



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat d'Informàtica de Barcelona



Sistemas Inteligentes Distribuidos

Agente de aprendizaje por refuerzo para el entorno Cliff Walking

Lluc Martínez Busquets

Eric Medina León

Àlex Rodríguez Rodríguez

2024/2025 Q2

Índice

1	Introducción	2
2	Value Iteration	2
2.1	Implementación	2
2.2	Experimentos	2
3	Direct Estimation	2
3.1	Implementación	2
3.2	Experimentos	2
4	Q-learning	2
4.1	Implementación	2
4.2	Experimentos	2
5	Reinforce	2
5.1	Implementación	2
5.2	Experimentos	2
6	Comparación de Métodos	2
7	Conclusiones	2

-
1. Introducción
 2. Value Iteration
 - 2.1. Implementación
 - 2.2. Experimentos
 3. Direct Estimation
 - 3.1. Implementación
 - 3.2. Experimentos
 4. Q-learning
 - 4.1. Implementación
 - 4.2. Experimentos
 5. Reinforce
 - 5.1. Implementación
 - 5.2. Experimentos
 6. Comparación de Métodos

Ventajas e inconvenientes de cada método, etc.
 7. Conclusiones