

Libro "Artificial Intelligence: A modern approach"

El libro comienza introduciendo la inteligencia artificial como el estudio de "agentes inteligentes". Estos son sistemas que perciben su entorno y toman decisiones para maximizar sus posibilidades de éxito. Los autores definen a un agente como cualquier cosa que pueda percibir su entorno a través de sensores y actuar en ese entorno como actuadores. En IA, un agente inteligente es aquel que puede tomar decisiones autónomas para alcanzar un objetivo basándose en información.

Tipos de agentes

Se exploran distintos tipos de agentes, según sus comportamientos y capacidades, entre ellos los agentes reactivos (que responden a estímulos sin un modelo interno complejo) y los agentes deliberativos (que planifican y toman decisiones considerando objetivos futuros). Además, se introduce la noción de "racionalidad", según la cual un agente es racional si selecciona las acciones que maximiza sus chances de alcanzar sus objetivos.

Problemas y soluciones en IA

La IA se presenta como la búsqueda de soluciones para problemas complejos, desde los más simples hasta aquellos con gran cantidad de variables. El libro abarca diferentes soluciones como las mediante algoritmos de búsqueda, desde las más simples (como la búsqueda en amplitud y profundidad) hasta algoritmos más avanzados como A* (una búsqueda optimizada) y algoritmos de búsqueda probabilística.

Resolver problemas mediante búsqueda

Un tema clave del libro es cómo resolver problemas mediante algoritmos de búsqueda. La búsqueda es un ensayo fundamental en la IA, ya que permite explorar posibles soluciones a un problema de manera estructurada. Los autores discuten métodos clásicos como la búsqueda ciega (que no utiliza información adicional) y la búsqueda informática (que utiliza heurística para guiar la búsqueda).

Aprendizaje automático

Una gran parte del libro es dedicada al aprendizaje automático (machine learning), que es la capacidad de un sistema para aprender a partir de datos. Este campo está en un auge debido a los avances en el procesamiento de gran volumen de datos y la mejora de algoritmos que permiten a las máquinas "aprender" de manera eficiente.

Russell y Norvig abordan los tres principales tipos de aprendizajes: supervisado, no supervisado y por refuerzo. El aprendizaje supervisado se basa en el uso de ejemplos etiquetados para entrenar un ~~modelo~~ modelo. El no supervisado intenta encontrar estructuras o patrones en datos no etiquetados. El aprendizaje por refuerzo, por su parte, está relacionado con el comportamiento de los agentes que interactúan con su entorno para maximizar una señal de ~~recompensa~~ recompensa.