

ML Introduction

The meaning of learning

In order to define what learning means as an ML concept we need three concepts:

- A certain task T that we want to improve on.
- Metrics to measure performance P .
- Experience E .

Learning means improving on a certain task T by gaining experience E in order to maximize performances P .

The thing we want to learn is a **target function**.

A target function can be defined in many ways and depends on our design choices. For example action functions outputs the next action that an agent should take, while a value function returns a value associated to a certain state.

The way we express a target function is assigning weights to meaningful components of our inputs. If we take for example the configuration of a board b , we can have a function that looks like this:

$$V(b) = w_0 + w_1 bp(b) + w_2 rp(b) + \dots$$

In this case $bp(b)$ is the number of black pieces on board b , $rp(b)$ is the number of red pieces on board b etc.

Our problem is about estimating the weights in order to obtain an optimal solution. In order to define what we mean by “finding an optimal solution” we have to define some concepts:

- $V(b)$ is the target function. Basically it represents perfection in our task (that we'll never reach).
- $V'(b)$ is our function that we will manipulate.
- $V_{train}(b)$ is the function given by the dataset.

We have to structure an algorithm such that given the estimate of the error:

$$error(b) = V'(b) - V(b)$$

will minimize it. This means that we want our function to get as close as possible to the ideal one by iterating the algorithm on the training dataset. The algorithm will work on our function's weights in order to achieve this.

The output of our algorithm will be the function $\hat{V}(b)$ that we will use in practice.

What is Machine Learning?

Machine learning is the set of all the various learning techniques. The ones we will study are:

- Supervised learning
 - Classification
 - Regression
- Unsupervised Learning
- Reinforcement Learning

In general ***a machine learning problem is the task of learning a function given a dataset.***

A function maps elements of a certain input domain to elements of a certain output domain:

$$f : X \rightarrow Y$$

A dataset D is a set of samplets that contains informations about the function.

Learning a function means to iteratively generate a function $\hat{f}(x)$ that best approximates the ideal function $f(x)$:

$$\hat{f}(x) \approx f(x) \text{ for } x \in X - X_d$$

Note that our goal during learning is achieving good prediction *outside* the dataset's domain, which means that we are able to generalize knowledge.

The input dataset will influence the kind of technique we will use. In fact we have:

- *Supervised learning* when the dataset contains information about inputs and their corresponding outputs: $D = \{(x_i, y_i)_{i=1}^n\}$
- *Unsupervised learning* when the dataset contains only the inputs: $D = \{(x_i)_{i=1}^n\}$

Moreover functions can be categorized based on the input and output domains.

Inputs can be expressed as a finite set of numbers $A_1 \times \dots \times A_m$ (*discrete*) or *continuous* if $X \in \mathbb{R}^m$.

When the output is finite we say that we are facing a problem of *classification*. In the continuous case we have a problem of *regression*.

We call *hypothesis space* H the space of all possible approximation of the optimal function c . The learning task consists in finding the best approximation $h^* \in H$ of $c : X \rightarrow Y$.

An hypothesis is called *consistent* only if:

$$h(x) = c(x) \quad \forall x \in D$$