

# Introduction to machine learning

Giacomo Fantoni

Telegram: @GiacomoFantoni

Github: <https://github.com/giacThePhantom/intro2ml>

6 aprile 2021

# Indice

<b>1</b>	<b>Introduzione</b>	<b>2</b>
1.1	Definizioni . . . . .	2
1.2	Processo . . . . .	2
1.2.1	Il processo di apprendimento . . . . .	2
1.3	Modello . . . . .	2
1.4	Deep learning . . . . .	3
<b>2</b>	<b>Machine learning basics</b>	<b>4</b>
2.1	Introduzione . . . . .	4
2.1.1	Processo di learning . . . . .	4
2.2	Dati . . . . .	4
2.2.1	Training e test set . . . . .	4
2.3	Tipi di learning . . . . .	5
2.3.1	Supervised learning . . . . .	5
2.3.2	Unsupervised learning . . . . .	5
2.3.3	Reinforcement learning . . . . .	5

# Capitolo 1

## Introduzione

### 1.1 Definizioni

Si intende per machine learning lo studio di algoritmi che migliorano autonomamente attraverso l'esperienza. È un campo dell'intelligenza artificiale. Stravolge il paradigma convenzionale della programmazione: un algoritmo di machine learning infatti prende come input un insieme di dati e risultati in modo da produrre un programma che fornisce un risultato appropriato. Coinvolge pertanto la scoperta automatica di regolarità nei dati attraverso algoritmi in modo da poter compiere azioni basate su di essi.

### 1.2 Processo

Il machine learning permette ai computer di acquisire conoscenza attraverso algoritmi che inferiscono e imparano da dati. Questa conoscenza viene rappresentata da un modello che può essere utilizzato su nuovi dati.

#### 1.2.1 Il processo di apprendimento

Il processo di apprendimento in particolare coinvolge diversi passaggi:

- Acquisizione dei dati dal mondo reale attraverso dispositivi di misurazione come sensori o database.
- Preprocessamento dei dati: filtraggio del rumore, estrazione delle feature e normalizzazione.
- Riduzione dimensionale: selezione e proiezione delle feature.
- Apprendimento del modello: classificazione, regressione, clustering e descrizione.
- Test del modello: cross-validation e bootstrap.
- Analisi dei risultati.

### 1.3 Modello

Un algoritmo di machine learning impara dall'esperienza  $E$  in rispetto di una classe di compiti  $T$  e di misurazione delle performance  $P$ , se la  $P$  di  $T$  aumenta con  $E$ . Si nota pertanto come un compito

di machine learning ben definito possiede una tripla:

$$\langle T, P, E \rangle$$

### 1.4 Deep learning

Il deep learning è un sottoinsieme del machine learning che permette a modelli computazionali composti di multipli strati di imparare la rappresentazione di dati con multipli livelli di astrazione. Si utilizza pertanto una rete neurale con diversi strati di nodi tra input e output. Questa serie di strati tra input e output computa caratteristiche rilevanti automaticamente in una serie di passaggi. Questi algoritmi sono resi possibili da:

- Enorme mole di dati disponibili.
- Aumento del potere computazionale.
- Aumento del numero di algoritmi di machine learning e della teoria sviluppata dai ricercatori.
- Aumento del supporto dall'industria.

## Capitolo 2

# Machine learning basics

### 2.1 Introduzione

Il machine learning permette ai computer di acquisire conoscenza attraverso algoritmi che imparano e inferiscono dai dati. Tale conoscenza viene rappresentata da un modello che viene poi utilizzato su dati futuri.

#### 2.1.1 Processo di learning

Si individua un processo di learning:

- Acquisizione di dati dal mondo reale attraverso sensori.
- Preprocessamento dei dati: eliminazione del rumore, estrazione delle features e normalizzazione.
- Riduzione di dimensionalità attraverso selezione e proiezione di features.
- Learning del modello: classification, regression, clustering e description.
- Test del modello attraverso cross-validation e bootstrap.
- Analisi dei risultati.

### 2.2 Dati

I dati disponibili ad un algoritmo di machine learning sono tipicamente un insieme di esempi. Questi esempi sono tipicamente rappresentati come un array di features, caratteristiche dei dati di interesse per lo studio in atto.

#### 2.2.1 Training e test set

In particolare per questi algoritmi si assume sempre che il training e il test set siano distribuiti secondo variabili indipendenti e identicamente distribuite (*i.i.d*)

## 2.3 Tipi di learning

### 2.3.1 Supervised learning

Nel supervised learning vengono dati in input a un modello o predittore un insieme di esempi che possiedono una label. Il modello poi impara a creare delle predizioni su un nuovo esempio.

#### 2.3.1.1 Classificazione

In un problema di classificazione si trova un insieme finito di label discrete. In particolare dato un training set  $\mathcal{T} = \{(x_1, u_1), \dots, (x_m, y_m)\}$ , si deve imparare una funzione  $f$  per predire  $y$  dato  $x$ .  $f$  sarà pertanto:

$$f : \mathbb{R}^d$$

Dove  $d$  è la dimensionalità di  $x$  e  $k$  il numero di labels distinte.

#### 2.3.1.2 Regression

Un problema di regressione presenta un insieme di label continue. Dato un training set  $\mathcal{T} = \{(x_1, y_1), \dots, (x_m, y_m)\}$ , si deve imparare una funzione  $f$  per predire  $y$  dato  $x$ .  $f$  sarà pertanto:

$$f : \mathbb{R}^d \rightarrow \mathbb{R}$$

Dove  $d$  è la dimensionalità di  $x$ .

#### 2.3.1.3 Ranking

Il ranking è un tipo particolare di classificazione in cui una label è un ranking.

### 2.3.2 Unsupervised learning

Nel unsupervised learning vengono dati in input a un modello o predittore un insieme di esempi senza label. Il modello impara a creare delle predizioni su un nuovo esempio.

#### 2.3.2.1 Clustering

Nel clustering, data  $\mathcal{T} = \{x_1, \dots, x_m\}$  si deve trovare la struttura nascosta che intercorre tra le  $x$  o i clusters.

#### 2.3.2.2 Dimensionality reduction

Nella dimensionality reduction si tenta di ridurre il numero di variabili sotto considerazione ottenendo un insieme di variabili principali.

### 2.3.3 Reinforcement learning

Nel reinforcement learning un agente impara dall'ambiente interagendo con esso e ricevendo premi per lo svolgimento di azioni particolari. In particolare, data una sequenza di esempi o stati e una reward dopo il completamento di tale sequenza si impara a predire l'azione da svolgere per uno stato o esempio individuale.