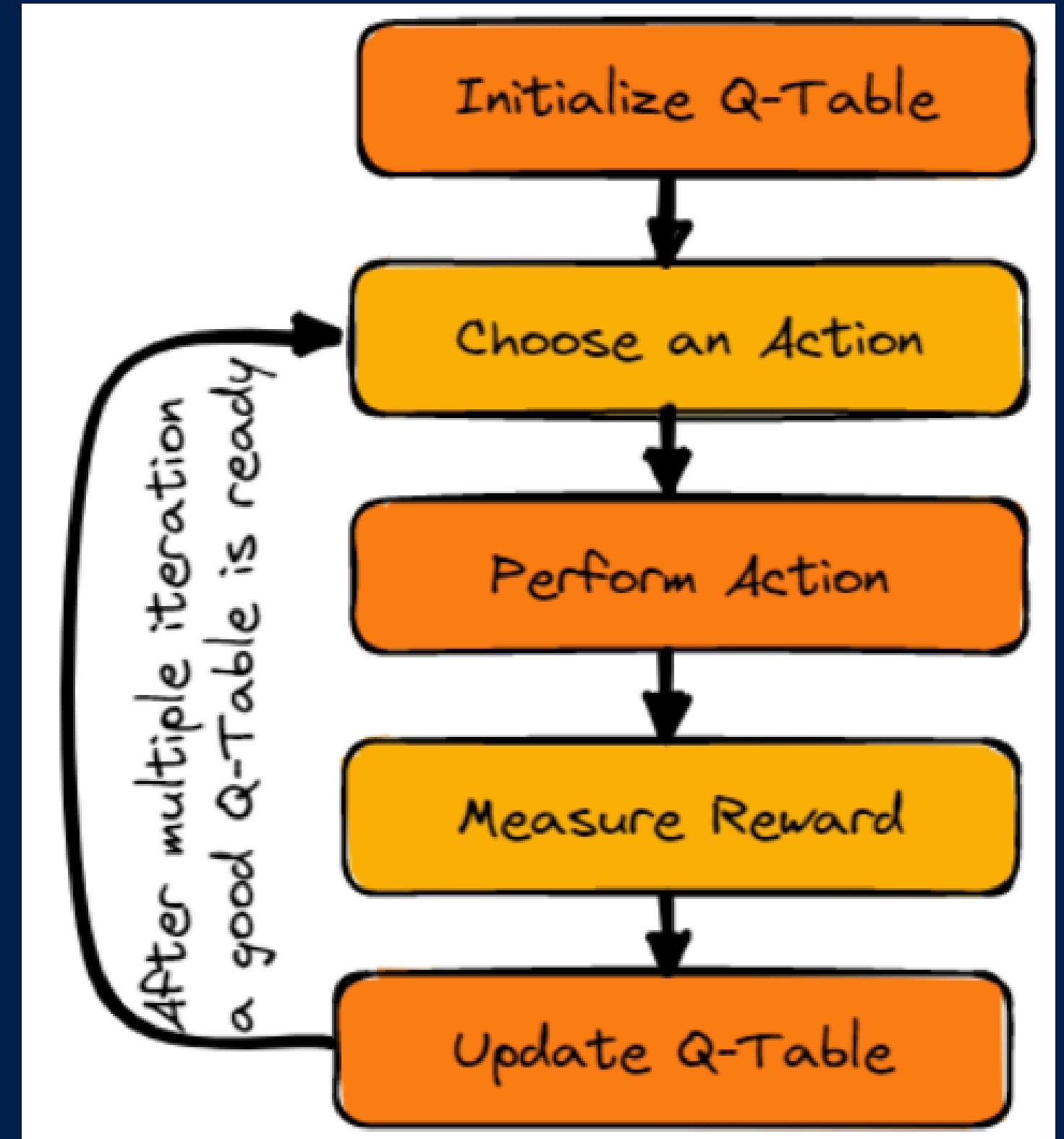
Mesa Q: El agente utilizará una tabla Q para tomar la mejor acción posible en función de la recompensa esperada para cada estado del entorno. En palabras sencillas, una tabla Q es una estructura de datos de conjuntos de acciones y estados, y utilizamos el algoritmo de aprendizaje Q para actualizar los valores de la tabla.

Función Q: La función Q utiliza la ecuación de Bellman y toma como entrada el estado(s) y la acción(a). La ecuación simplifica el cálculo de los valores de estado y de estado-acción.

$$Q^\pi(s_t, a_t) = \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots | s_t, a_t]$$

↓ ↓ ↓

Q-Values for the state
given a particular state Expected discounted
cumulative reward Given the state and action



ALGORITMO DE APRENDIZAJE Q



CASO DE PRUEBA

Iniciar Q-Table

Inicializaremos la tabla Q con valores a 0.

	→	←	↑	↓
Start	0	0	0	0
Idle	0	0	0	0
Hole	0	0	0	0
End	0	0	0	0

Elija una acción



Realizar una acción

	→	←	↑	↓
Start	0	0	0	1
Idle	0	0	0	0
Hole	0	0	0	0
End	0	0	0	0

Medir las recompensas

- La recompensa por alcanzar el objetivo es +1
- La recompensa por tomar el camino equivocado (caer en el agujero) es 0
- La recompensa por inactividad o movimiento en el lago helado también es 0.

Actualizar tabla Q

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$$

New Q-value estimation
 Former Q-value estimation
 Learning Rate
 Immediate Reward
 Discounted Estimate optimal Q-value of next state
 Former Q-value estimation

TD Target
 TD Error

	→	←	↑	↓
Start	0	1	0	0
Idle	2	0	0	3
Hole	0	2	0	0
End	1	0	0	0



Explotacion y Exploracion



Exploración (Exploration):

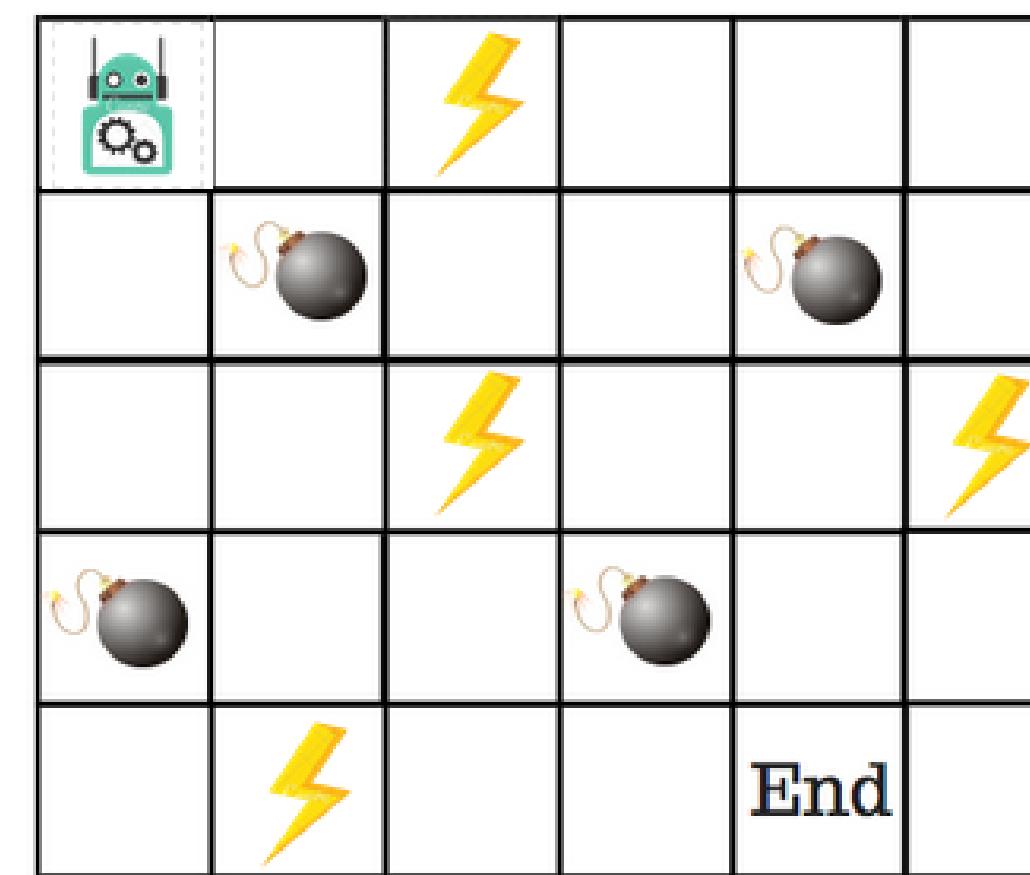
La exploración significa que el agente decide probar nuevas acciones, incluso si no está seguro de cuál será el resultado, para aprender más sobre el entorno.

Objetivo: Descubrir nuevas opciones que podrían conducir a mejores recompensas a largo plazo.

Explotación (Exploitation):

La explotación significa que el agente elige la mejor acción posible basándose en lo que ha aprendido hasta el momento, es decir, en su tabla Q.

Objetivo: Maximizar la recompensa a corto plazo utilizando el conocimiento actual.





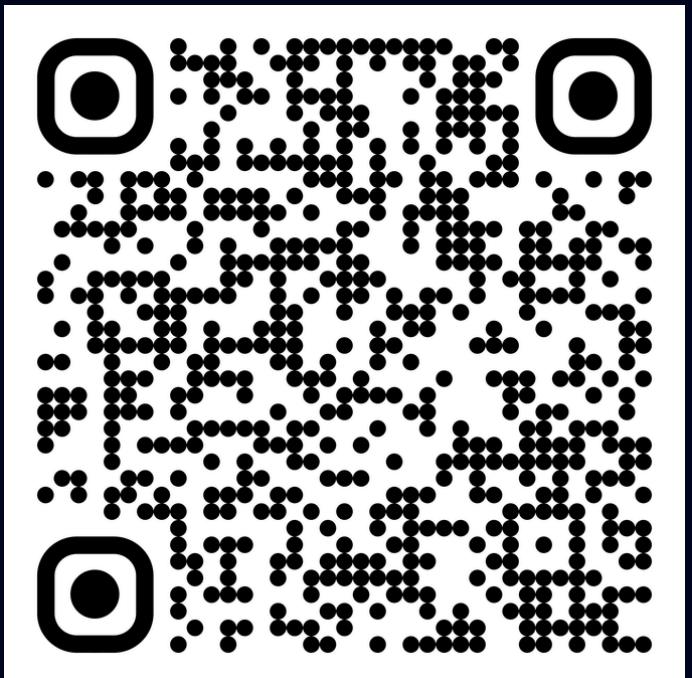
CONCLUSIÓN

El presente proyecto demuestra la aplicabilidad de dos enfoques de aprendizaje no supervisado, el Clustering GMM y el aprendizaje por refuerzo Q-Learning, en la predicción del rendimiento de equipos de fútbol de la Premier League y su aplicación en un caso de prueba. Mediante la implementación de estos modelos, se logra identificar patrones ocultos en los datos y optimizar las decisiones en un entorno dinámico. Los resultados obtenidos muestran la capacidad de los modelos para adaptarse a escenarios cambiantes y mejorar su rendimiento con el tiempo. Sin embargo, para optimizar aún más la precisión, es necesario ajustar continuamente los parámetros y explorar nuevas técnicas complementarias.





GRACIAS!!



REPOSITORIO DE GITHUB