

Tarallucci, Vino e Machine Learning  
Introduzione agli speech di L. Massaron

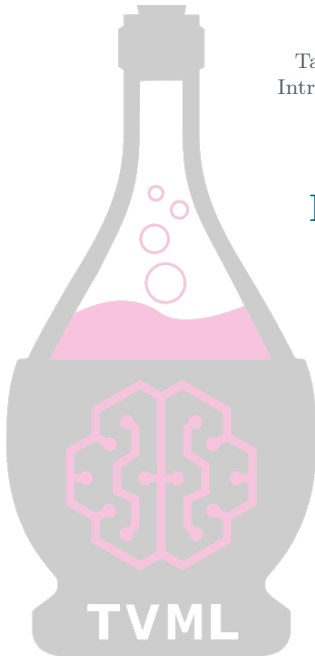
# Machine Learning e Probabilità

Fabio Mardero

[fabio.mardero@gmail.com](mailto:fabio.mardero@gmail.com)

[github.com/fmardero](https://github.com/fmardero)

21 marzo 2019





## Il Machine Learning

Apprendimento

Compito

## Accenni di Inferenza Statistica

Approccio frequentista

Approccio bayesiano



# Il Machine Learning



Il machine learning è un gruppo di modelli matematici in grado di “apprendere” dai dati allo scopo di eseguire, nel modo migliore possibile, un dato compito.

### Caratterizzazione di un modello di ML

Un modello di machine learning è quindi caratterizzato da

- ▶ un apprendimento
- ▶ un compito



In termini matematici, un modello apprende quando modifica la sua struttura, o i suoi parametri, per ridurre gli errori delle sue previsioni.

Può essere paragonato ad un agente collocato in un dato ambiente (*environment*). È esattamente ciò che accade per un algoritmo di ML messo in produzione.

**L'algoritmo può interagire con l'ambiente o “subirlo”.**



L'ambiente e/o l'interazione con esso produce un fenomeno i cui effetti misurabili sono raccolti come dati.

I dati possono essere

- ▶ strutturati, organizzati in database detti *dataset*,
- ▶ non strutturati, conservati senza alcuno schema,
- ▶ semi-strutturati.

Ogni informazione con carattere diverso è detta variabile.



## Metodi di apprendimento

L'apprendimento di un modello di machine learning può avvenire in tre diversi modi:

- ▶ per rinforzo (*reinforcement learning*),
- ▶ in maniera supervisionata (*supervised learning*),
- ▶ in maniera non supervisionata (*unsupervised learning*).



### Italiano

L'agente interagisce con l'environment e ogni sua azione modifica l'ambiente stesso.

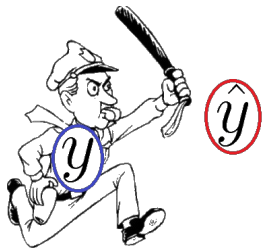
### Matematiche

Il modello interagisce con il sistema e ogni sua previsione modifica lo stato dello stesso.





Nel tempo, non necessariamente ad ogni interazione con l'ambiente, l'agente riceve un *feedback* sul suo comportamento. Egli modifica quindi le sue future azioni, sulla base delle precedenti, tentando di massimizzare quelle che hanno portato a risultati positivi e minimizzando quelle risultate negative. L'apprendimento dipende quindi da un sistema di *rewards* e *punishments*.



Il modello subisce l'ambiente.  
Nel caso dell'apprendimento supervisionato il modello mira a predire il comportamento di una o più variabili osservate rispetto alle altre.

Indicata con  $\hat{y}$  la previsione e con  $y$  il valore osservato, il modello apprende a minimizzare l'errore tra  $\hat{y}$  e  $y$ . L'apprendimento è, informalmente, “supervisionato” dai valori di  $y$ .



Il modello subisce l'ambiente ma non è allenato per fornire una previsione.

L'apprendimento non supervisionato prevede che l'algoritmo ricerchi strutture informative (*pattern*) tra i dati.



Il compito definisce su cosa il modello è allenato e con quali intenzioni. L'oggetto di analisi sono dati strutturati o semi-strutturati.

Si riconoscono due casi:

1. si individuano delle variabili più importanti, dette **variabili target**/risposta, rispetto alle altre, chiamate **variabili esplicative**/covariate/features,
2. tutte le variabili sono intese come significative (o potenzialmente tali).

Dato un insieme di dati, spetta all'osservatore decidere come intende interpretarli e se assegnare particolare importanza a qualcuna delle variabili disponibili.



## Caso 1

Compiti di regressione o classificazione.

Mirando a fornire una previsione accurata delle variabili target, il modello spiega il fenomeno che genera  $y$ .

## Caso 2

Compiti legati all'estrazione di informazione dai dati e ad una loro rappresentazione, ad esempio il clustering.

Ad esempio si individuano somiglianze tra informazioni presenti nel dataset.

Il modello di machine learning, a discapito del compito, fornisce un'interpretazione del fenomeno che genera i dati. Cambia la finalità esplicativa.



## Accenni di Inferenza Statistica



L'inferenza statistica mira a descrivere processi/esperimenti non deterministici tramite l'osservazione di un campione.

L'inferenza, anche storicamente, può essere descritta secondo l'approccio

- ▶ frequentista
- ▶ bayesiano



Data una variabile aleatoria  $X$ , sia  $H = \{X = x\}$  evento di interesse/di “successo”.

La probabilità che si verifichi  $H$  è pari a

$$\lim_{n \rightarrow +\infty} \frac{n_H}{n} = p(H)$$

con  $n$  in numero di prove/esperimenti e  $n_H$  il numero dei casi di successo.

*La probabilità che si verifichi un dato evento è interpretata come la frequenza di successo.*





Ogni variabile aleatoria può essere

- ▶ discreta con funzione/legge di probabilità

$$P(\{X = x\}) = p_x$$

- ▶ continua con distribuzione di probabilità

$$P(\{a < X < b\}) = \int_a^b f(x)dx$$

con  $f$  funzione di densità di probabilità tale che

$$f(a) = \lim_{h \rightarrow 0} \frac{P(X > a) - P(X > a + h)}{h}$$



**Distribuzione condizionata di  $A$  dato  $B$**

$$P(\{X \in A\} | \{X \in B\}) = P(A|B) = \frac{P(A \cap B)}{P(B)}$$



## Distribuzione congiunta di $A$ e $B$

$$P(\{X \in A\} \cap \{X \in B\}) = P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

I due eventi si dicono indipendenti se

$$P(\{X \in A\} | \{X \in B\}) = P(A|B) = P(A)$$

$$P(\{X \in B\} | \{X \in A\}) = P(B|A) = P(B)$$

allora

$$P(\{X \in A\} \cap \{X \in B\}) = P(A \cap B) = P(A) P(B)$$

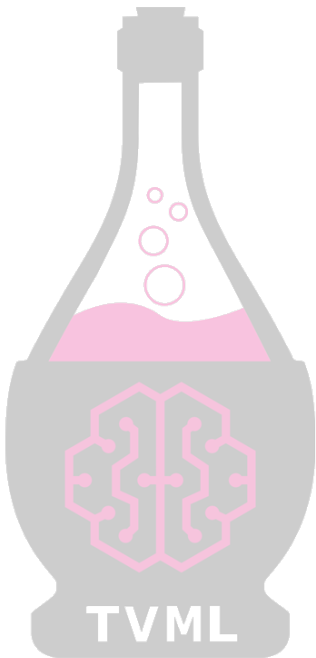


Data una variabile aleatoria  $X$ , sia  $H = \{X = x\}$  evento di successo. L'**osservatore**, prima degli esperimenti, possiede un grado di fiducia in merito al verificarsi di  $H$  che è descritto da una *probabilità a priori*  $P(H)$ . Ad ogni prova  $E$ , osservando i risultati, l'osservatore aggiorna la sua opinione in merito alla probabilità di verificarsi dell'evento. Si parla dunque di *probabilità a posteriori*  $P(H|E)$  che è appunto “condizionata” dalle evidenze pratiche.

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

Con  $P(E|H)$  è indicata la probabilità che si verifichi l'evento osservato secondo le ipotesi a priori dell'osservatore.

*La probabilità che si verifichi un dato evento è interpretata come il livello di fiducia che l'osservatore ne associa a riguardo.*



Grazie dell'attenzione!