Fundamentals of Artificial Intelligence and Knowledge Representation

Academic Year 2023 – 2024
Alma Mater Studiorum · University of Bologna

Contents

1 Introduction

1.1 AI systems classification

1.1.1 Intelligence classification

Intelligence is defined as the ability to perceive or infer information and to retain the knowledge for future use.

Weak AI aims to build a system that acts as an intelligent system. Weak AI

Strong AI aims to build a system that is actually intelligent. Strong AI

1.1.2 Capability classification

General AI systems able to solve any generalized task. General AI

Narrow AI systems able to solve a particular task. Narrow AI

1.1.3 AI approaches

Symbolic AI (top-down) Symbolic representation of knowledge, understandable by humans. Symbolic AI

Connectionist approach (bottom up) Neural networks. Knowledge is encoded and not understandable by humans. Connectionist approach

1.2 Symbolic AI

Deductive reasoning Conclude something given some premises (general to specific). It is unable to produce new knowledge. Deductive reasoning

    Example. "All men are mortal" and "Socrates is a man" → "Socrates is mortal"

Inductive reasoning A conclusion is derived from an observation (specific to general). Produces new knowledge, but correctness is not guaranteed. Inductive reasoning

    Example. "Several birds fly" → "All birds fly"

Abduction reasoning An explanation of the conclusion is found from known premises. Differently from inductive reasoning, it does not search for a general rule. Produces new knowledge, but correctness is not guaranteed. Abduction reasoning

    Example. "Socrates is dead" (conclusion) and "All men are mortal" (knowledge) → "Socrates is a man"

Reasoning by analogy Principle of similarity (e.g. k-nearest-neighbor algorithm). Reasoning by analogy

    Example. "Socrates loves philosophy" and Socrates resembles John → "John loves philosophy"

Constraint reasoning and optimization Constraints, probability, statistics. Constraint reasoning

1.3 Machine learning

1.3.1 Training approach

Supervised learning Trained on labeled data (ground truth is known).
Suitable for classification and regression tasks.

Unsupervised learning Trained on unlabeled data (the system makes its own discoveries).
Suitable for clustering and data mining.

Semi-supervised learning The system is first trained to synthesize data in an unsupervised manner, followed by a supervised phase.

Reinforcement learning An agent learns by simulating actions in an environment with rewards and punishments depending on its choices.

1.3.2 Tasks

Classification Supervised task that, given the input variables X and the output (discrete) categories Y, aims to approximate a mapping function $f : X \to Y$.

Regression Supervised task that, given the input variables X and the output (continuous) variables Y, aims to approximate a mapping function $f : X \to Y$.

Clustering Unsupervised task that aims to organize objects into groups.

1.3.3 Neural networks

A neuron (**perceptron**) computes a weighted sum of its inputs and passes the result to an activation function to produce the output.
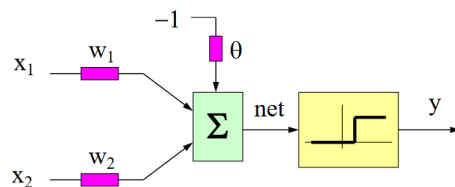
Figure 1.1: Representation of an artificial neuron

A **feed-forward neural network** is composed of multiple layers of neurons, each connected to the next one. The first layer is the input layer, while the last is the output layer. Intermediate layers are hidden layers.
The expressivity of a neural networks increases when more neurons are used:

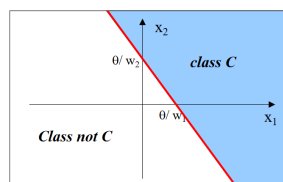**Single perceptron** Able to compute a linear separation.

Figure 1.2: Separation performed by one perceptron

Three-layer network Able to separate a convex region ($n_{\text{edges}} \leq n_{\text{hidden neurons}}$)
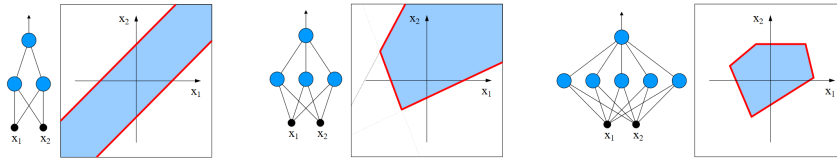

Figure 1.3: Separation performed by a three-layer network

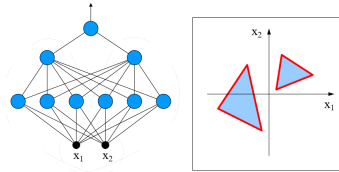**Four-layer network** Able to separate regions of arbitrary shape.

Figure 1.4: Separation performed by a four-layer network

Theorem 1.3.1 (Universal approximation theorem). A feed-forward network with one hidden layer and a finite number of neurons is able to approximate any continuous function with desired accuracy.

Deep learning Neural network with a large number of layers and neurons. The learning process is hierarchical: the network exploits simple features in the first layers and synthesis more complex concepts while advancing through the layers.

1.4 Automated planning

Given an initial state, a set of actions and a goal, **automated planning** aims to find a partially or totally ordered sequence of actions to achieve a goal.
An **automated planner** is an agent that operates in a given domain described by:

- Representation of the initial state

- Representation of a goal

- Formal description of the possible actions (preconditions and effects)

1.5 Swarm intelligence

Decentralized and self-organized systems that result in emergent behaviors.