### Conclusión

El proyecto implementa exitosamente un agente de Q-learning para resolver un rompecabezas numérico en un entorno simulado. Aquí están los puntos clave:

1. \*\*Definición del Entorno\*\*:

- El rompecabezas es una matriz de 4x5 con números del 1 al 20.

- El agente puede moverse en cuatro direcciones y recibe recompensas o penalizaciones según sus movimientos.

2. \*\*Agente de Q-learning\*\*:

- Utiliza una tabla Q para aprender valores de acción-estado.

- Emplea una política epsilon-greedy para equilibrar exploración y explotación.

- Mejora sus decisiones basándose en las recompensas recibidas.

3. \*\*Entrenamiento y Resultados\*\*:

- El agente mejora su rendimiento con el tiempo, necesitando menos pasos para resolver el rompecabezas y acumulando más recompensas.

- Demuestra la efectividad del Q-learning para problemas de control y toma de decisiones.

### Beneficios y Aplicaciones Futuras

- \*\*Aprendizaje Autónomo\*\*: El agente aprende sin supervisión, mostrando el potencial del aprendizaje por refuerzo en situaciones complejas.

- \*\*Escalabilidad\*\*: La metodología puede adaptarse a otros problemas en diversos dominios como robótica y planificación de rutas.

- \*\*Optimización de Parámetros\*\*: Futuros trabajos pueden optimizar los parámetros de aprendizaje para mejorar la eficiencia.

En resumen, este proyecto demuestra cómo el Q-learning puede resolver problemas secuenciales y proporciona una base para futuras investigaciones en aprendizaje por refuerzo.