

L'APPRENDIMENTO AUTOMATICO

MACHINE LEARNING

Alessandro Sperduti
Fabio Aiolli

<http://www.math.unipd.it/~sperduti/ml.html>

Orario

Mercoledì e Venerdì, dalle 15:30 alle 17:30
Aula IBC45, Torre Archimede

Lezioni in laboratorio PO36 (15:30 - 17:30):
12 e 19 Dicembre; 9 e 16 Gennaio

Calendario completo
delle lezioni sul sito
web del corso di
laurea

Esame: scritto (compitini/compiti), con eventuale orale

Compitini

- 21 Novembre: I compitino
- 23 Gennaio: II compitino

Introduzione



Principali paradigmi di apprendimento



Ingredienti fondamentali



Alcuni Modelli di Apprendimento

Quando è necessario l'Apprendimento Automatico ?

quando il sistema deve

- **adattarsi** all'ambiente in cui opera
(anche personalizzazione automatica)
- **migliorare** le sue prestazioni rispetto ad un particolare compito
- **scoprire** regolarità e nuova informazione (conoscenza) a partire da dati empirici
- **acquisire** nuove capacità computazionali

Perché non usare un approccio algoritmico tradizionale ?

- impossibile formalizzare esattamente il problema (e quindi dare una soluzione algoritmica)
- presenza di rumore e/o incertezza
- complessità alta nel formulare una soluzione: impossibile "fare a mano"
- mancanza di conoscenza "compilata" rispetto al problema da risolvere

Rule 1: If Income (w_9) is Very High (VH) then the sensitivity to price (quality) parameter s_{11} (s_q) is Very Insensitive (VI) (Very Sensitive (VS)).

Che cifre sono rappresentate in questa immagine?

7	2	1	0	4	1	4	9	5	9
0	6	9	0	1	5	9	7	3	4
9	6	6	5	4	0	7	4	0	1
3	1	3	4	7	2	7	1	2	1
1	7	4	2	3	5	1	2	4	4

Patient103 time=1



Patient103 time=2



Patient103 time=n

Age: 23

FirstPregnancy: no

Anemia: no

Diabetes: no

PreviousPrematureBirth: no

Ultrasound: ?

Elective C-Section: ?

Emergency C-Section: ?

...

Age: 23

FirstPregnancy: no

Anemia: no

Diabetes: YES

PreviousPrematureBirth: no

Ultrasound: abnormal

Elective C-Section: no

Emergency C-Section: ?

...

Age: 23

FirstPregnancy: no

Anemia: no

Diabetes: no

PreviousPrematureBirth: no

Ultrasound: ?

Elective C-Section: no

Emergency C-Section: **Yes**

...

Rule 1:	<i>If</i> Income (w_j) is <i>Very High</i> (VH) <i>then</i> the sensitivity to price (quality) parameter s_{pj} (s_{qj}) is <i>Very Insensitive</i> (VI) (<i>Very Sensitive</i> (VS)).
Rule 2:	<i>If</i> Income (w_j) is <i>High</i> (HG) <i>then</i> the sensitivity to price (quality) parameter s_{pj} (s_{qj}) is <i>Insensitive</i> (IS) (<i>Sensitive</i> (ST)).
Rule 3:	<i>If</i> Income (w_j) is <i>Medium</i> (MD) <i>then</i> the sensitivity to price (quality) parameter s_{pj} (s_{qj}) is <i>Medium Sensitive</i> (MS).
Rule 4:	<i>If</i> Income (w_j) is <i>Low</i> (LW) <i>then</i> the sensitivity to price (quality) parameter s_{pj} (s_{qj}) is <i>Sensitive</i> (<i>Insensitive</i> (IS)).
Rule 5:	<i>If</i> Income (w_j) is <i>Very Low</i> (VL) <i>then</i> the sensitivity to price (quality) parameter s_{pj} (s_{qj}) is <i>Very Sensitive</i> (<i>Very Insensitive</i> (VS)).

Ruolo dei dati

Tipicamente

- si hanno a disposizione (molti ?) dati
 - ottenuti una volta per per tutte **batch learning**
 - acquisibili interagendo direttamente con l'ambiente **on-line learning**
- (forse) conoscenza del dominio applicativo, ma
 - incompleta
 - imprecisa (rumore, ambiguità, incertezza, errori, ...)

Desiderio: usare i dati per

- ottenere **nuova conoscenza**
- **raffinare** la conoscenza di cui si dispone
- **correggere** la conoscenza di cui si dispone

Principali Paradigmi di Apprendimento

Un paradigma specifica

- quali dati sono disponibili per l'apprendimento
- con che modalità il sistema riceve i dati
- che tipo di output ci si aspetta dal sistema

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Supervised Learning (apprendimento supervisionato)

$\{(x_i, f(x_i))\}$

- dato un insieme di dati preclassificati (esempi di apprendimento) apprendere una descrizione generale che incapsula l'informazione contenuta negli esempi
- tale descrizione deve poter essere usata in modo predittivo
dato un nuovo ingresso $\tilde{x}$, predire $f(\tilde{x})$
- si assume che un esperto (o maestro) ci fornisca la supervisione
i valori $f(x_i)$



Face recognition technologies



AT&T

LeNet 5

RESEARCH

answer: 4

4
433

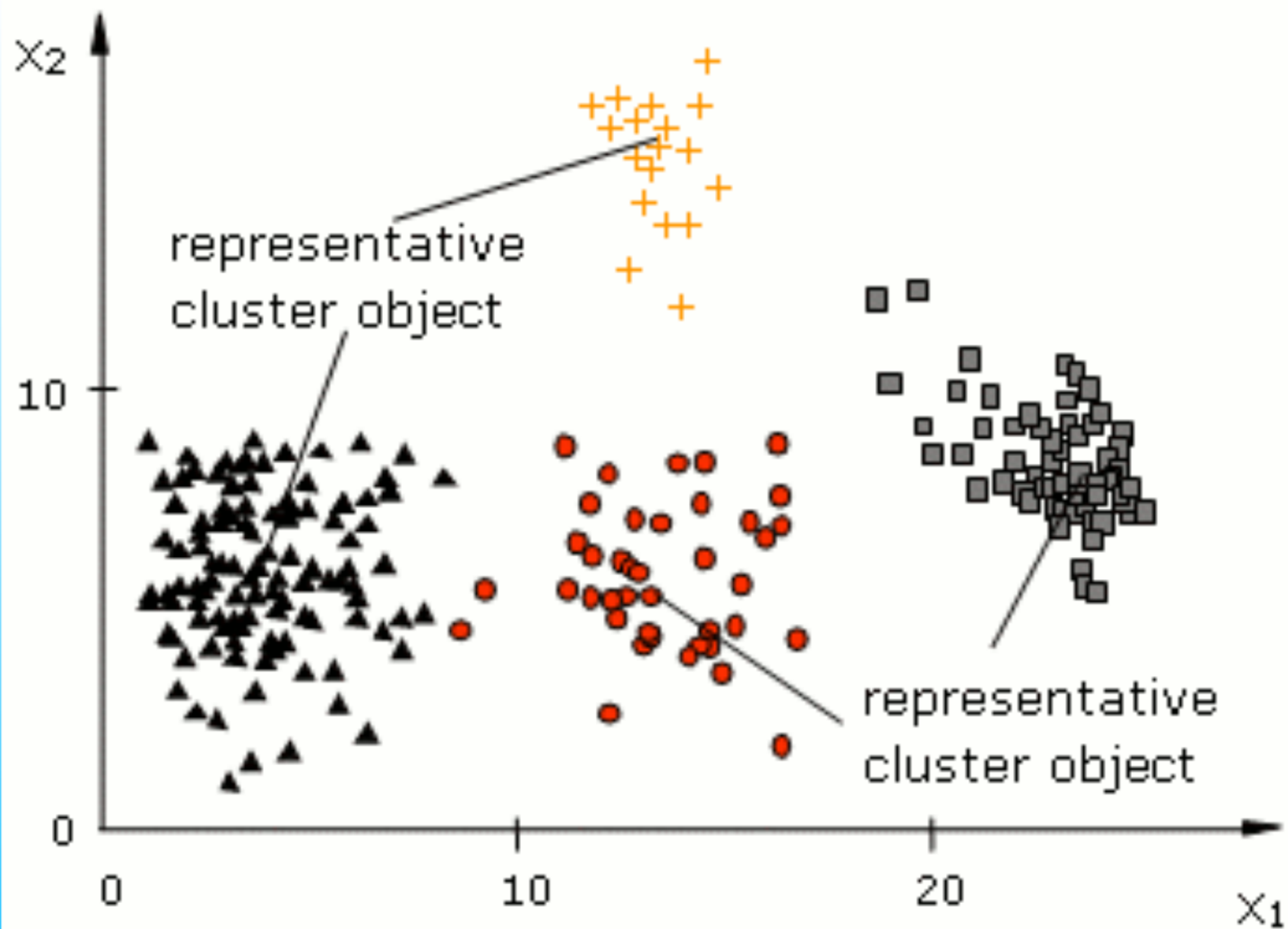
4

- si assume che un esperto (o maestro) ci fornisca supervisione i valori $f(x_i)$

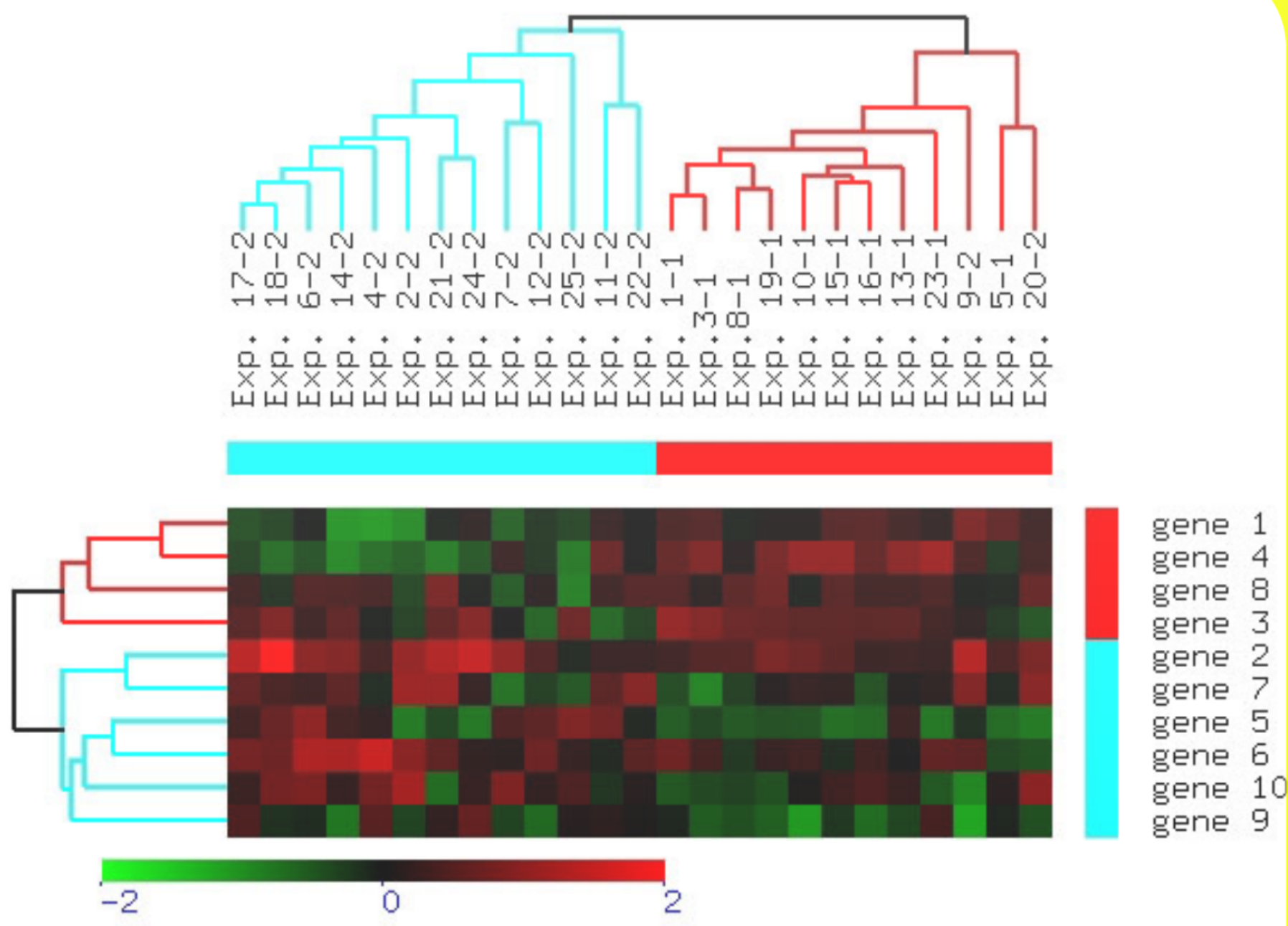
Unsupervised Learning (apprendimento non-supervisionato)

- dato un insieme di dati $\{x_i\}$, estrarre regolarità e/o pattern valide su tutto il dominio di ingresso
- non esiste un maestro che ci fornisca un aiuto

Learning (apprendimento con rinforzo)
PRIMO A LEZIONE



number of clusters: 4



Reinforcement Learning (apprendimento con rinforzo) NON LO VEDREMO A LEZIONE

Sono dati



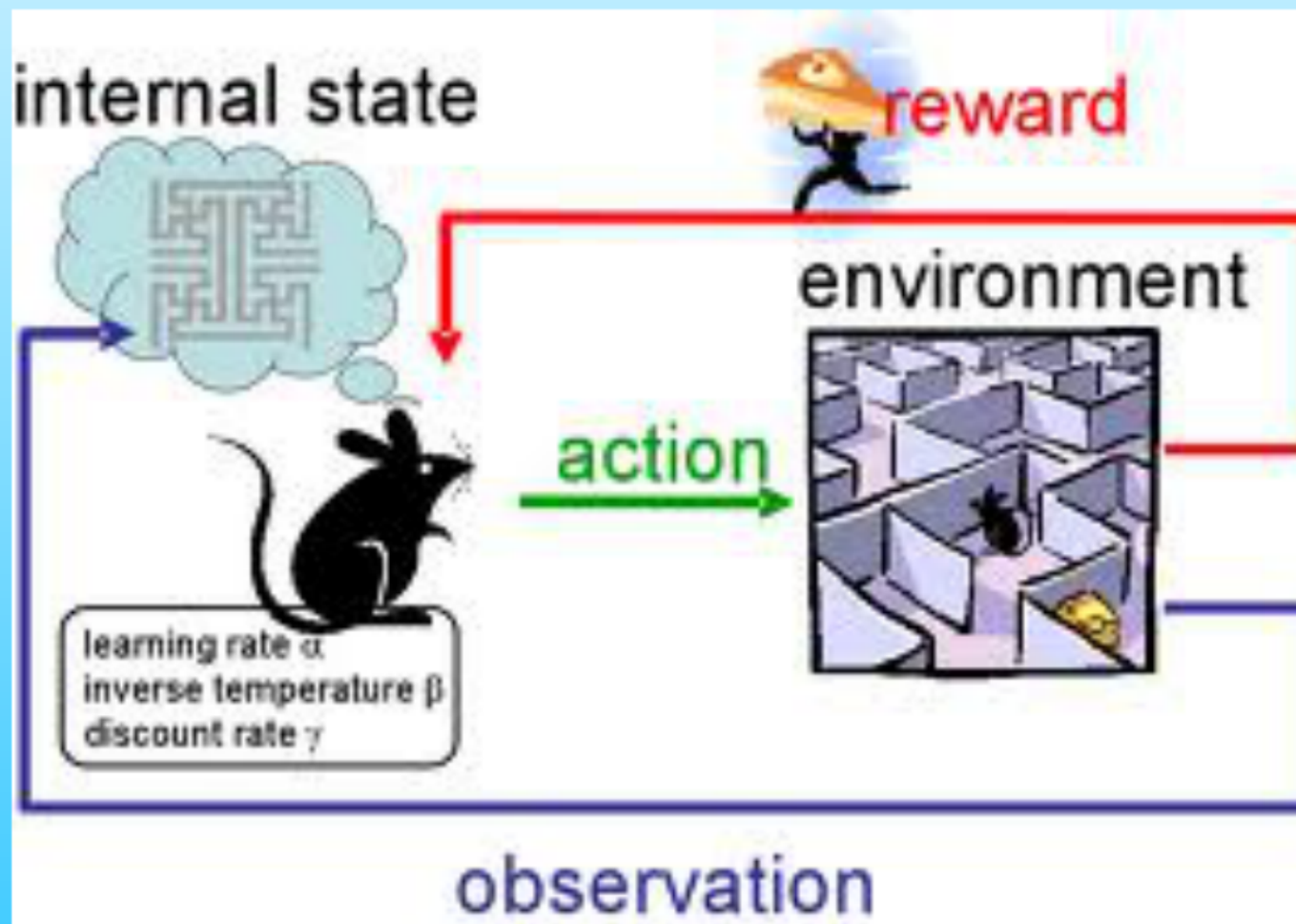
- un **agente** (intelligente?), che può
 - trovarsi in uno **stato** s
 - eseguire una **azione** a (all'interno delle azioni possibili nello stato corrente)
- un **ambiente** tale che quando l'agente applica una azione a nello stato s restituisce
 - lo **stato successivo**
 - una **ricompensa** r , che può essere positiva (+), negativa (-), o neutra (0)

Scopo dell'agente è quello di
massimizzare la ricompensa



$$\sum_{t=0}^{\infty} \gamma^t r_{t+1} \text{ dove } 0 \leq \gamma < 1)$$

scelta della strategia ottima



Ingredienti Fondamentali Apprendimento

Dati

(spazio delle istanze)



Spazio delle Ipotesi H

- costituisce l'insieme delle funzioni che possono essere realizzate dal sistema di apprendimento
- si assume che la funzione da apprendere f possa essere rappresentata da una ipotesi

$h \in H$

(selezione di h attraverso i dati di apprendimento)

- o che almeno una ipotesi h sia simile a f
(approssimazione)



Algoritmo di ricerca
nello spazio delle Ipotesi
(algoritmo di apprendimento)



h "ottima"

ipotesi restituita
dall'apprendimento

ATTENZIONE!!

lo spazio delle ipotesi H non può coincidere con l'insieme di tutte le funzioni possibili e la ricerca essere esaustiva



essere esaustiva



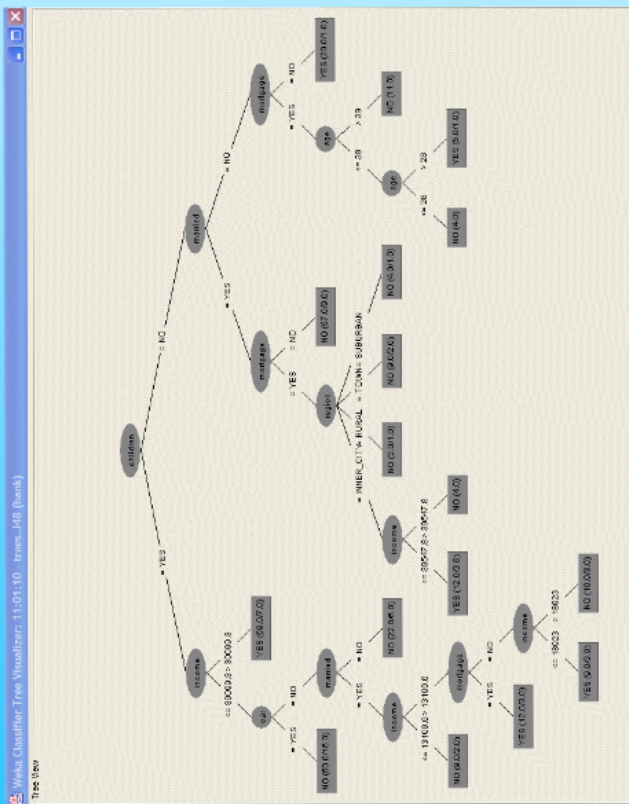
Apprendimento Inutile !!

Bias Induttivo

Bias Induttivo

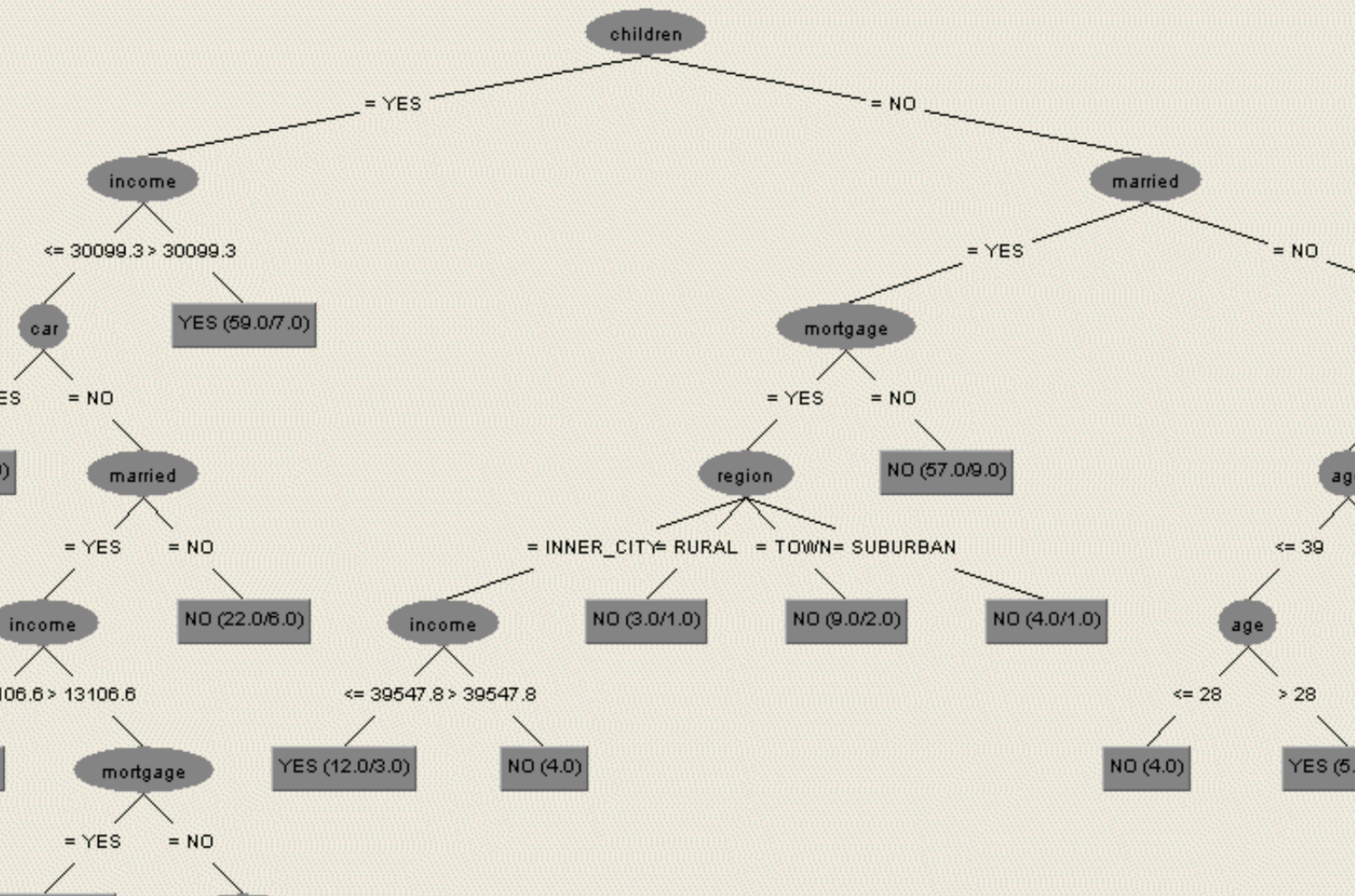
- sulla rappresentazione (H)
- sulla ricerca (alg. apprendimento)

Decision Trees

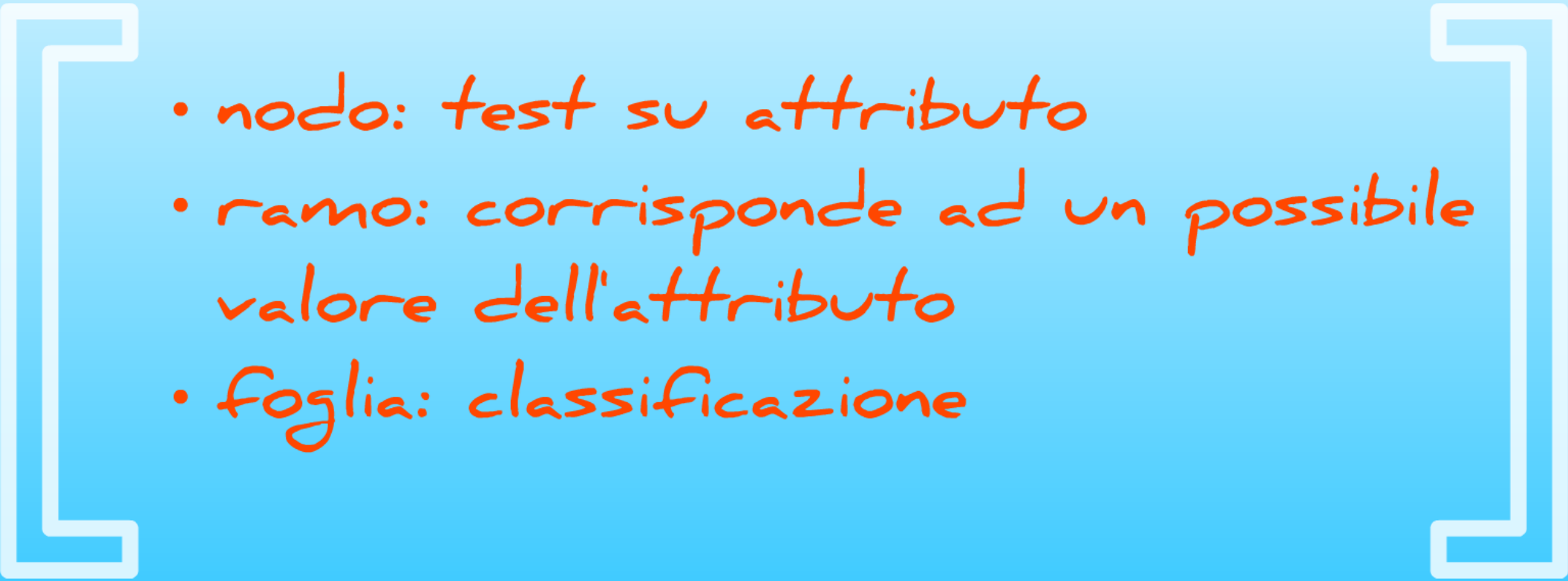


- istanze rappresentate da coppie attributo-valore
- classificazione multiclasse
- esempi di apprendimento possono contenere errori e/o valori mancanti

- nodo: test su attributo
- ramo: corrisponde ad un possibile valore dell'attributo
- foglia: classificazione

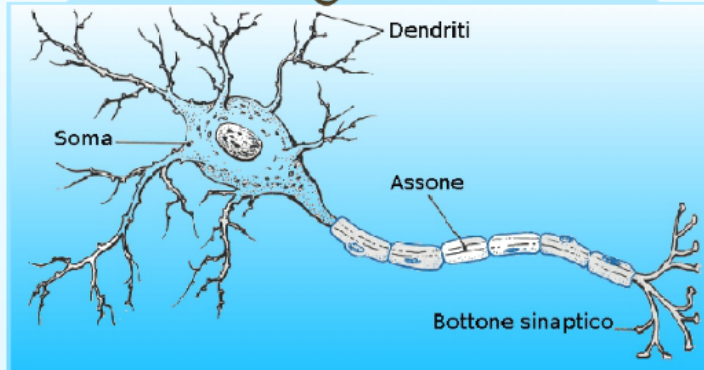


- esempi di apprendimento possono contenere errori e/o valori mancanti

- 
- nodo: test su attributo
 - ramo: corrisponde ad un possibile valore dell'attributo
 - foglia: classificazione

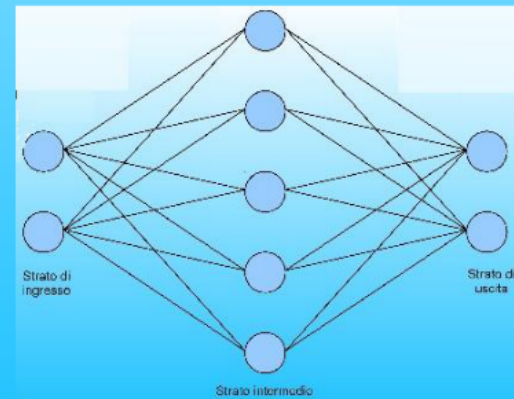
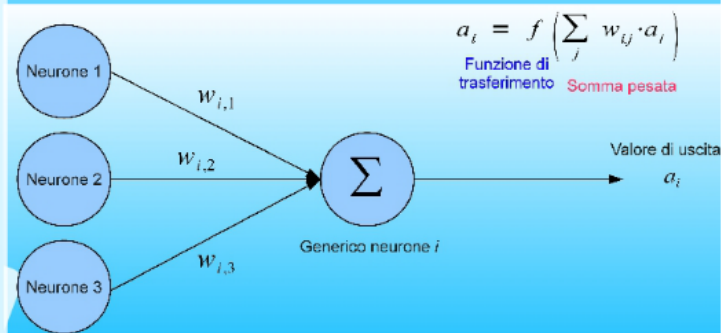
Neural Networks

Neurone biologico

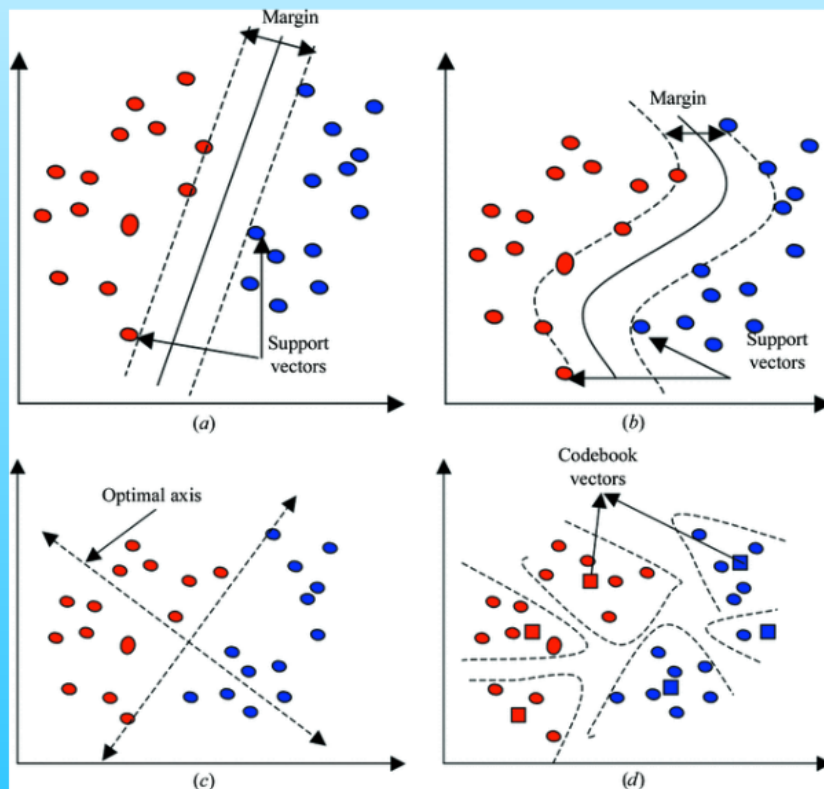


- si ispirano al cervello umano
- sia classificazione che regressione
- sia supervised che unsupervised
- adatte ad approssimare funzioni reali continue

Neurone artificiale



Kernel Methods



- sia supervised che unsupervised
- sfruttano le funzioni kernel
- possono trattare direttamente dati strutturati
- fra le migliori tecniche di apprendimento automatico

Vedremo anche...

- Rappresentazione delle istanze ("feature selection")
- Apprendimento Probabilistico (bayesiano)
- Algoritmi "semplici" di clustering
- Esperienze pratiche in aula informatica



e non finisce qui...