

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.

Perchè 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: non si può fare a mano;
- mancanza di **conoscenza "compilata"** rispetto al problema da risolvere;

Ruolo dei Dati

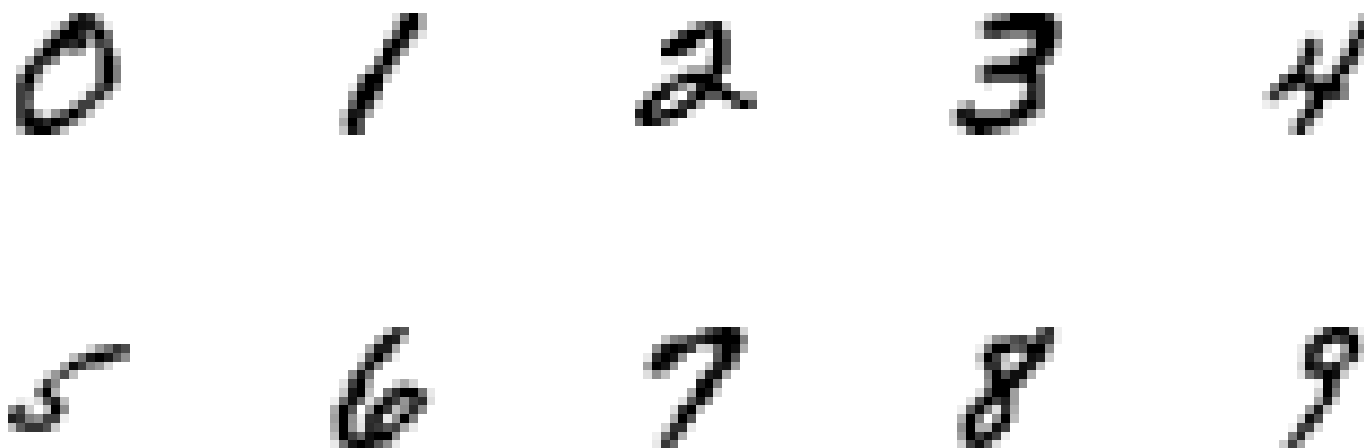
Tipicamente...

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

Desiderio: usare i dati per

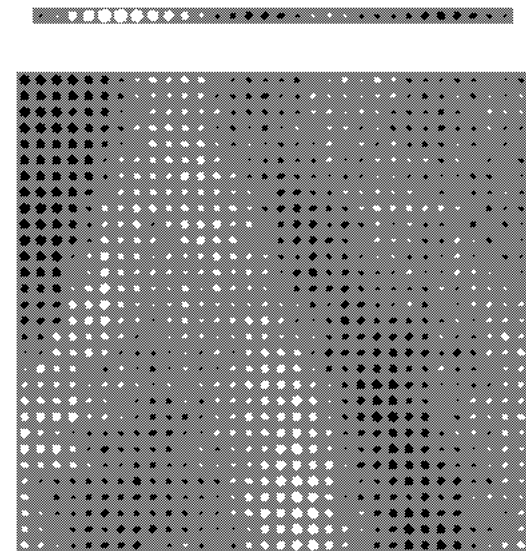
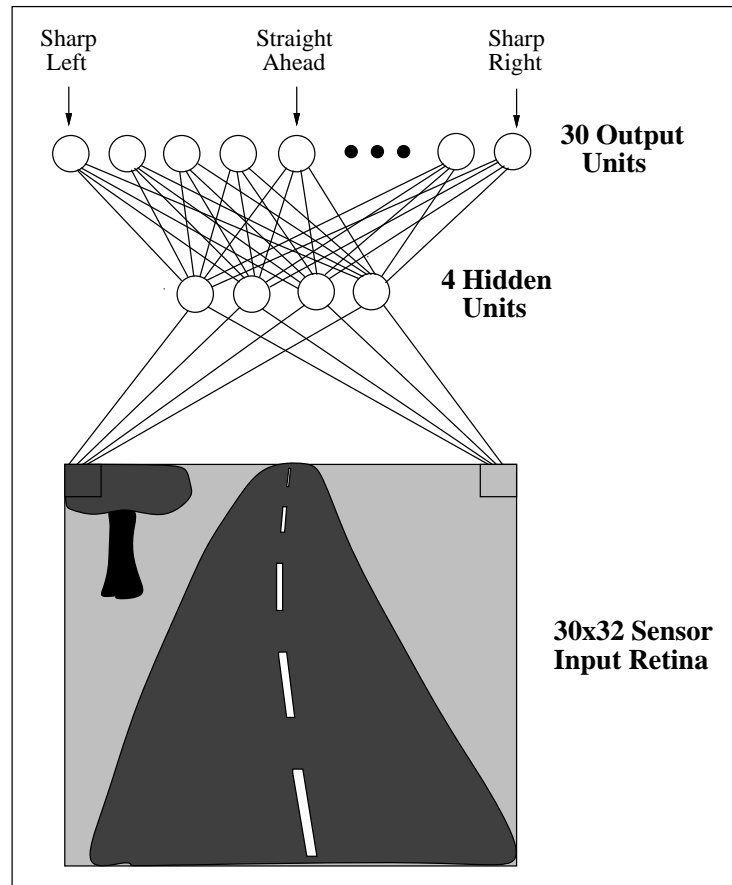
- ottenere **nuova conoscenza**;
- **raffinare** la conoscenza di cui si dispone;
- **correggere** la conoscenza di cui si dispone;

Es. - Riconoscimento di Cifre Manoscritte



- impossibile **formalizzare** esattamente il problema: disponibili solo esempi;
- possibile presenza di **rumore** e **dati ambigui**;

Es. - Guidare una Automobile



Es. - Estrarre Conoscenza Medica dai Dati

<i>Patient103</i> time=1	→	<i>Patient103</i> time=2	...	→	<i>Patient103</i> time=n
Age: 23		Age: 23			Age: 23
FirstPregnancy: no		FirstPregnancy: no			FirstPregnancy: no
Anemia: no		Anemia: no			Anemia: no
Diabetes: no		Diabetes: YES			Diabetes: no
PreviousPrematureBirth: no		PreviousPrematureBirth: no			PreviousPrematureBirth: no
Ultrasound: ?		Ultrasound: abnormal			Ultrasound: ?
Elective C-Section: ?		Elective C-Section: no			Elective C-Section: no
Emergency C-Section: ?		Emergency C-Section: ?			Emergency C-Section: Yes
...	

Linee di Ricerca all'interno dell' Apprendimento Automatico

- induzione di regole/alberi di decisione,
- algoritmi connessionisti (reti neurali),
- “clustering” & “discovery”,
- apprendimento basato sulle istanze
- apprendimento Bayesiano,
- apprendimento basato sulla spiegazione,
- apprendimento con rinforzo,
- apprendimento induttivo guidato dalla conoscenza,
- ragionamento per analogia & basato sui casi,
- algoritmi genetici,
- programmazione logica induttiva, . . .

Principali Paradigmi di Apprendimento

Apprendimento Supervisionato:

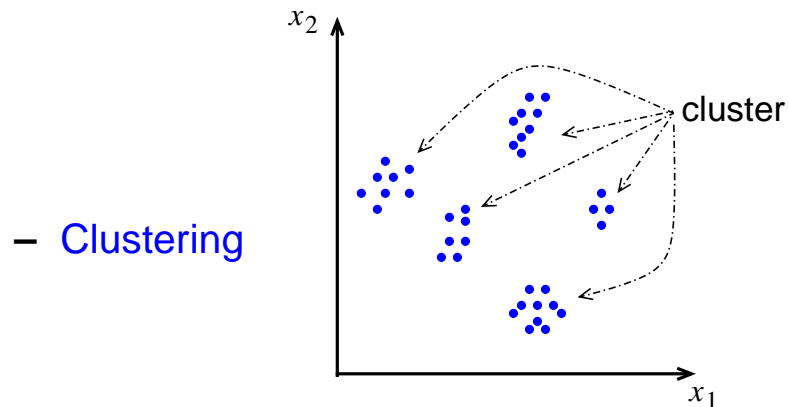
- dato in insieme di esempi pre-classificati, $Tr = \{(x^{(i)}, f(x^{(i)}))\}$, apprendere una descrizione generale che incapsula l'informazione contenuta negli esempi (regole valide su tutto il dominio di ingresso)
- tale descrizione deve poter essere usata in modo predittivo (dato un nuovo ingresso \tilde{x} predire l'output associato $f(\tilde{x})$)
- si assume che un esperto (o maestro) ci fornisca la supervisione (cioè i valori della $f()$ per le istanze x dell'insieme di apprendimento)

Esempio di applicazione: classificazione di caratteri manoscritti

Principali Paradigmi di Apprendimento

Apprendimento Non-supervisionato:

- dato in insieme di esempi $Tr = \{x^{(i)}\}$, estrarre regolarità e/o pattern (valide(i) su tutto il dominio di ingresso)
- non esiste nessun esperto (o maestro) che ci fornisca un aiuto



- Scoperta di Regole (Discovery)

Esempio di applicazione: data mining su database strutturati

Principali Paradigmi di Apprendimento

Apprendimento con Rinforzo:

- Sono dati:
 - agente (intelligente ?), che può
 - * trovarsi in uno stato s , ed
 - * eseguire una azione a (all'interno delle azioni possibili nello stato corrente)
 - ed opera in un ambiente e , che applicando una azione a nello stato s restituisce
 - * lo stato successivo, e
 - * una ricompensa r , che può essere positiva (+), negativa (-), o neutra (0).
- Scopo dell'agente è quello di massimizzare una funzione delle ricompense (es. ricompensa scontata: $\sum_{t=0}^{\infty} \gamma^t r_{t+1}$ dove $0 \leq \gamma < 1$)

Esempio di applicazione: navigare sul Web alla ricerca di informazione focalizzata

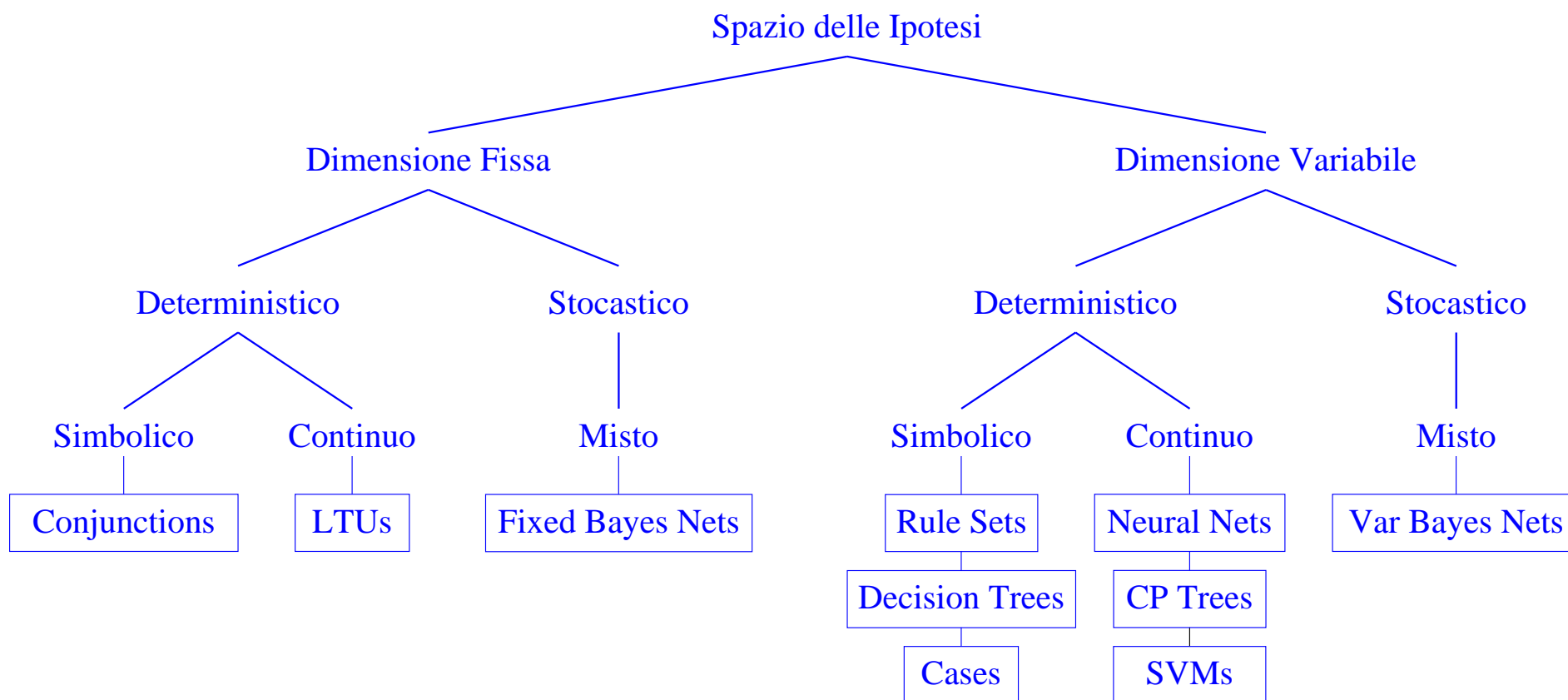
Ingredienti Fondamentali

- Dati di Allenamento
- Spazio delle Ipotesi, \mathcal{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 $h \in \mathcal{H}$... (selezione di h attraverso i dati di apprendimento)
 - o che almeno una ipotesi $h \in \mathcal{H}$ sia simile a f (approssimazione);
- Algoritmo di Ricerca nello Spazio delle Ipotesi, alg. di apprendimento

ATTENZIONE: \mathcal{H} non può coincidere con l'insieme di tutte le funzioni possibili e la ricerca essere esaustiva → **Apprendimento è inutile!!!**

Si parla di **Bias Induttivo**: sulla rappresentazione (\mathcal{H}) e/o sulla ricerca (alg. di apprendimento)

Tassonomia (non completa) dello Spazio delle Ipotesi



Tassonomia (non completa) degli Algoritmi di Apprendimento

