

# Computazione per l'interazione naturale: macchine che apprendono



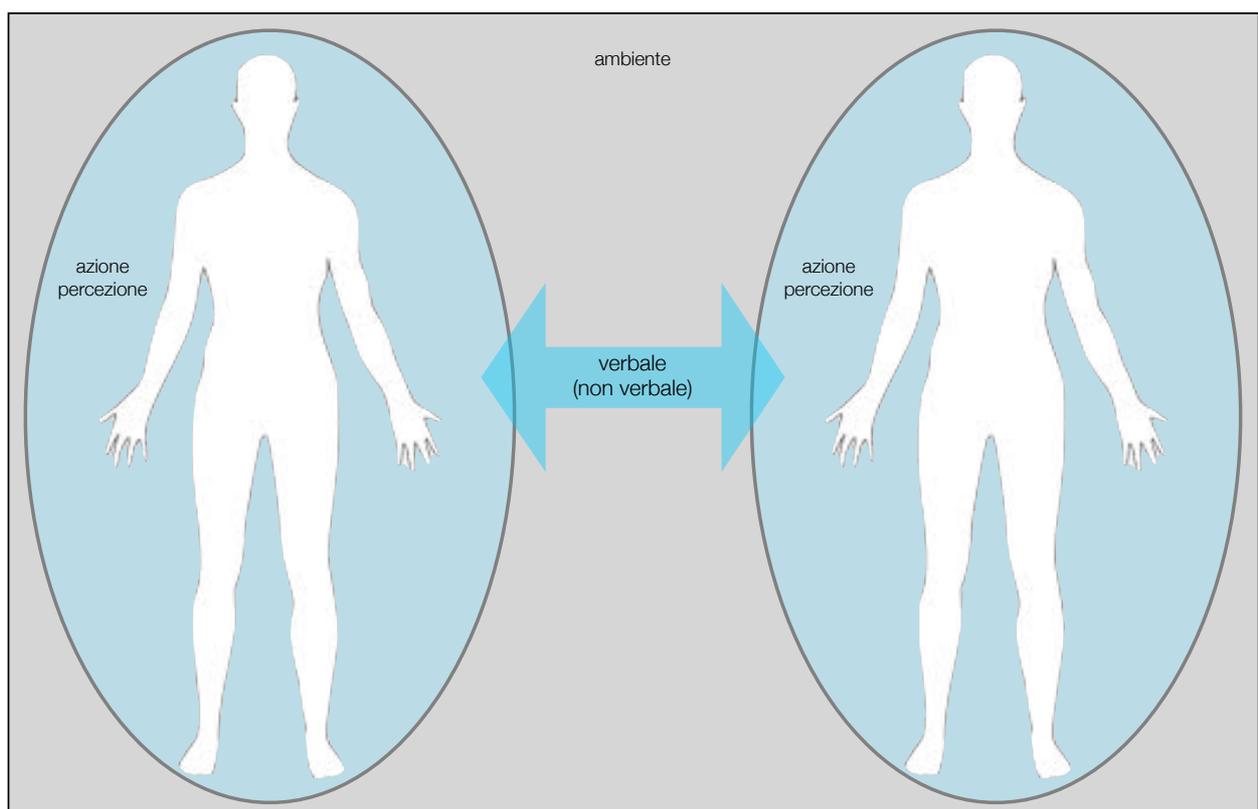
Corso di Interazione uomo-macchina II

Prof. Giuseppe Boccignone

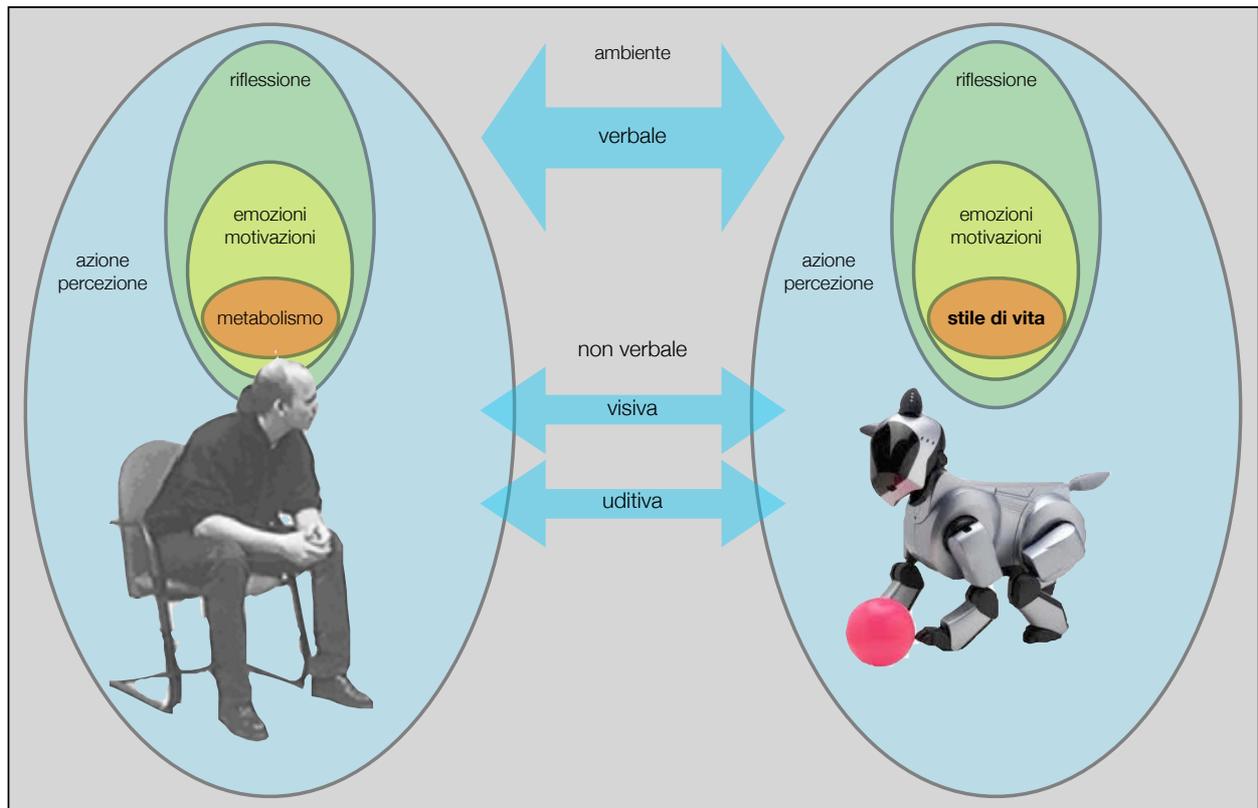
Dipartimento di Informatica  
Università di Milano

[boccignone@di.unimi.it](mailto:boccignone@di.unimi.it)  
[http://boccignone.di.unimi.it/IUM2\\_2014.html](http://boccignone.di.unimi.it/IUM2_2014.html)

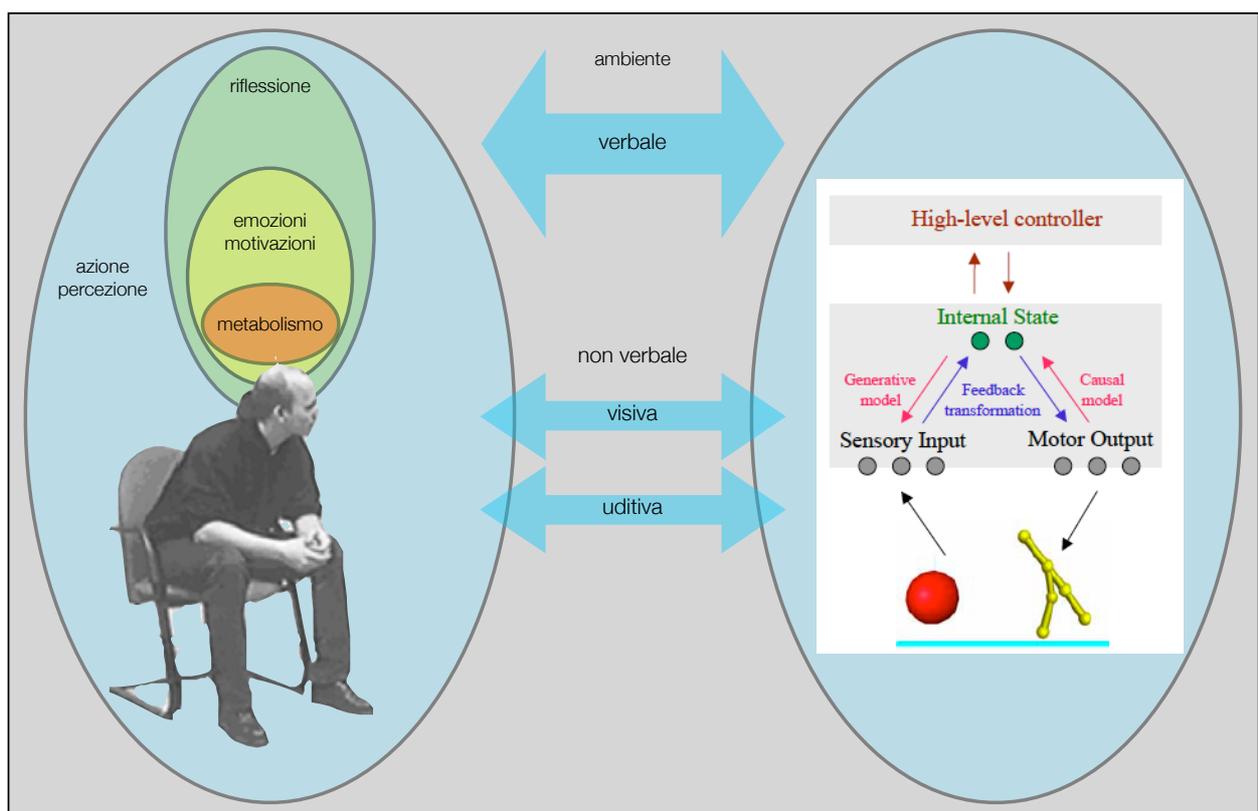
Ipotesi di lavoro: interazione fra organismi



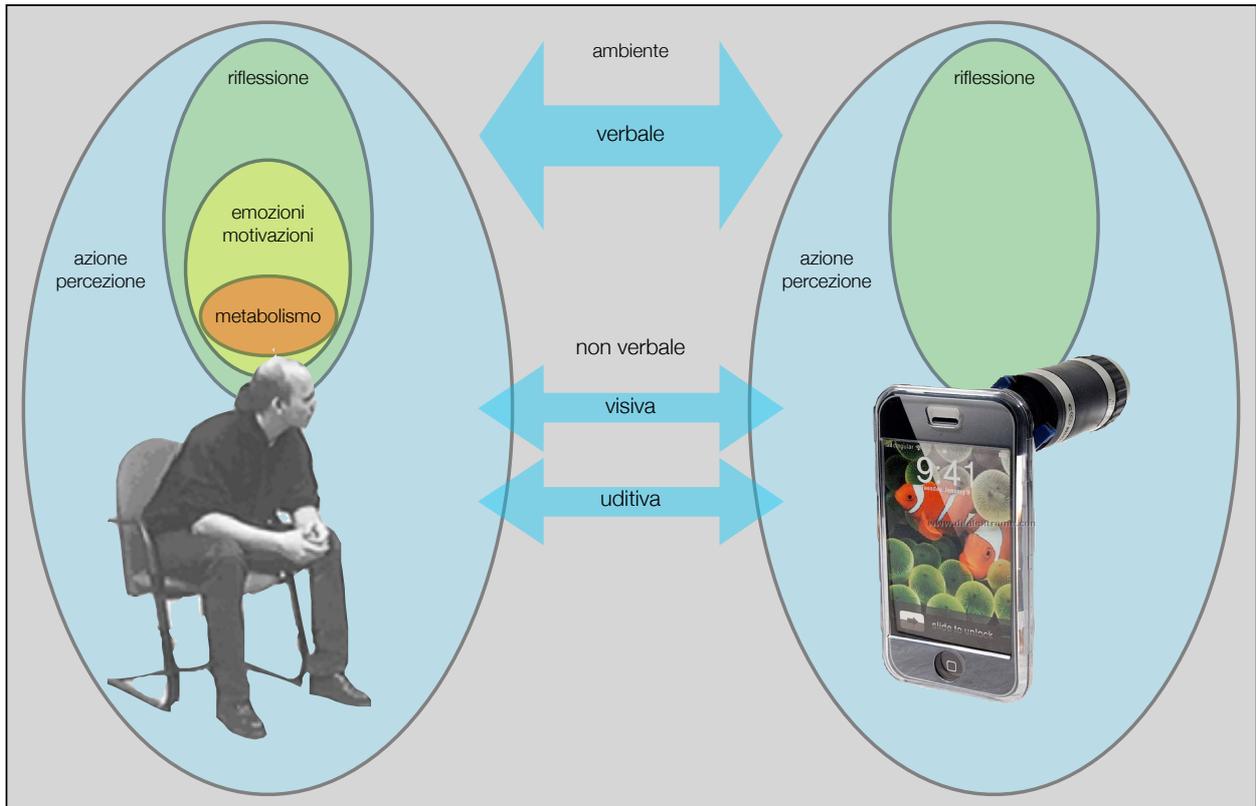
# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



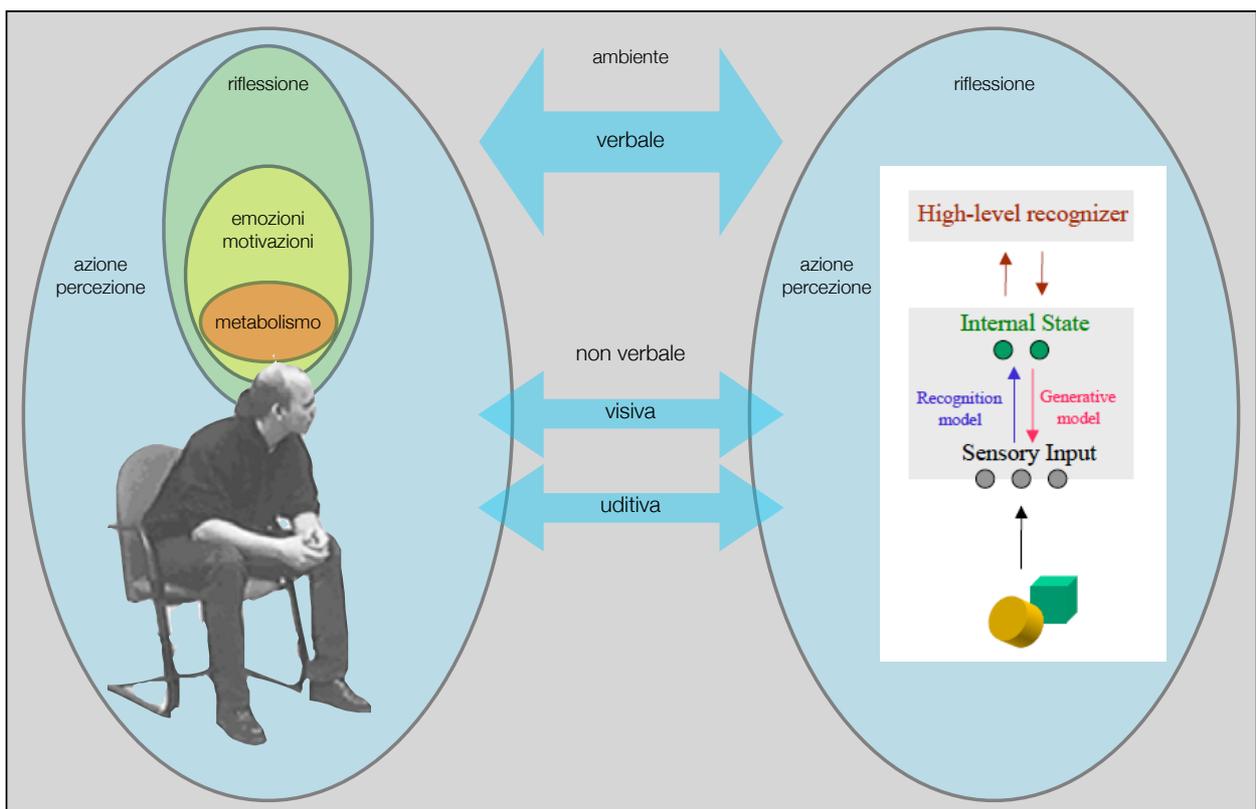
# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



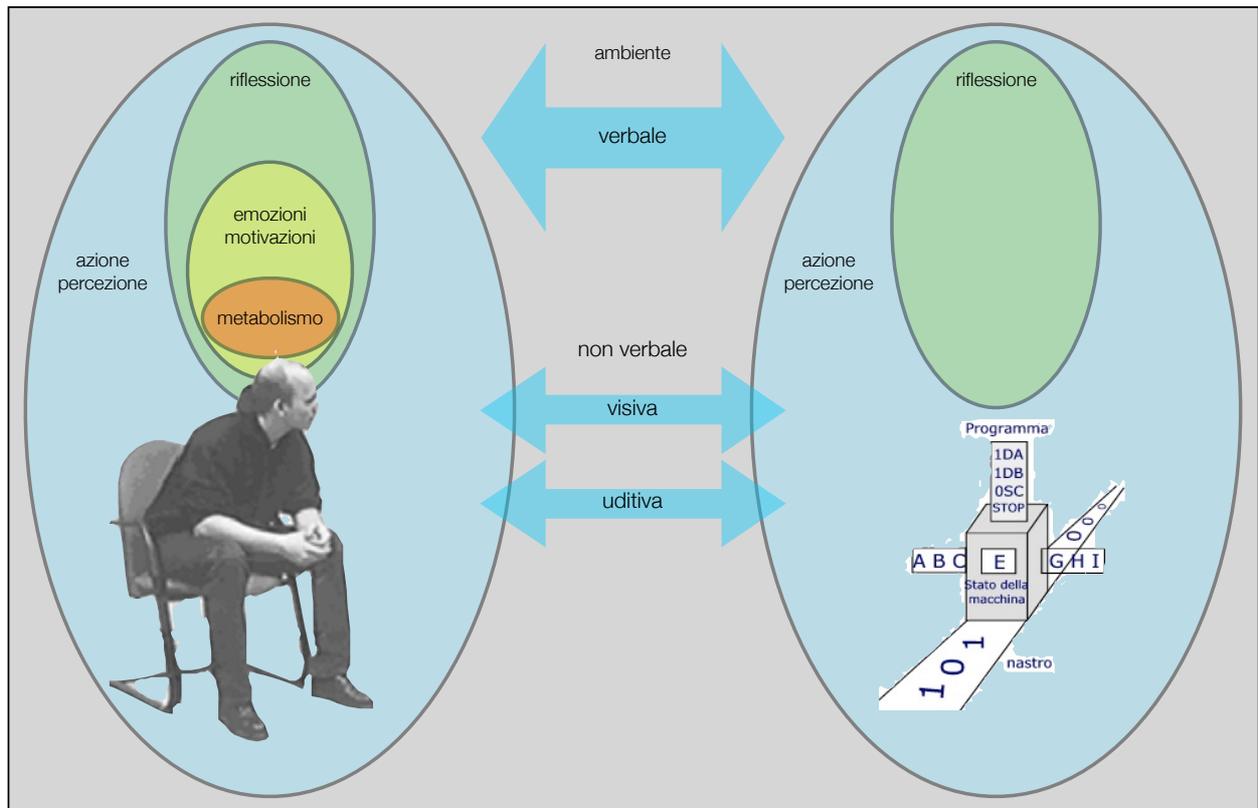
# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



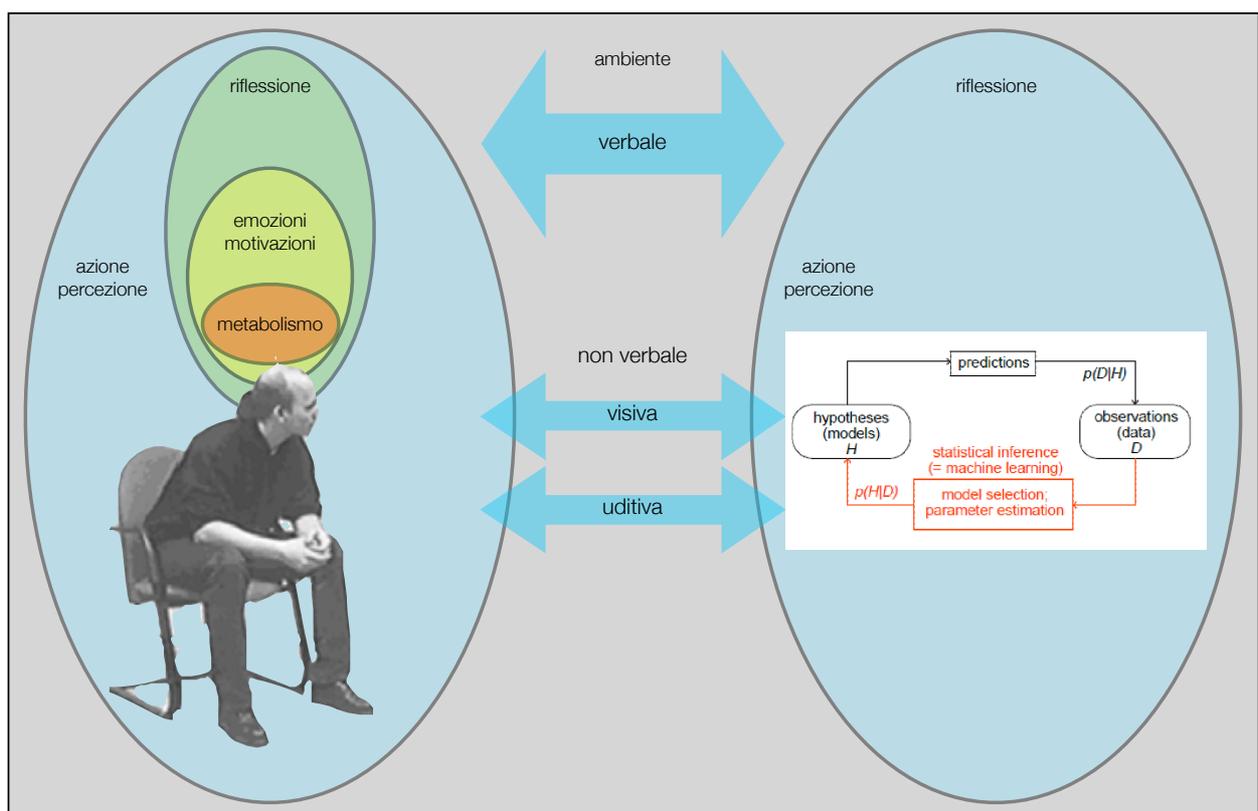
# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



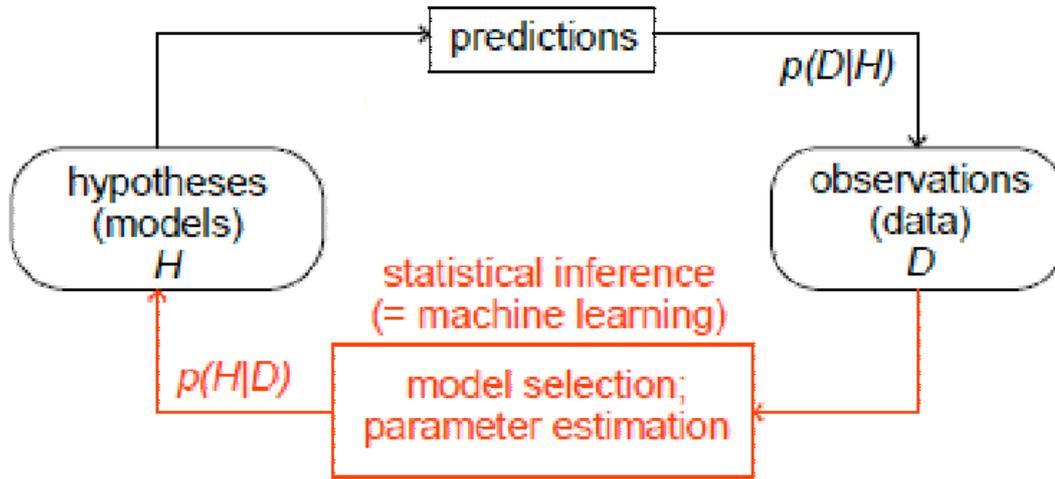
# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



# Ipotesi di lavoro: interazione fra organismi //variazioni di contesto



# Modelli di computazione per l'interazione naturale



## Esempi

Riconoscere  
espressioni facciali

$$E \begin{array}{c} \xleftarrow{p(E | D)} \\ \xrightarrow{p(D | E)} \end{array}$$

Riconoscere gesti

$$G \begin{array}{c} \xleftarrow{p(G | D)} \\ \xrightarrow{p(D | G)} \end{array}$$

Riconoscere azioni

$$A \begin{array}{c} \xleftarrow{p(A | D)} \\ \xrightarrow{p(D | A)} \end{array}$$

$D$  = dati percepibili



## Il problema dell' apprendimento statistico

---

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |

## Il problema dell' apprendimento statistico

---

- Per learning si intende il processo di individuare, a partire da un insieme di dati, un modello che li descriva.
  - Per esempio, nel caso del supervised learning si vuole individuare una corrispondenza tra input e output.
- Un modo per fare ciò è postulare l'esistenza di un qualche tipo di meccanismo parametrico per la generazione dei dati, di cui non conosciamo però i valori esatti dei parametri.
- Questo processo fa tipicamente riferimento a tecniche di tipo statistico.

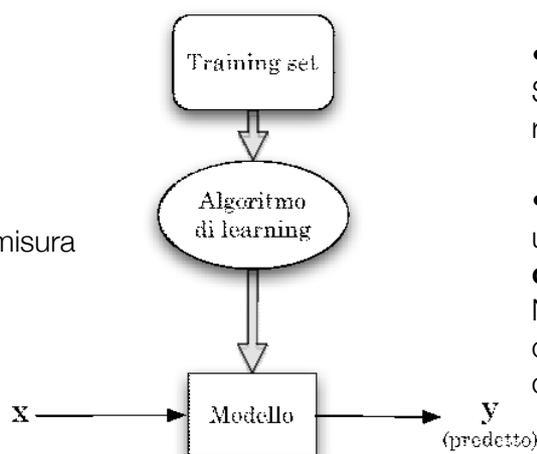
# Classificazione dei problemi di learning

## //apprendimento supervisionato

- I problemi di learning vengono divisi in due tipologie principali: supervised learning e unsupervised learning:
- Supervised: prevedere, dato un elemento di cui si conoscono un insieme di parametri (features), il valore di un diverso parametro di output relativo all'elemento stesso. Viene definito (mediante apprendimento da insiemi di esempi) un modello

Valori assunti dalle singole features  $\mathbf{x}$ :

- Quantitativi: forniscono la misura di una grandezza
- Qualitativi: specificano una classe di appartenenza



Valori di output  $\mathbf{y}$

- Quantitativi: **regressione**. Se  $y$  è un vettore, si parla di regressione multivariata.
- Qualitativi: l'assegnazione ad una classe (categoria), e si parla di **classificazione**. Numero di possibili classi = 2, si classifica binaria, altrimenti di classificazione multi-classe.

## Apprendimento supervisionato

### //Classificazione

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |

# Apprendimento supervisionato

## //Classificazione

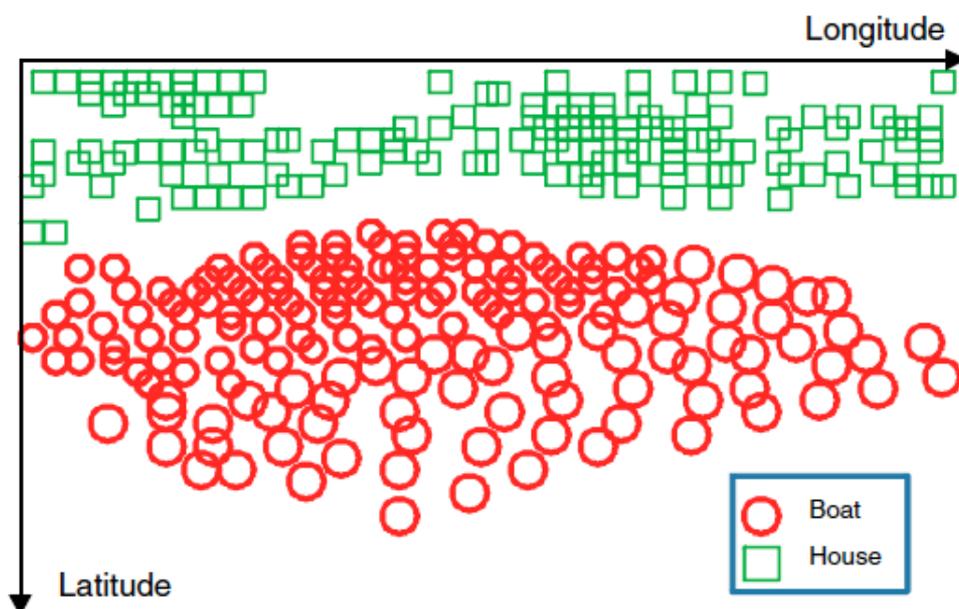
---



# Apprendimento supervisionato

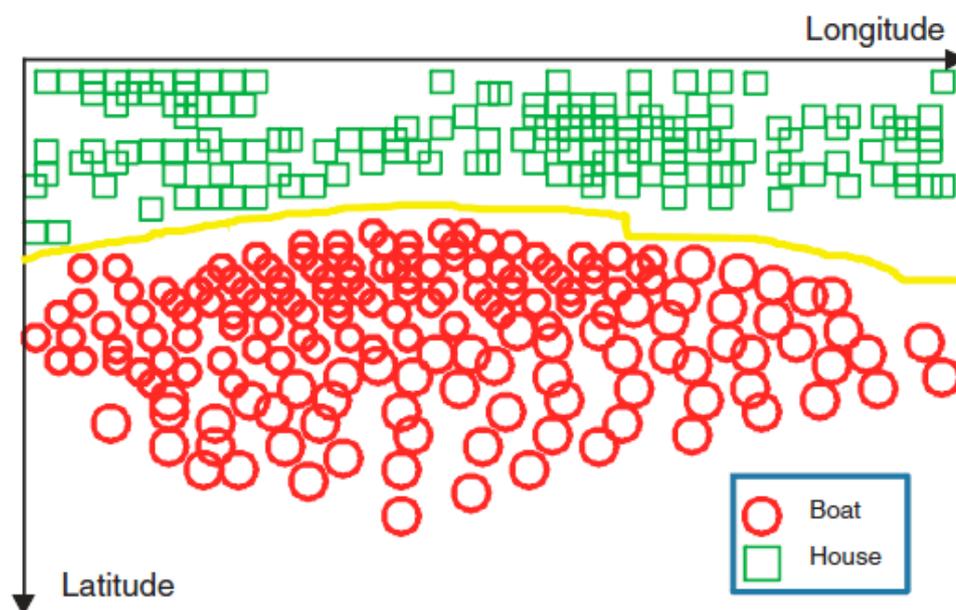
## //Classificazione

---



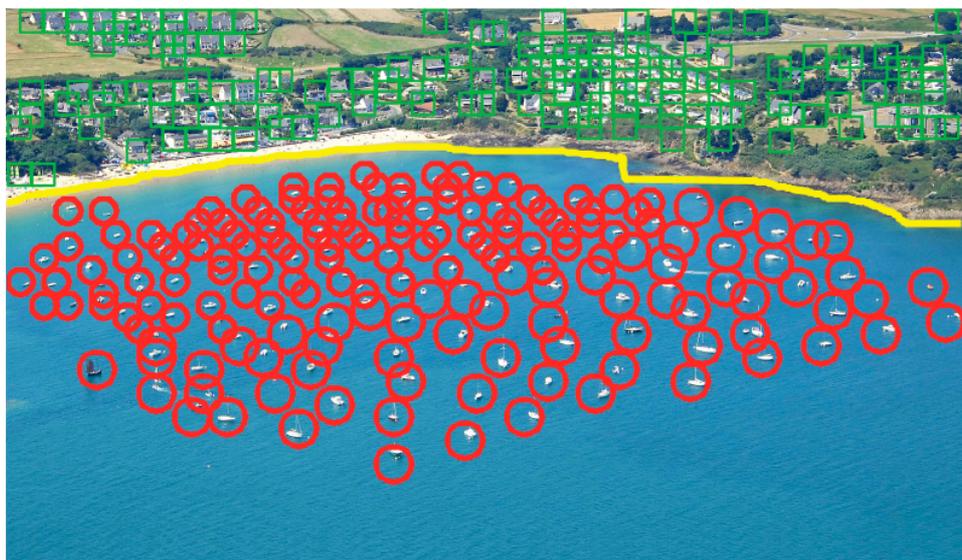
# Apprendimento supervisionato //Classificazione

---



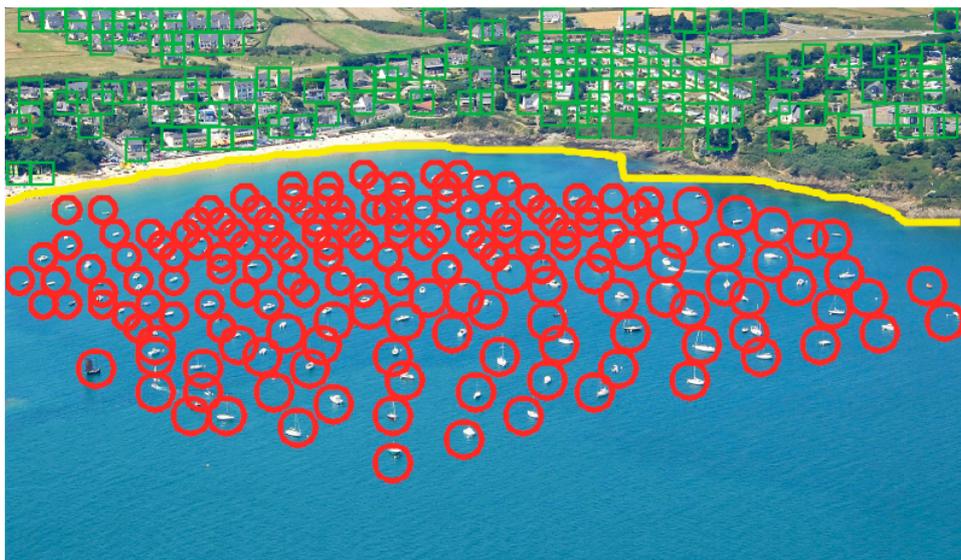
# Apprendimento supervisionato //Classificazione

---



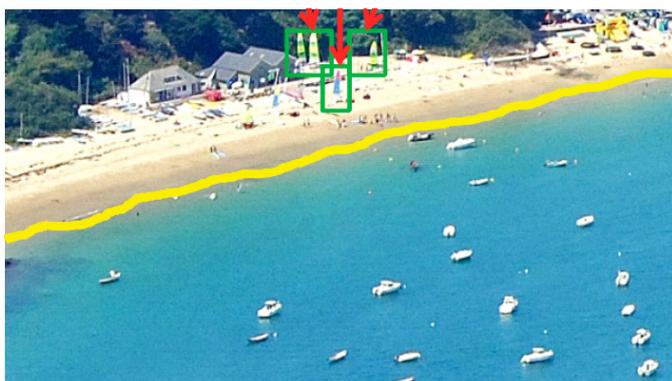
# Apprendimento supervisionato //Classificazione

---



# Apprendimento supervisionato //Classificazione

---



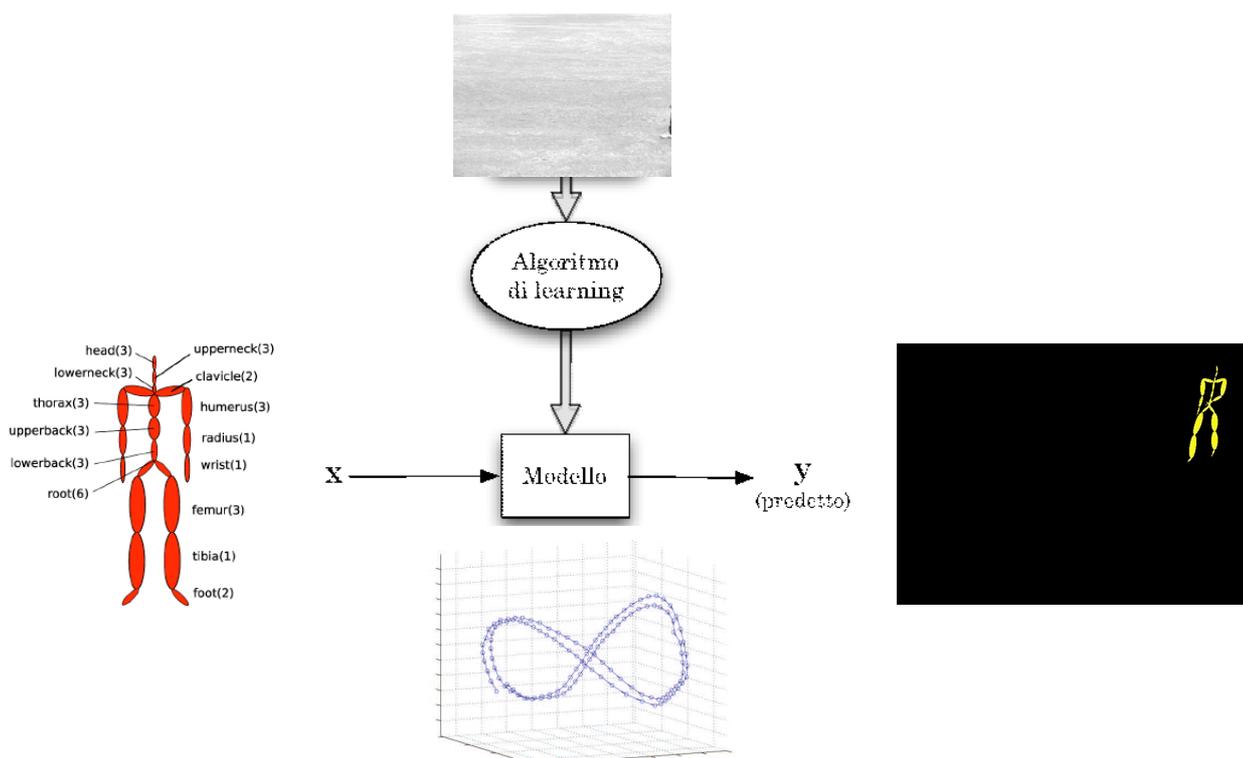
# Apprendimento supervisionato

## //Regressione

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |

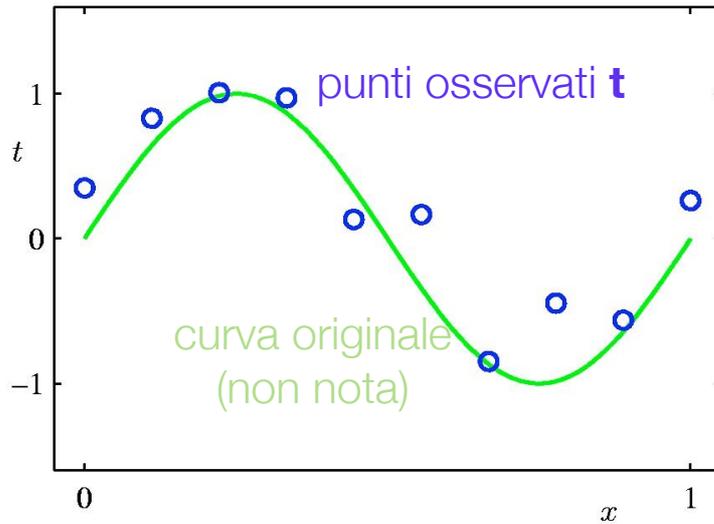
# Apprendimento supervisionato

## //Regressione



# Apprendimento supervisionato

## //Regressione: fitting di una curva



Trovare i coefficienti  $\mathbf{w}$

$$\mathbf{t} = \mathbf{y}(x; \mathbf{w}) + \mathbf{u}$$

↑  
rumore

Polinomio  
approssimante

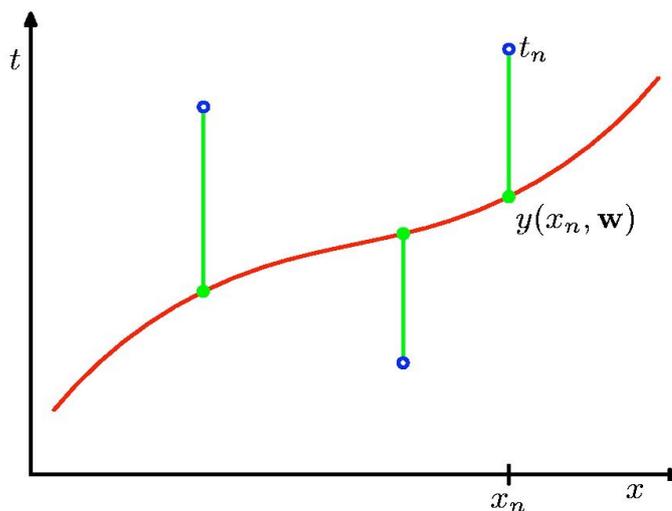
$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j$$

Problema di  
apprendimento

Trovare i coefficienti  $\mathbf{w}$  (parametri) noti i punti target  $\mathbf{t}$

# Apprendimento supervisionato

## //Regressione: fitting di una curva



$$\mathbf{t} = \mathbf{y}(x; \mathbf{w}) + \mathbf{u}$$

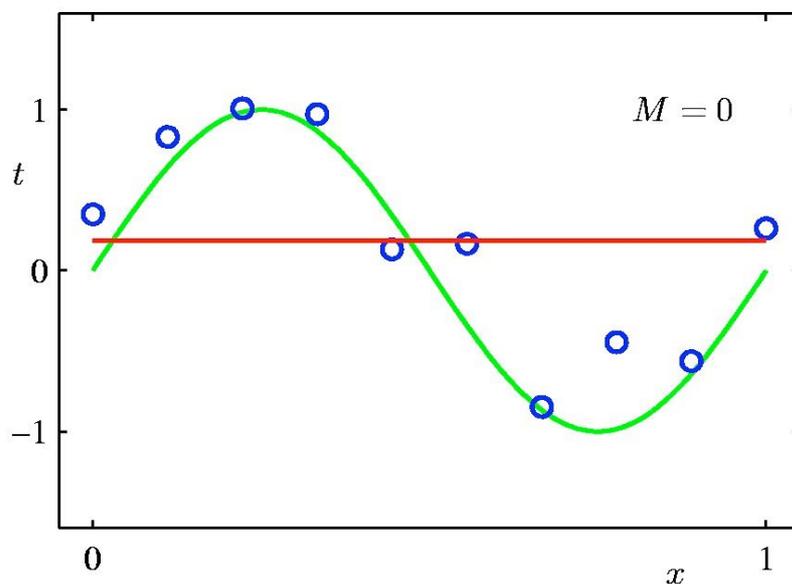
Funzione di  
loss (errore)  
quadratica

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

# Apprendimento supervisionato

## //Regressione: fitting di una curva

---

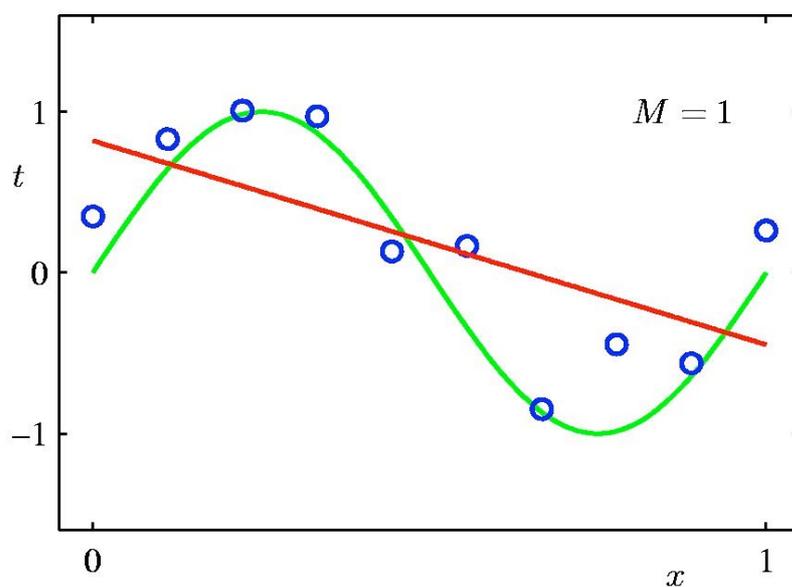


Polinomio di ordine 0

# Apprendimento supervisionato

## //Regressione: fitting di una curva

---

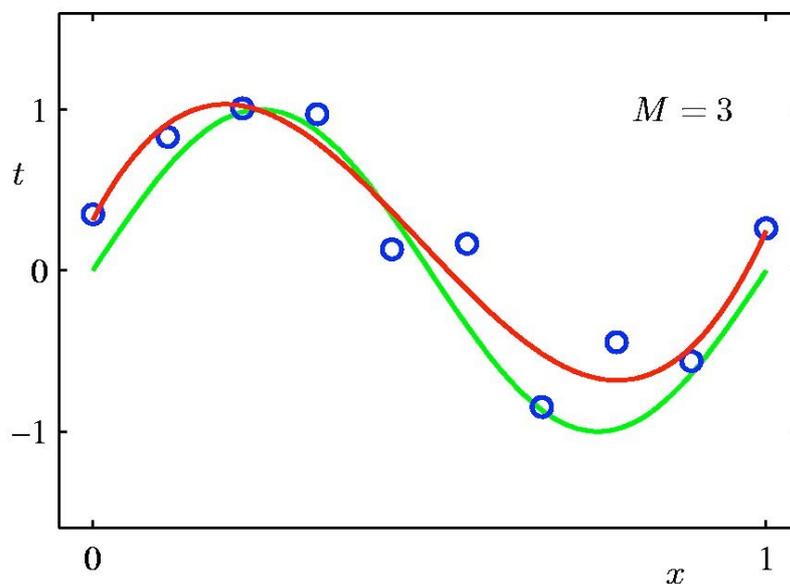


Polinomio di ordine 1

# Apprendimento supervisionato

//Regressione: fitting di una curva

---

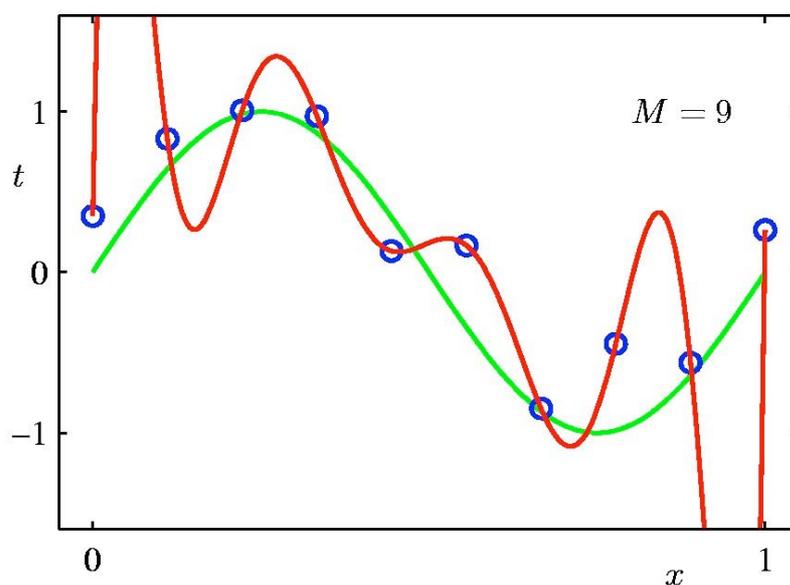


Polinomio di ordine 3

# Apprendimento supervisionato

//Regressione: fitting di una curva

---

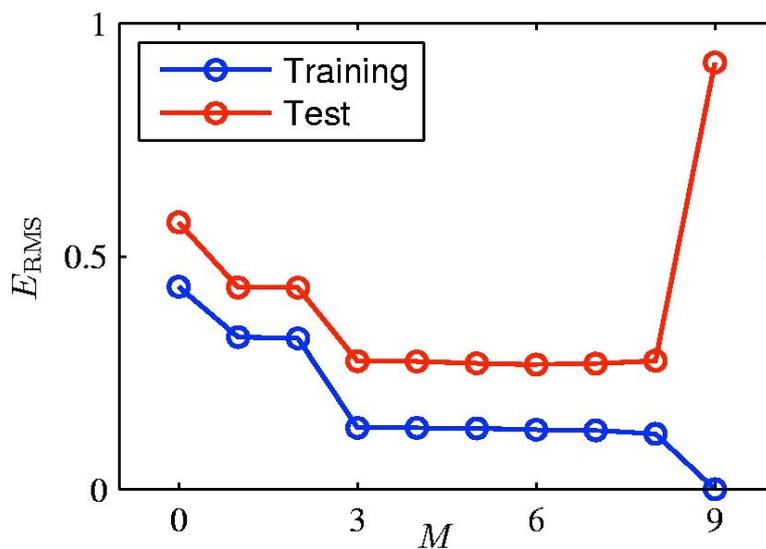


Polinomio di ordine 9

# Apprendimento supervisionato

## //Regressione: il problema dell'over-fitting

---



Root-Mean-Square (RMS) Error:  $E_{\text{RMS}} = \sqrt{2E(\mathbf{w}^*)/N}$

## Apprendimento non supervisionato

---

- I dati di training non sono etichettati (unlabeled)
  - Le labels sono variabili nascoste (hidden variables)
- Clustering
- Riduzione di dimensionalità

# Apprendimento supervisionato

## //Clustering

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |

Apprendimento non supervisionato  
\\Il problema delle variabili latenti:clustering

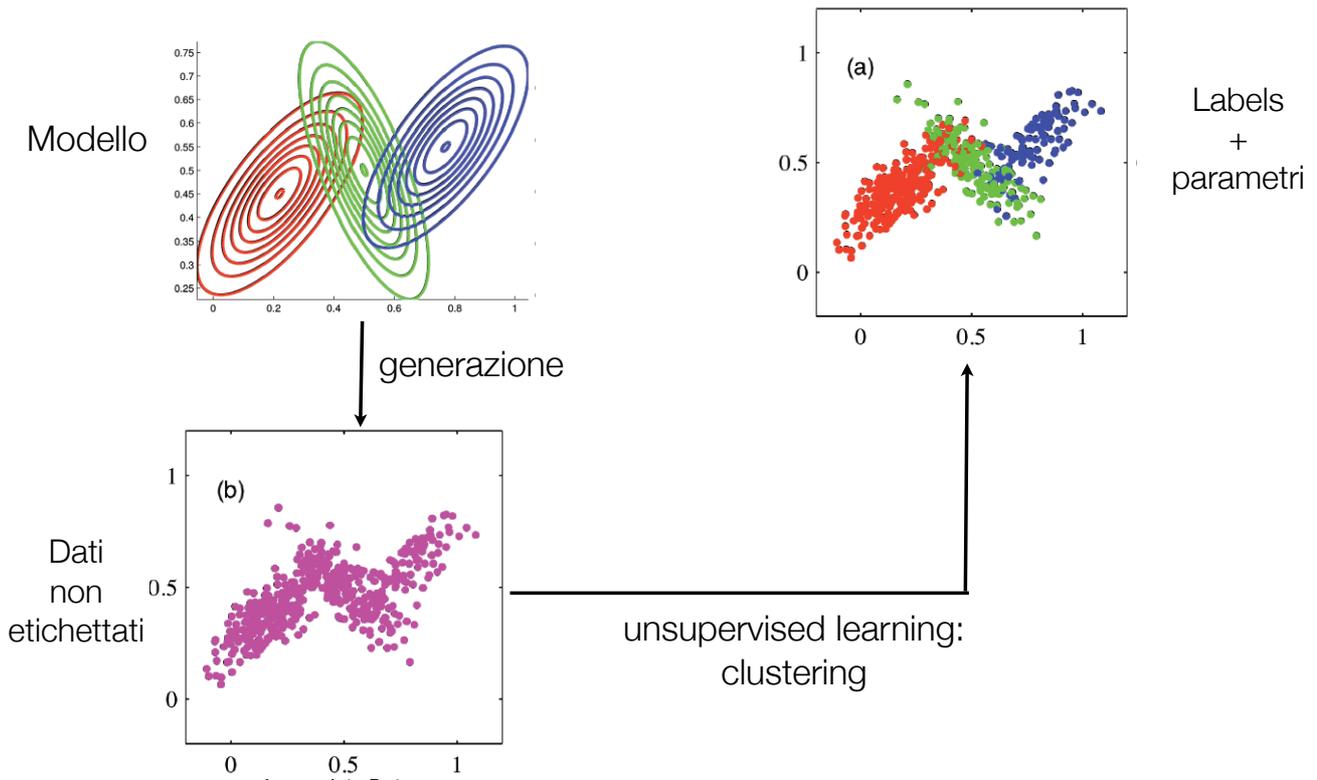
- Parte dei dati sono sconosciuti
  - intrinsecamente inaccessibili.:
  - dati mancanti o errati.:
  - Esempio: rumore.
- Esempio: per un'immagine da segmentare l'insieme completo di dati è:

misura/osservazione su ciascun pixel +

labels che dicono a quale componente il pixel appartiene

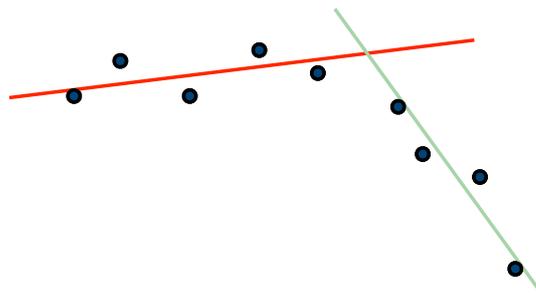
- Se i dati mancanti sono correlati con quelli osservati, possiamo sperare di recuperarli lavorando sui dati osservati.
- Se sono indipendenti, non c'è speranza

Apprendimento non supervisionato  
\\Il problema delle variabili latenti: clustering

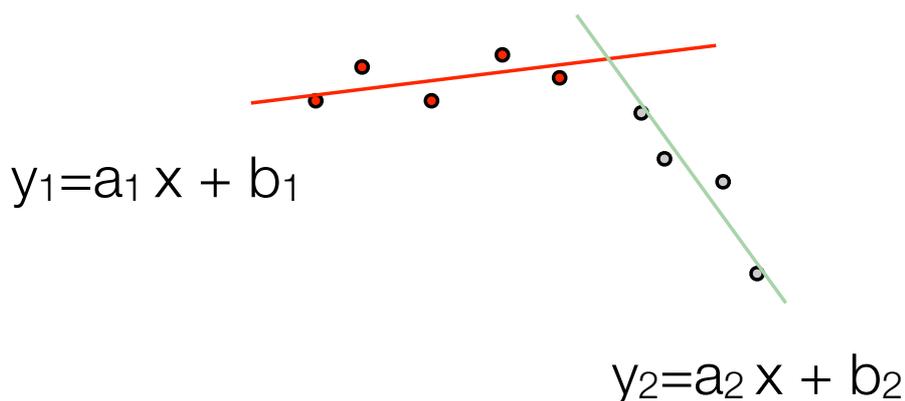


Apprendimento non supervisionato  
\\Il problema delle variabili latenti: clustering

- Supponiamo di voler raggruppare dei punti in due segmenti di linea “best-fit”

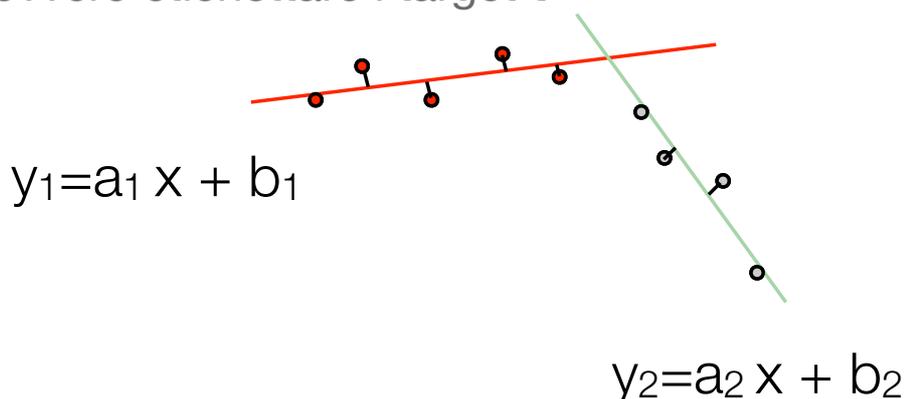


- E' un problema uovo/gallina: se conoscessimo a quale linea ciascun punto appartiene (cioè conosciamo i target  $\mathbf{t}$ ) potremmo calcolare il best fit (un semplice problema di regressione), ovvero i parametri  $\mathbf{w} = \{\mathbf{a}, \mathbf{b}\}$



35

- Se sapessi dove sono collocate le due linee best-fit, ovvero conosco i parametri  $\mathbf{w} = \{\mathbf{a}, \mathbf{b}\}$ , potrei classificare:
  - assegnare ciascun punto alla linea (minima distanza), ovvero etichettare i target  $\mathbf{t}$



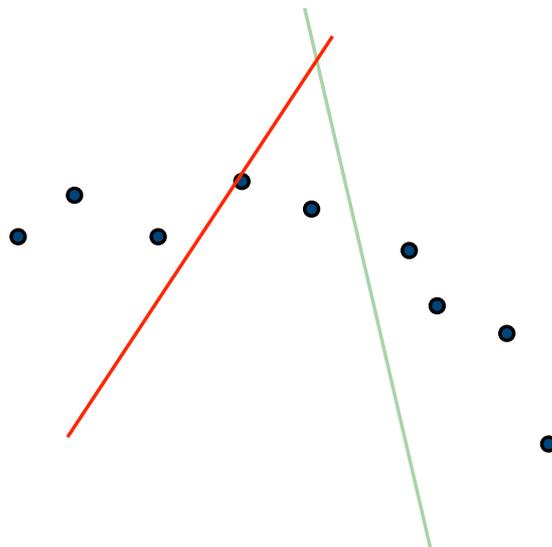
36

## Soluzione iterativa

- Inizializzazione: effettuare una scelta random delle linee
- Repeat
  - Trovare la linea più vicina a ciascun punto e raggruppare in due insiemi (**Classificazione**)
  - Ricalcolare la linea best-fit per ciascun insieme (**Fitting dei parametri**)
- Until (il processo converge)

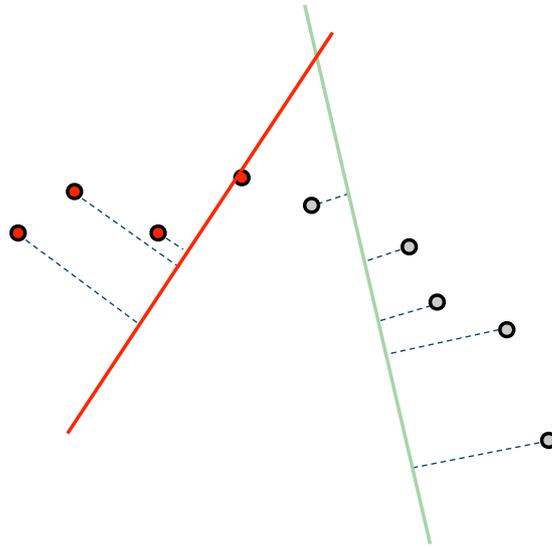
L'algoritmo è garantito essere **convergente**

37

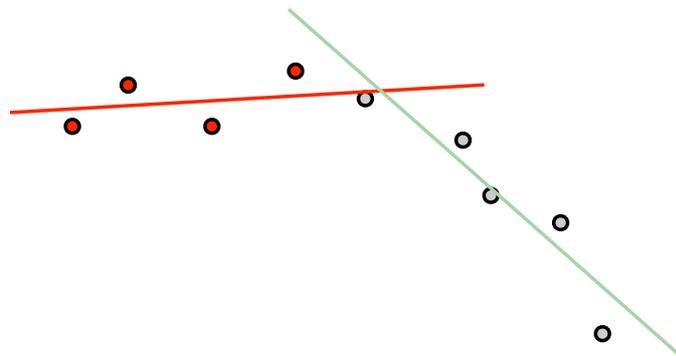


38

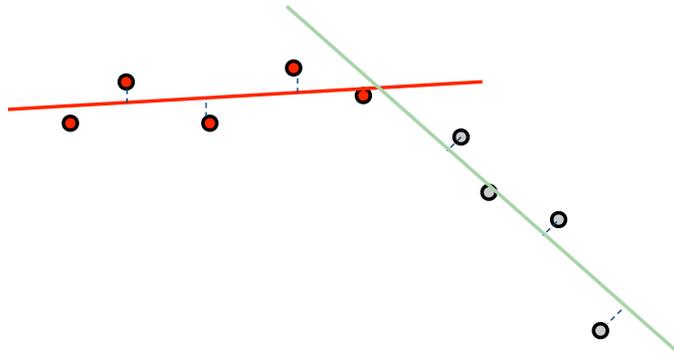
Passo E:  
Classifico i  
punti



Passo M:  
Fitting delle  
rette

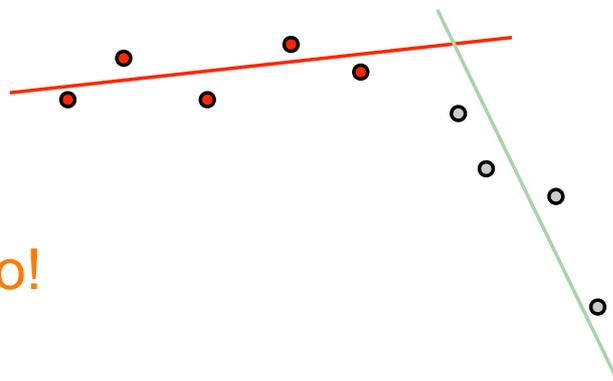


Passo E:  
Classifico i  
punti



41

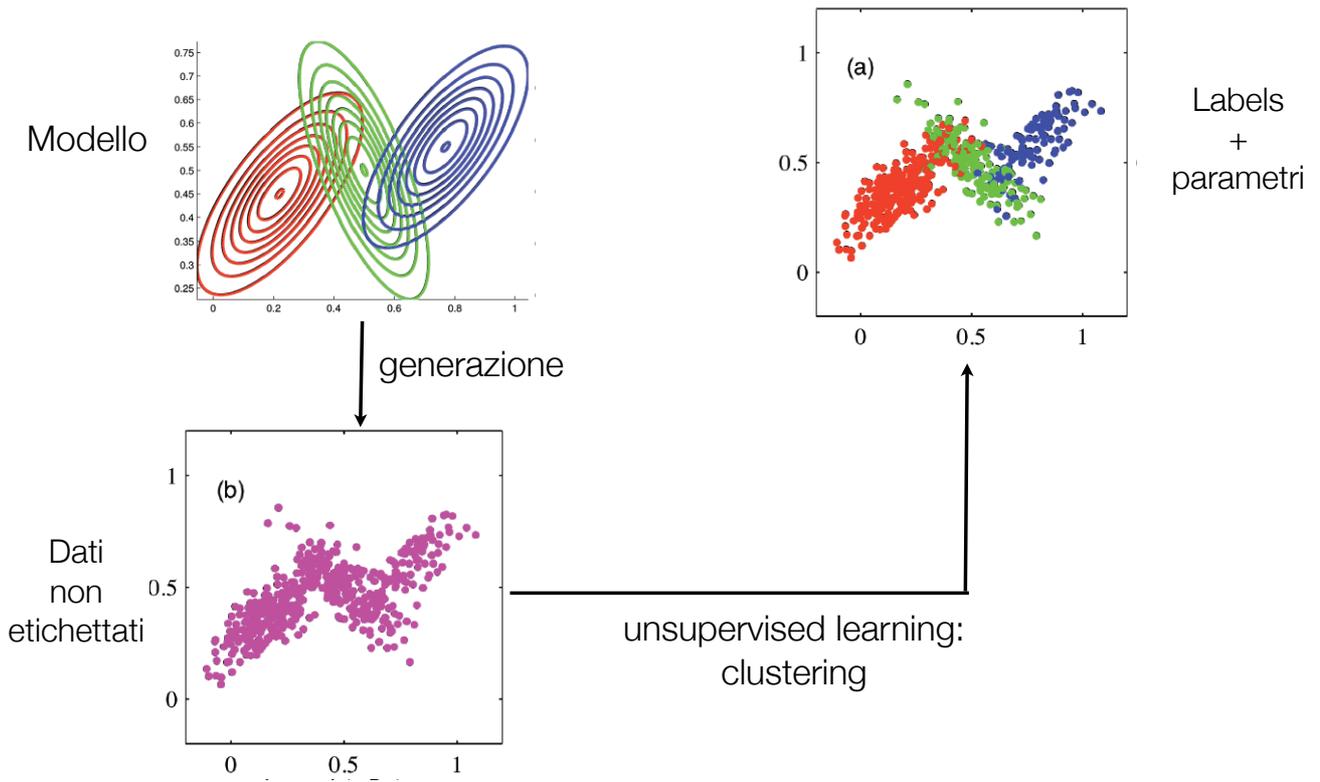
Passo M:  
Fitting delle  
rette



**Terminato!**

42

Apprendimento non supervisionato  
 \Il problema delle variabili latenti: clustering



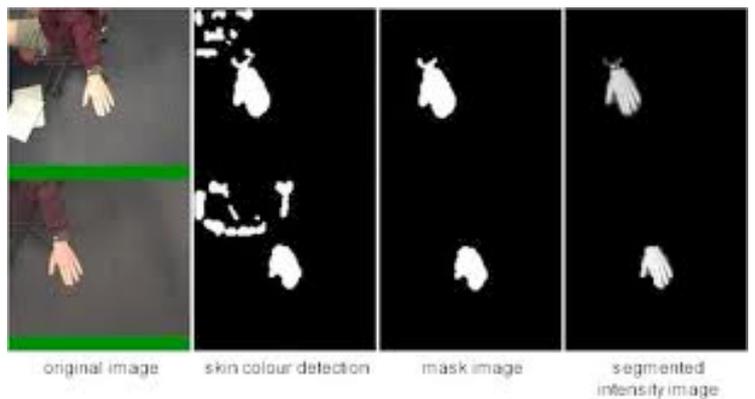
Apprendimento non supervisionato  
 \Esempio: rilevare la pelle



volti



gesti



# Apprendimento supervisionato

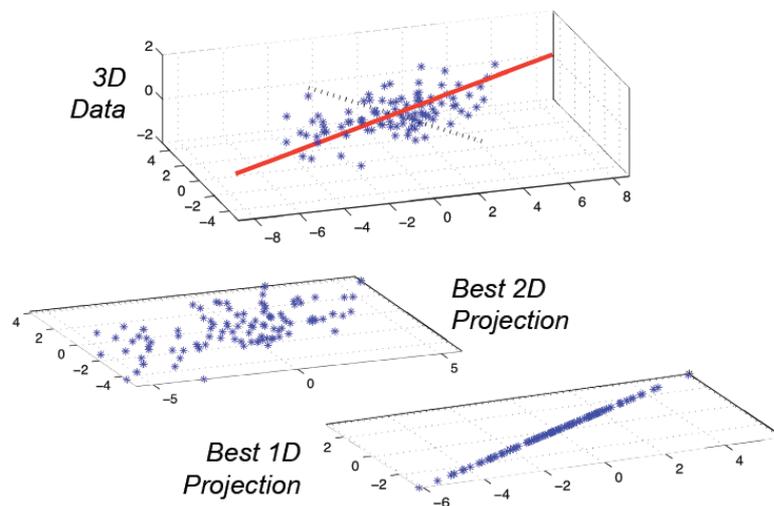
## //Riduzione di dimensionalità

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |

## Apprendimento non supervisionato

### \\Riduzione di dimensionalità

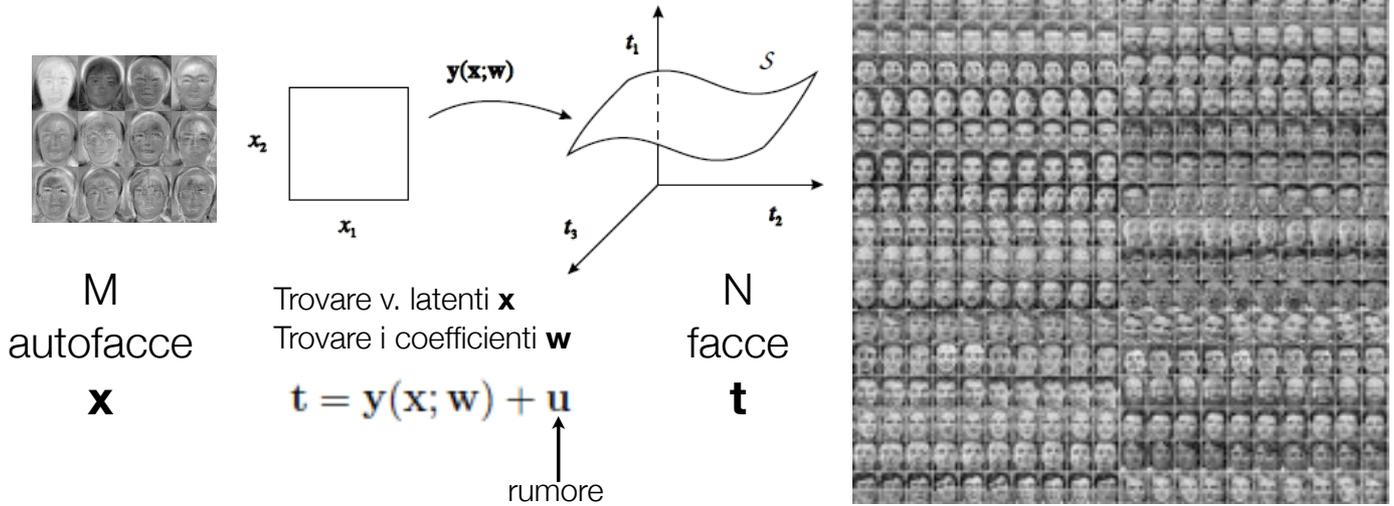
- Trovare delle labels / output per “comprimere” le variabili osservate
- Si passa da uno spazio di dimensionalità  $N$  a uno di dimensionalità ridotta  $M$ 
  - $M \ll N$



# Apprendimento non supervisionato

## \\Riduzione di dimensionalità

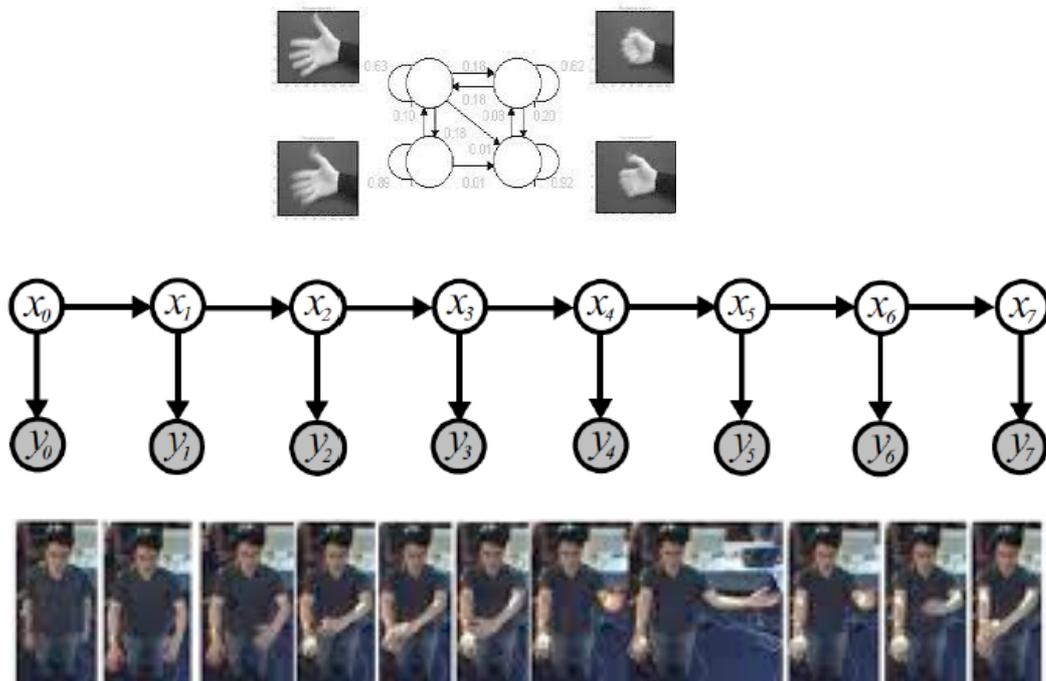
- Trovare delle labels / output per “comprimere” le variabili osservate
- Si passa da uno spazio di dimensionalità N a uno di dimensionalità ridotta M



# Apprendimento non supervisionato

## \\Modelli dinamici: Hidden Markov Models

- Modelli temporali a variabili latenti



# Il problema dell' apprendimento statistico

---

|                   | <i>Supervised Learning</i>  | <i>Unsupervised Learning</i> |
|-------------------|---|------------------------------|
| <i>Discrete</i>   | classification or categorization<br>Discrete Output $y \in \{1, \dots, K\}$ | clustering                   |
| <i>Continuous</i> | regression<br>Continuous Output $y \in R$                                   | dimensionality reduction     |