

Machine Learning

Pierciro Caliandro

10 marzo 2022

Indice

1	Introduzione del corso	2
1.1	Introduzione	2
1.2	Presentazione del ML	2

Capitolo 1

Introduzione del corso

1.1 Introduzione

Per il progetto: sviluppare un notebook in cui la parte espositiva sia presente, quindi intercalare markdown + codice, dove il markdown deve esserci e deve essere eloquente.

NON solo codice, esporre i ragionamenti, le limitazioni etc...

1.2 Presentazione del ML

L'informatica è risolvere problemi da un punto di vista algoritmico, risolvere un problema attraverso algoritmi si può dividere in fasi:

- analizzare il problema
- vederlo in modo matematico
- pensare ad un algoritmo
- implementarlo in un linguaggio
- verificare la correttezza (attività di testing) e valutarne l'efficienza

Il caso elementare può essere un testo dove si richiede di trovare i caratteri *i* nel testo: l'algoritmo può scorrere un carattere alla volta e contare ogni *i*. L'algoritmo funziona sulla base di un assunto, ovvero che sappiamo riconoscere la *i*, quindi che l'algoritmo sappia cosa è una *i* o cosa non lo sia, quindi abbiamo bisogno di una definizione del carattere che permetta alle persone di riconoscere il carattere.

Possiamo dare varie definizioni:

- riconoscere la *i* come na' sbarretta co' un puntino sopra
- sequenza ASCII, UNICODE, ovvero una sequenza di bit

A volte non è facile: il problema può essere simile ma in un contesto diverso, ovvero ad esempio se occorre riconoscere quante volte in una sequenza di foto compaia una certa persona (NDR: Emma Stone, bello sticchio).

Il metodo che si applica può essere triviale e non può essere codificato, come ad esempio riconoscere una faccia o capire un testo parlato.

Non abbiamo quindi una chiara definizione di cosa cerchiamo di identificare, quindi come facciamo a definirla: una caratterizzazione di una persona per attributi è un po' complicata e non esaustiva. Se

non possiamo fare così, possiamo farlo per esempi: forniamo una serie di foto di quella persona, non siamo in grado di definirla formalmente, possiamo poi dare degli esempi di foto che non ritraggono quella persona. Ci si muove quindi in modo induttivo, caratterizzando sulla base di singoli specifici esempi, in ML si cerca di fare questo: problemi per cui non si riesce a fornire una soluzione in quanto troppo complessi si prova a risolverli a partire da esempi.

Entra in gioco l'induzione, quindi non so riconoscere la persona ma posso mostrare solo degli esempi e quindi potrò avere un algoritmo derivato dall'insieme di esempi, tale per cui sottoponendo una foto questo ci dice se è della persona o meno con una certa precisione.

Quindi, nel ML, **a partire dai dati vogliamo derivare un algoritmo che ci dia delle perdizioni**, per poter derivare un algoritmo serve un altro algoritmo.

Abbiamo quindi:

- i dati
- l'algoritmo da derivare
- l'altro algoritmo, che è un **sistema o modello di ML**

Vorremmo quindi un algoritmo che **apprenda** come fare una cosa, che è un modo accorciato perché in realtà l'algoritmo sulla base dei dati, ne produce un altro per rispondere ad una domanda ben precisa.

Un esempio classico è considerare il problema in contesti più complessi come ad esempio riconoscere un carattere scritto a mano: è l'"Hello World" del ML, dove abbiamo dei caratteri scritti a mano (delle immagini), da cui voglio ottenere un algoritmo che sia in grado di riconoscere la cifra corrispondente, sulla base di esempi che sono la coppia $\langle \text{immagine}, \text{cifra corrispondente} \rangle$, come nel caso precedente della classificazione di persone.

È in questo caso una classificazione **multiclasse**, ma come prima abbiamo un elemento per cui abbiamo la risposta corretta, da cui vogliamo realizzare un sistema per cui la risposta corretta non c'è ma deve essere probabilmente corretta.

Normalmente a grandi linee, un sistema di ML cosa fa:

- costruisce un modello, l'informazione va trattata da algoritmi e quindi va formalizzata matematicamente.
Si costruisce ad esempio un range 0-9, abbiamo per esempio immagini 28x28 in scale di grigi, abbiamo quindi un vettore di $28 \times 28 = 784$ valori, ognuno è un intero compreso fra 0, 255. Un elemento è quindi un punto in uno spazio di 784 dimensioni
- geometricamente, l'insieme di tutti i caratteri è un insieme di punti in uno spazio a 784 dimensioni, vorremmo ottenere un modo per tagliare lo spazio in modo che da una parte ci siano solo 0, da una solo 1 e così via... generando così delle zone associate ad un singolo valore così che qualunque punto che cade in una zona è associato ad un certo valore
- in altro modo, vogliamo una funzione f da $\{0,1,\dots,255\}^{784}$ in $\{0,1,2,\dots,9\}$