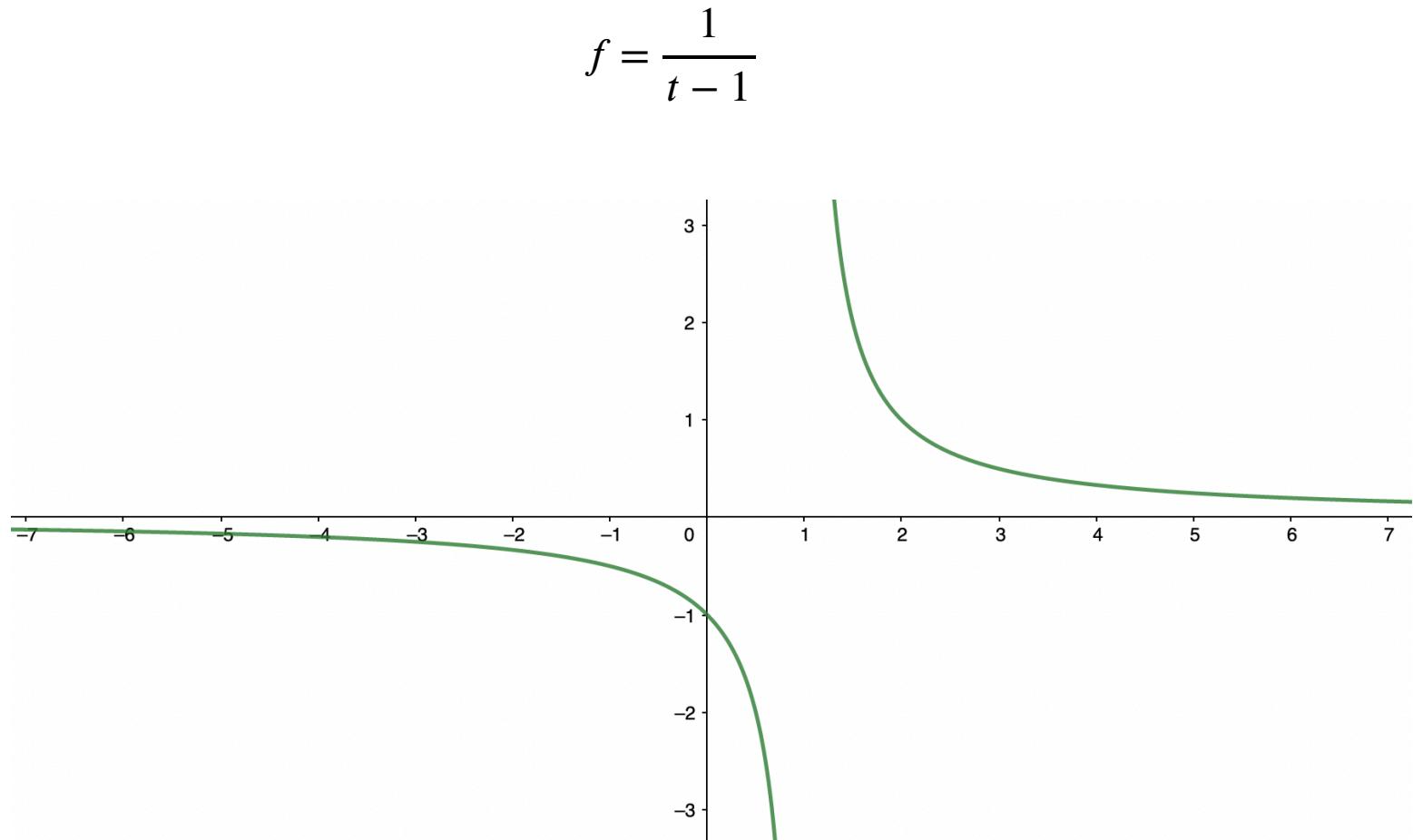
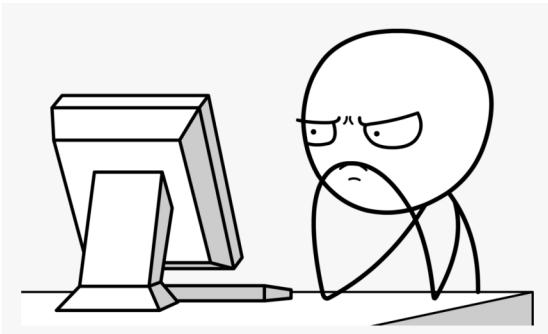


Programma la tua prima intelligenza artificiale

Alessandro Cossard

Introduzione

- Problema dei tre corpi
- Uso del **computer**
- NON è un corso di puro Python

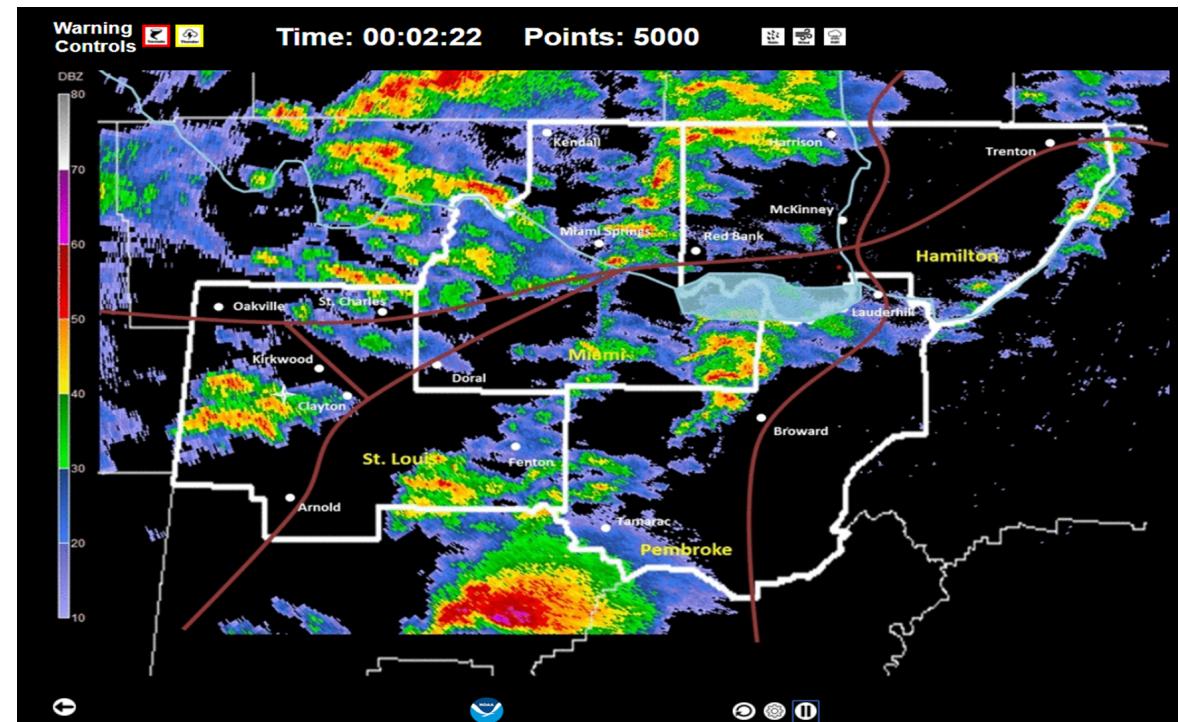


Modellizzazione

- **Descrivere** un sistema complesso
- Trovare gli aspetti significativi
- Adattarlo ad altri sistemi e non specifico
- Simulazioni al pc = programmare



Giorgio Parisi



Programmare

- Parte teorica
- Anaconda e Google Colab per tensorflow
- Notebook

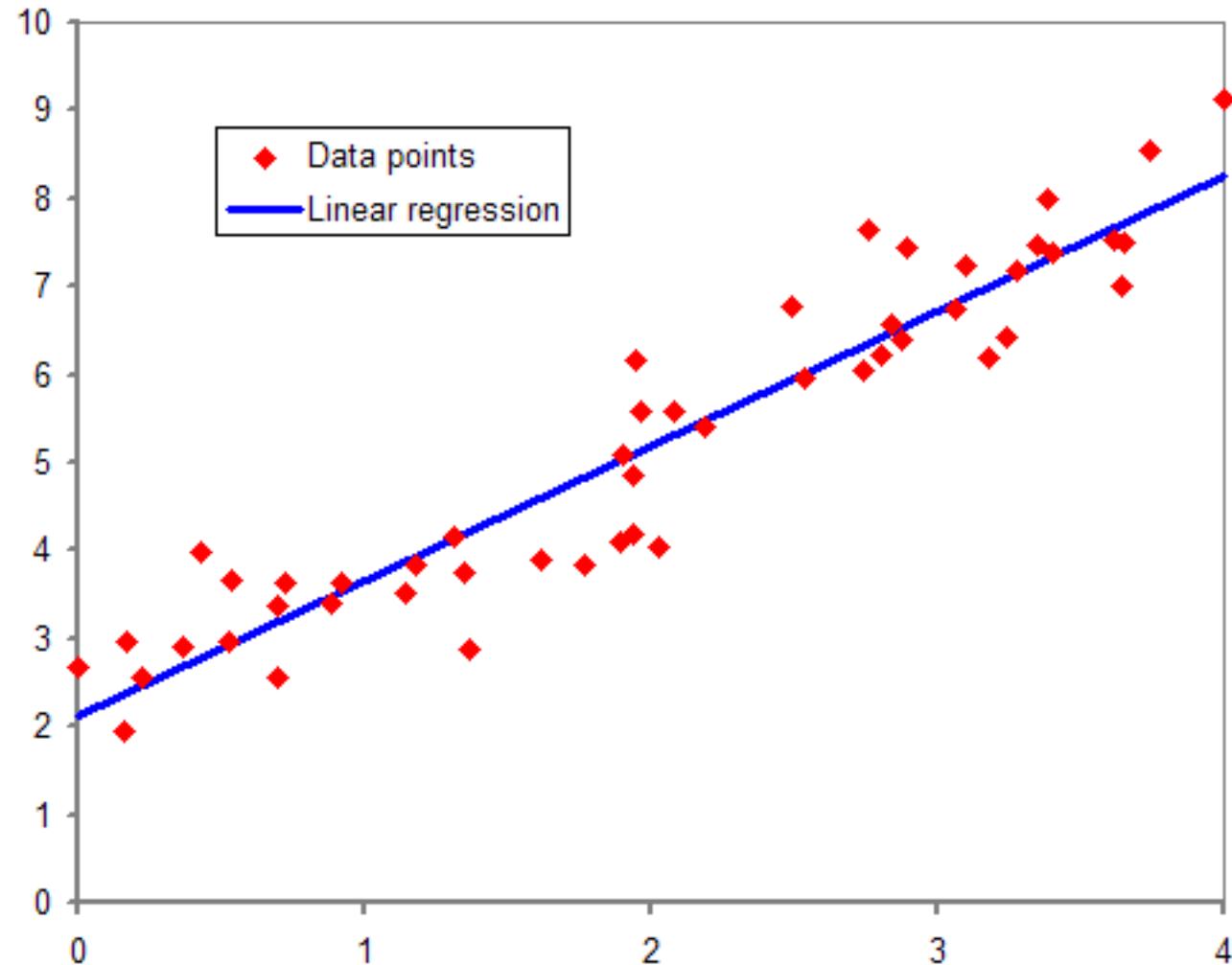


Data Science

- Machine Learning Vs Artificial Intelligence
 - Machine Learning: apprendimento automatico dai dati (predizione)
 - Usato **ovunque** (riconoscimento volti Facebook, suggeriti Spotify, importante l'applicazione in medicina)
 - Scopo: creare un modello **predittivo** generalizzabile e automatico
 - Fondamentale l'utilizzo e lo studio dei **dati**
-
- **Notebook**

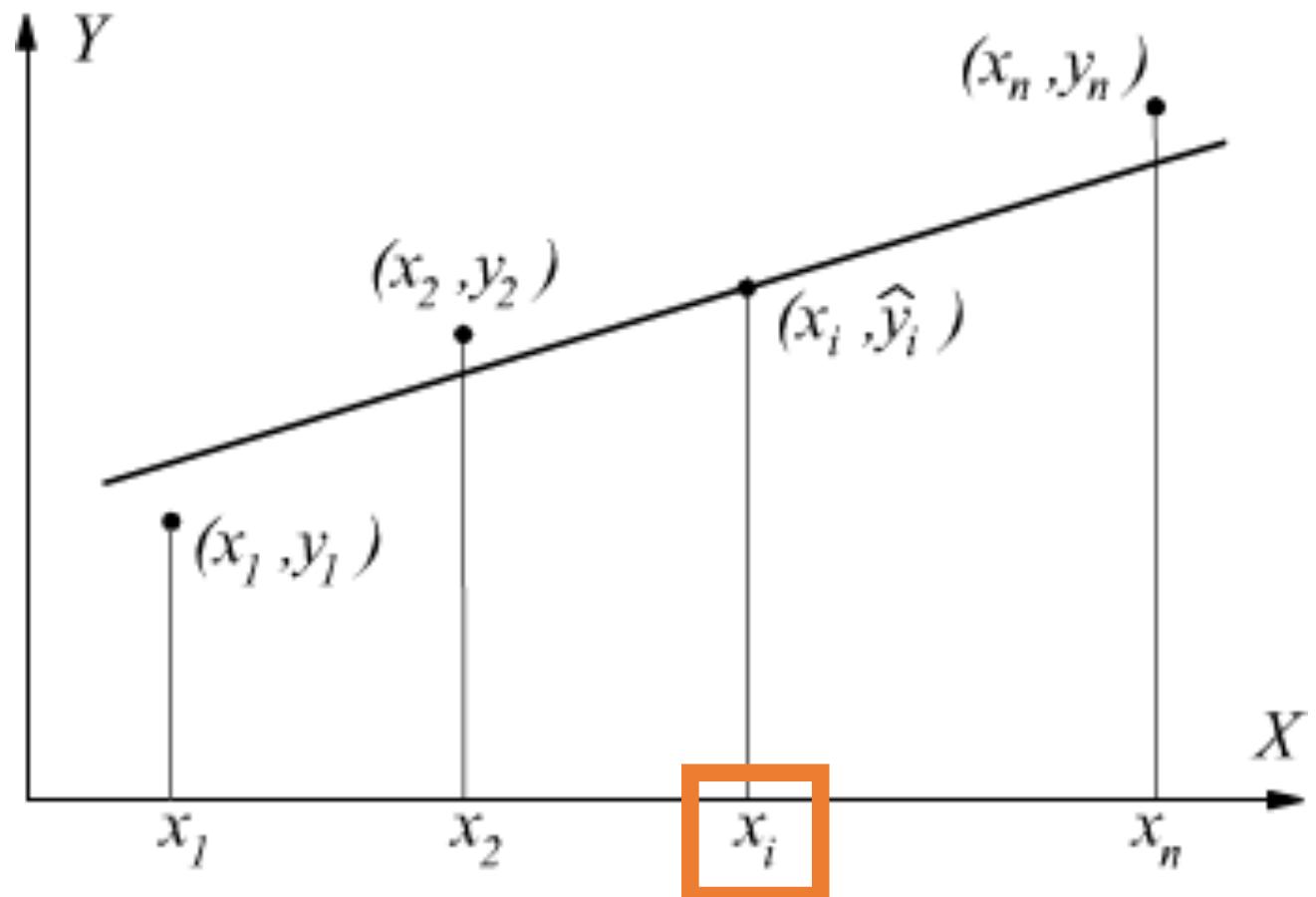
Regressione lineare

- Metodo **supervised**
- Regressione perché valore reale
- Quantificare la **dipendenza** tra due variabili



Regressione lineare

- Predire valori nuovi a partire da quelli noti
- Uso la retta di regressione come **predict**



Regressione lineare: l'algoritmo

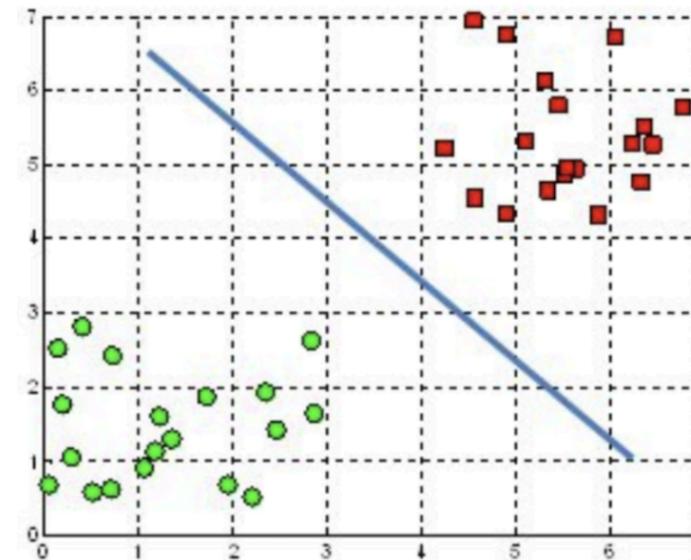
- $$h(x) = \sum_{i=0}^d w_i x_i$$
- Quanto $h(x)$ **approssima** bene $f(x)$
- Uso l'errore quadratico
- Minimizzo l'errore modificando i pesi
(geometricamente sposto l'iperpiano)
- Training Vs Test

$$E = \frac{1}{N} \sum_{i=1}^n [h(x_n) - y_n]^2 = \frac{1}{N} \sum_{i=1}^n \left[\sum_{i=0}^d w_i x_i - y_n \right]^2$$

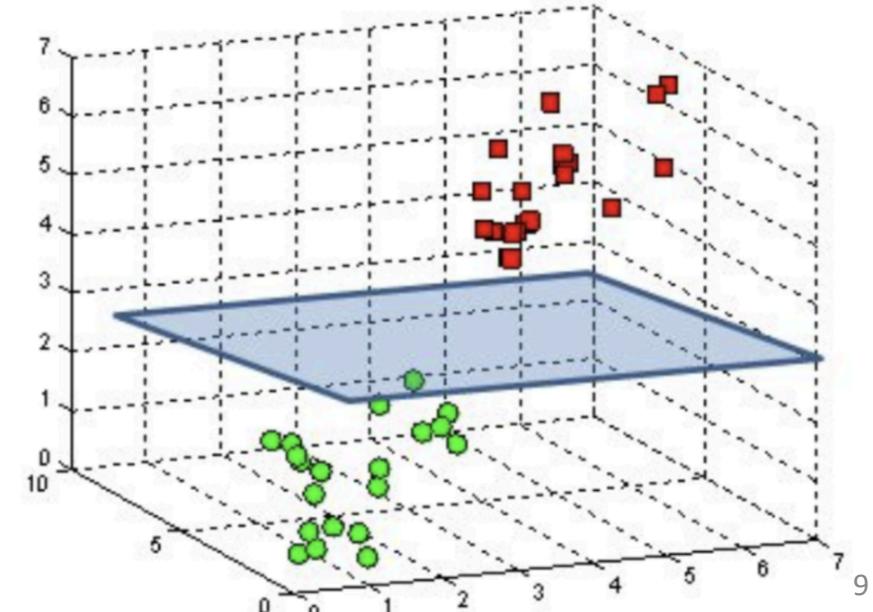
Statistica Multivariata

- La statistica multivariata guarda come cambiano più variabili in modo **indipendente**
- Quindi dopo i 3 assi non posso più fare i plots perchè è come se fossi in dimensioni maggiori a 3

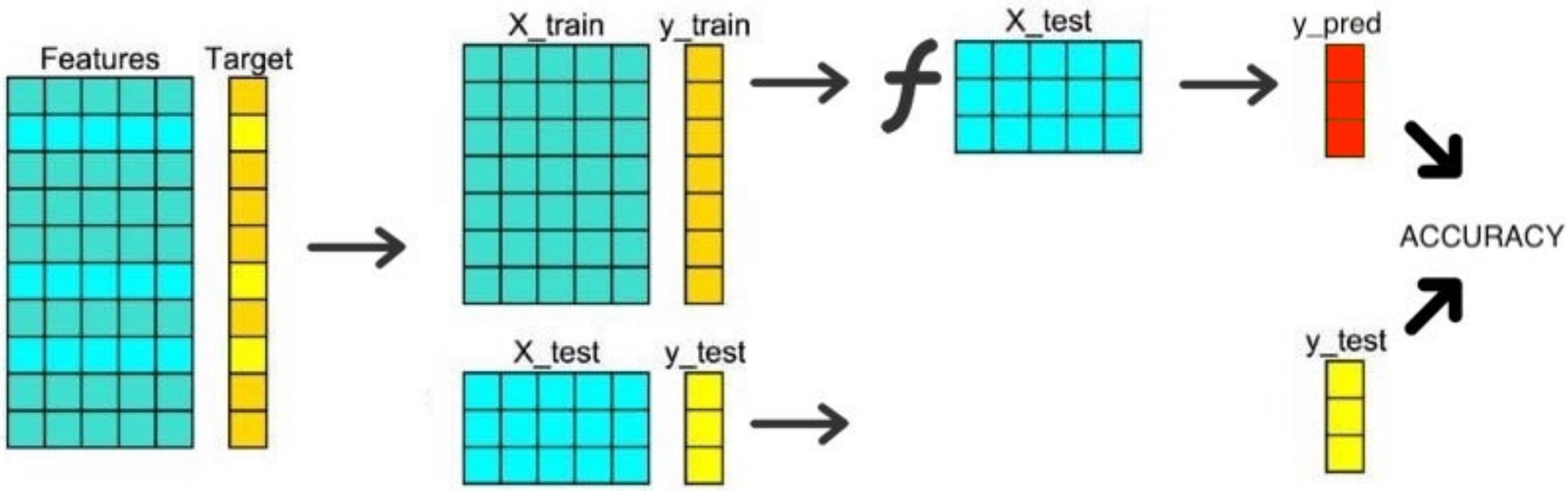
A hyperplane in \mathbb{R}^2 is a line



A hyperplane in \mathbb{R}^3 is a plane

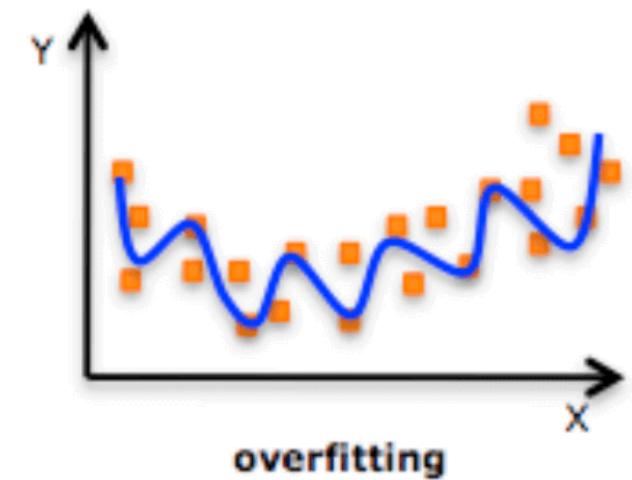
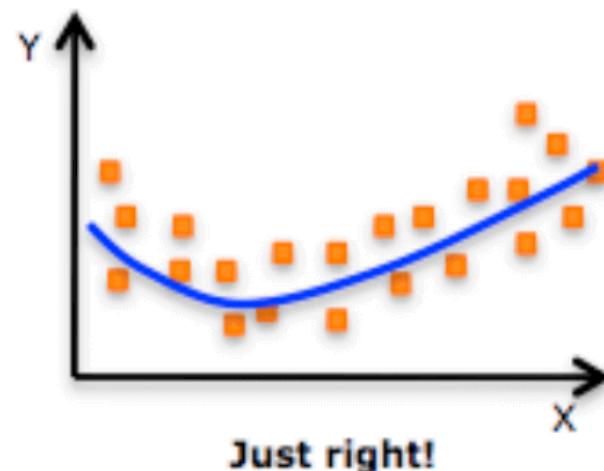


Regressione lineare - Train Vs Test



Overfitting

- Non voglio il 100% di accuracy perchè significa che il mio modello, classificando tutto correttamente, si è **adattato troppo ai dati**
- Quindi **perde** di generalizzazione nel caso di out-of-sample data



Esempio della banca

- Problema
- Dati
- Soluzione
- Banca

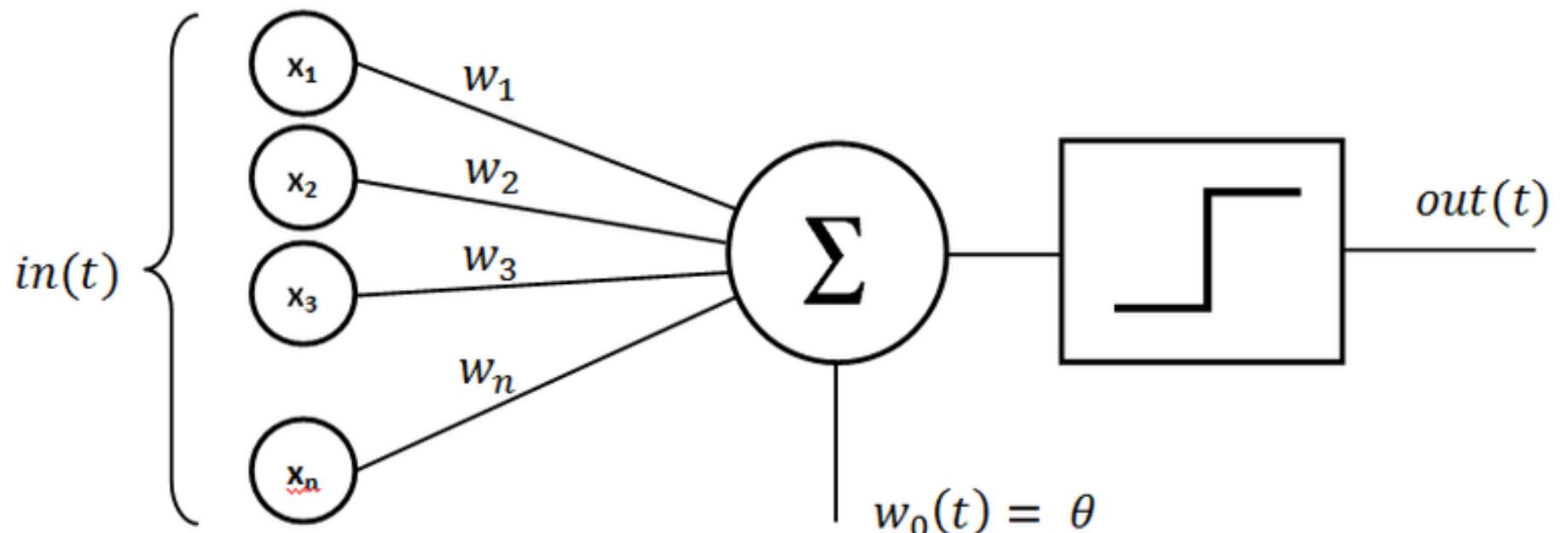
$$h(x) = \text{sign}\left(\sum_{i=0}^d w_i x_i\right)$$



Clienti	Casa	Macchina	Stipendio	Nucleo	Salute	Label
Cliente 1	si	si	15k	4	si	si
Cliente 2	no	Si	12k	1	si	si
Cliente 3	si	no	8k	3	si	no
Cliente N	no	si	16k	6	no	no

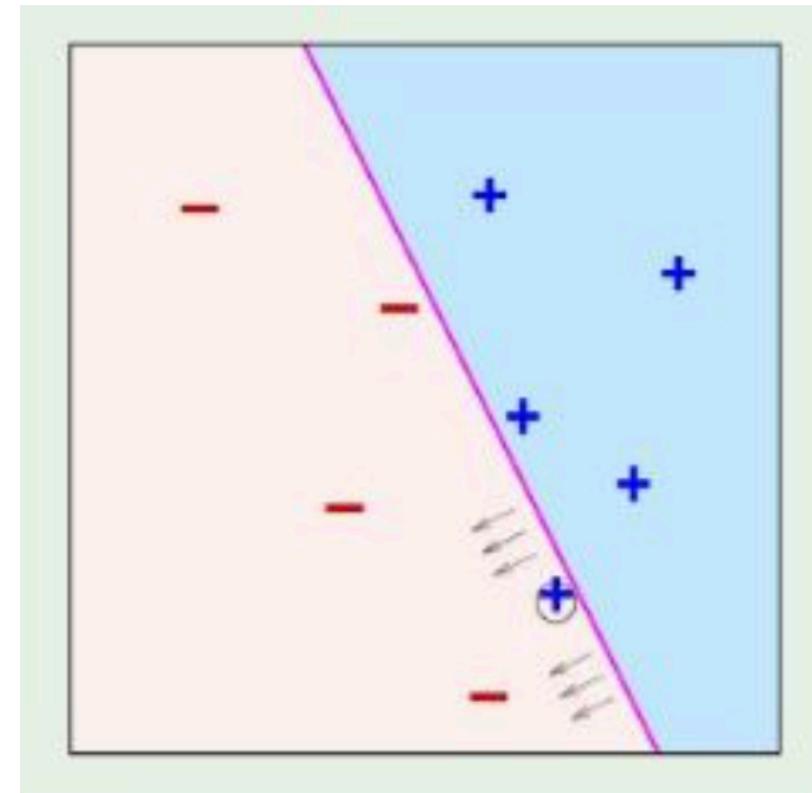
Perceptron

- Neurone artificiale
- Modello lineare che richiede la separabilità dei dati
- Input, pesi, somma, funzione di attivazione, output



Perceptron Learning Algorithm

- Inizializza i pesi e la threshold random
- Calcola la somma pesata della features
- Applica la funzione di attivazione per avere la Y_{pred}
- Confronta la Y_{pred} con la Y di quel singolo punto
- Se $Y_{pred} \neq Y$ aggiorna i pesi secondo
$$w_i = w_i + \alpha(Y_{true} - Y_{pred})x_i$$
dove α è il learning rate
- Ripetere tutti gli step fino alla convergenza

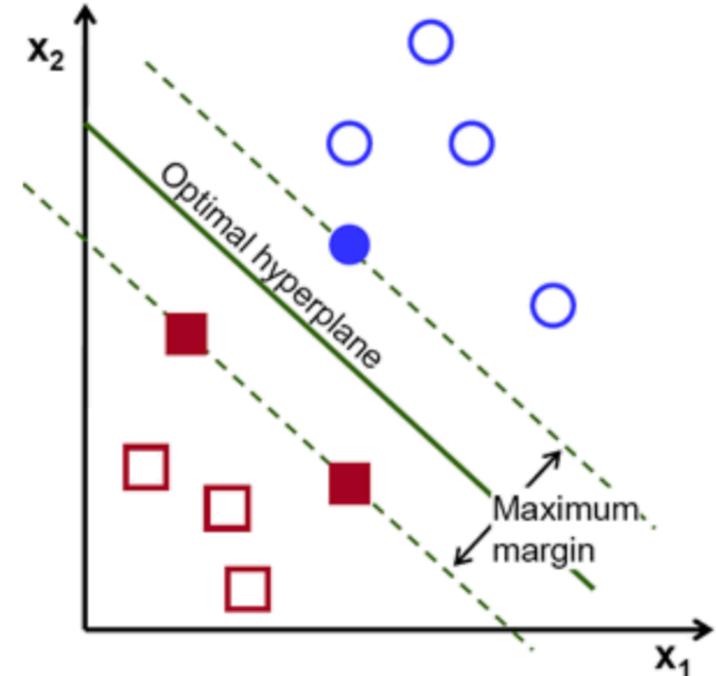
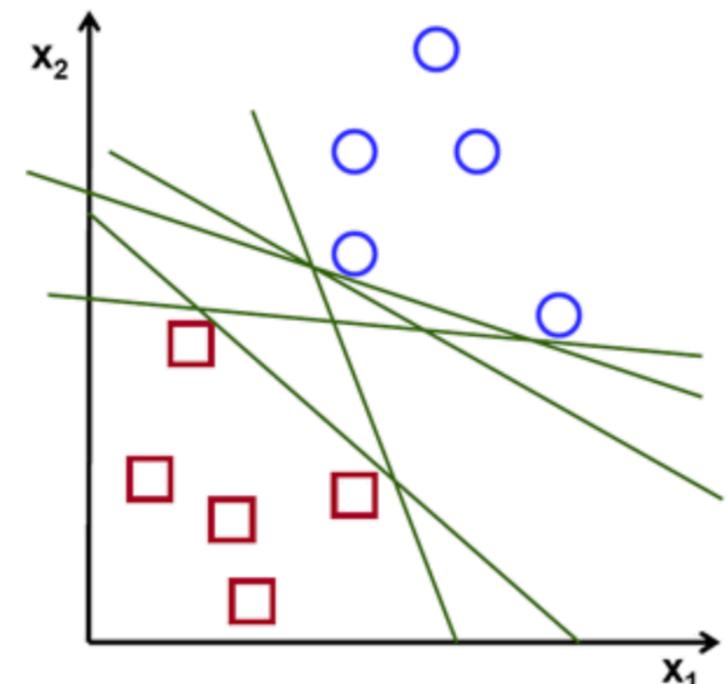


Scelta e Morale

- Neuromarketing
- Eyetracking: <https://www.webmarketingstrategico.net/eye-tracking/>
- https://en.wikipedia.org/wiki/Vance_Packard
- <https://www.snopes.com/fact-check/subliminal-advertising/>
- Morale?

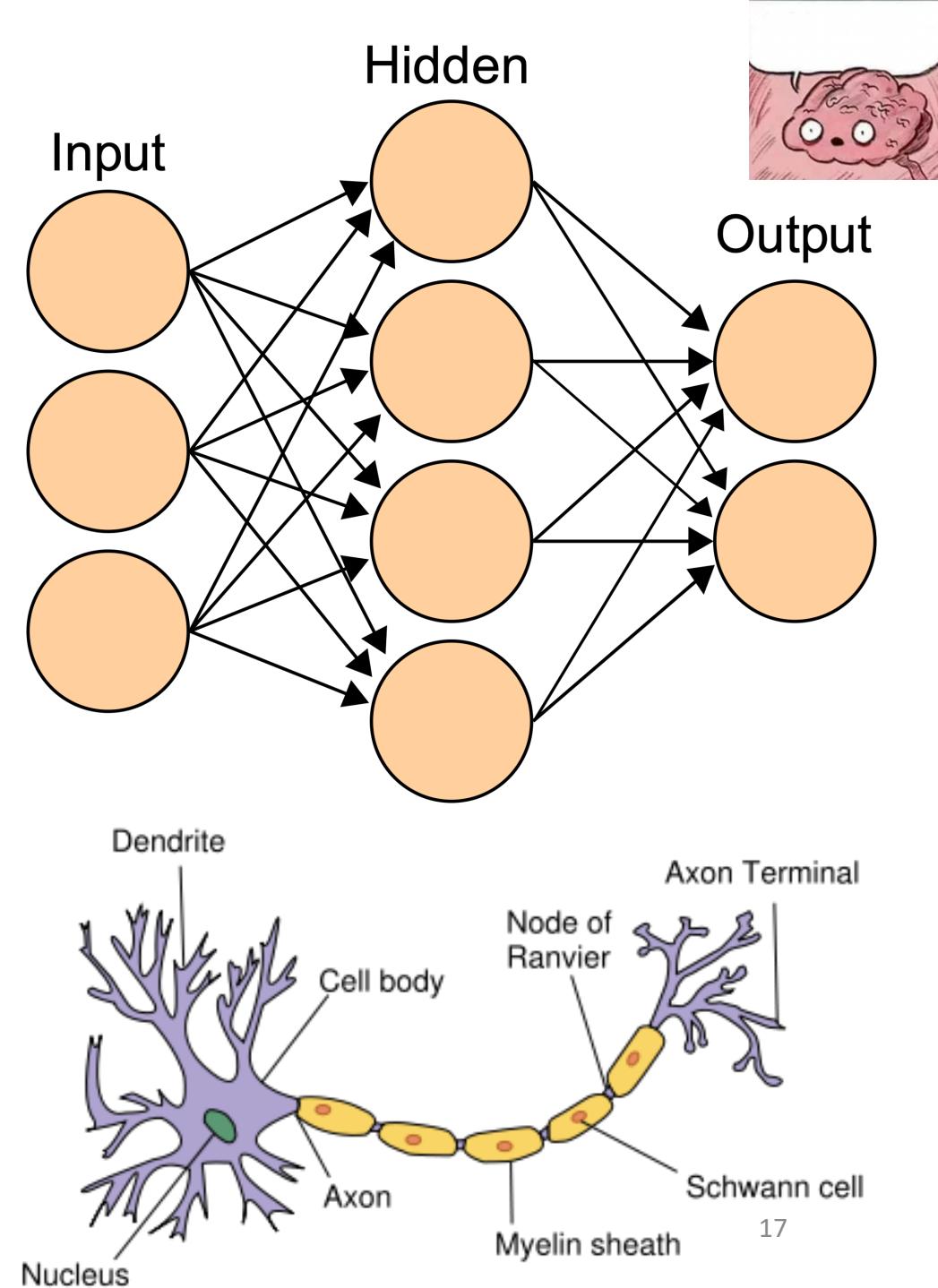
Support Vector Machine

- Simile al Perceptron MA
1) Non richiede la separabilità dei dati perchè li **proietta** usando il Kernel
2) Al perceptron basta un piano qualunque mentre SVM cerca i **Support Vectors** per selezionarlo



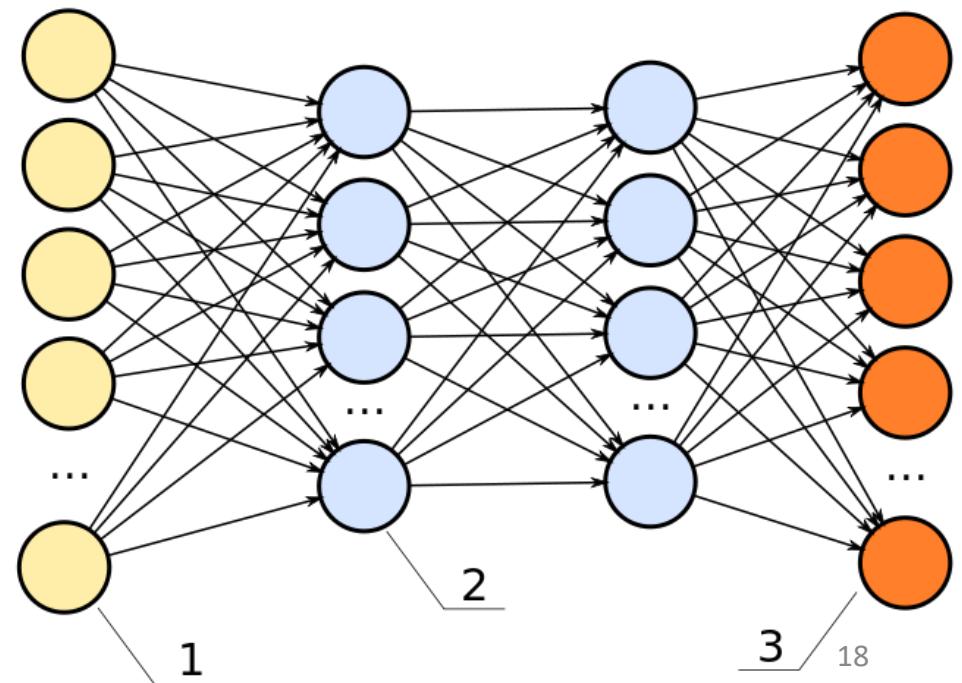
Rete Neurale

- Multilayer perceptron
- Sistema di apprendimento formato da diverse sottocomponenti, alcune di input, altre di output e altre ancora nascoste (comunicative)
- Simile al cervello umano: se il segnale dei dendriti supera una certa soglia, il neurone emette un segnale elettrico
- La rete va allenata modificando i pesi (Black Box)



Deep Neural Network

- Almeno due o tre **hidden layers (deep learning)**
- Risultati migliori ma più difficile da gestire
- Supervised: posso controllare il risultato
- Non linearità gestita da diverse soglie di attivazione
- Loss function: minimizza con i gradienti che hanno sia la magnitude che la direction



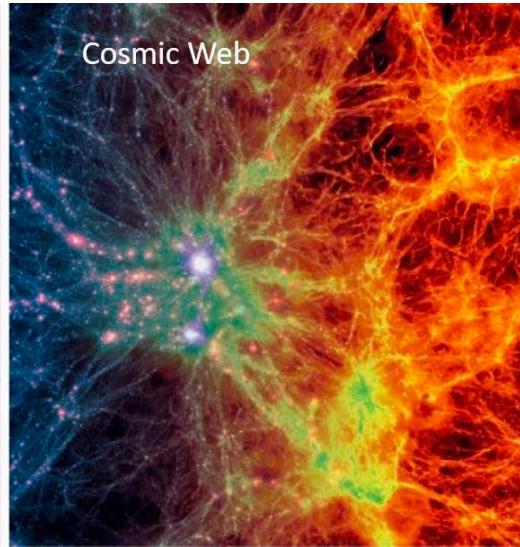
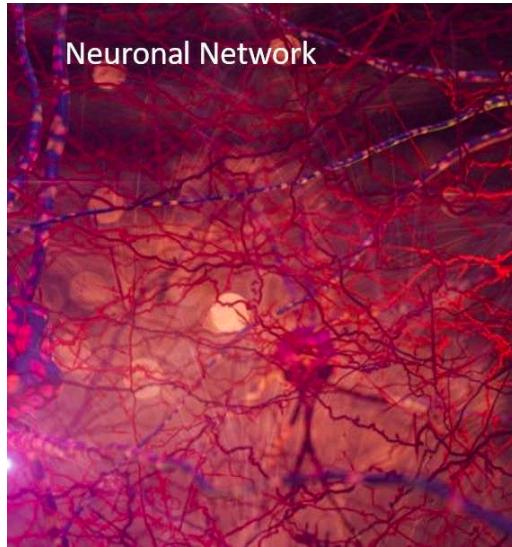
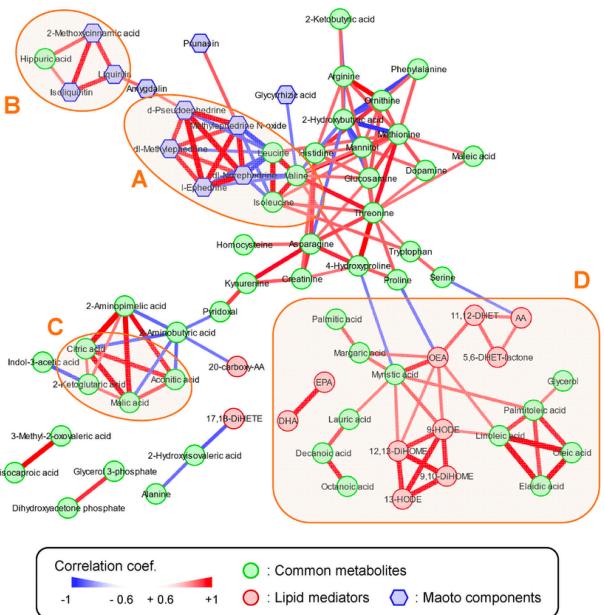
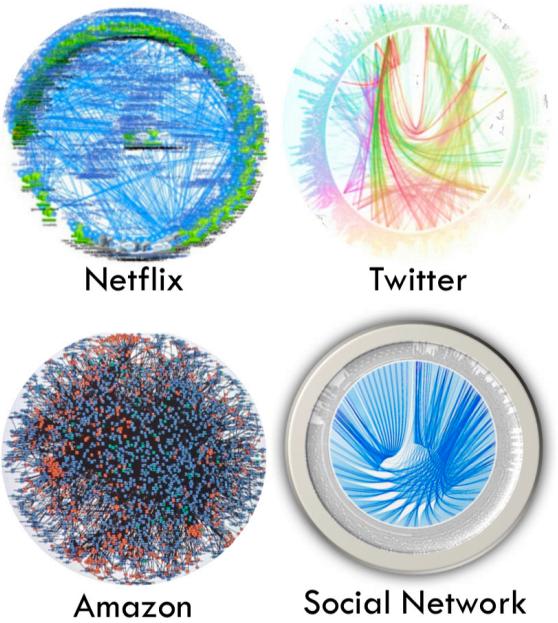
Backpropagation Algorithm

- Reti a Backpropagation Algorithm:
 1. Inizializzazione pesi random
 2. Fase feed forward
 3. Confronto labels (supervised)
 4. Valutazione dell'errore
 5. Aggiornamento dei pesi
 6. Repeat
- L'algoritmo è interrotto quando l'errore è più piccolo della soglia desiderata
- Backpropagation = propaga indietro l'errore per aggiustare i pesi
- Vado avanti così finche il la rete non classifica bene e poi test set

Classificazione di Iris

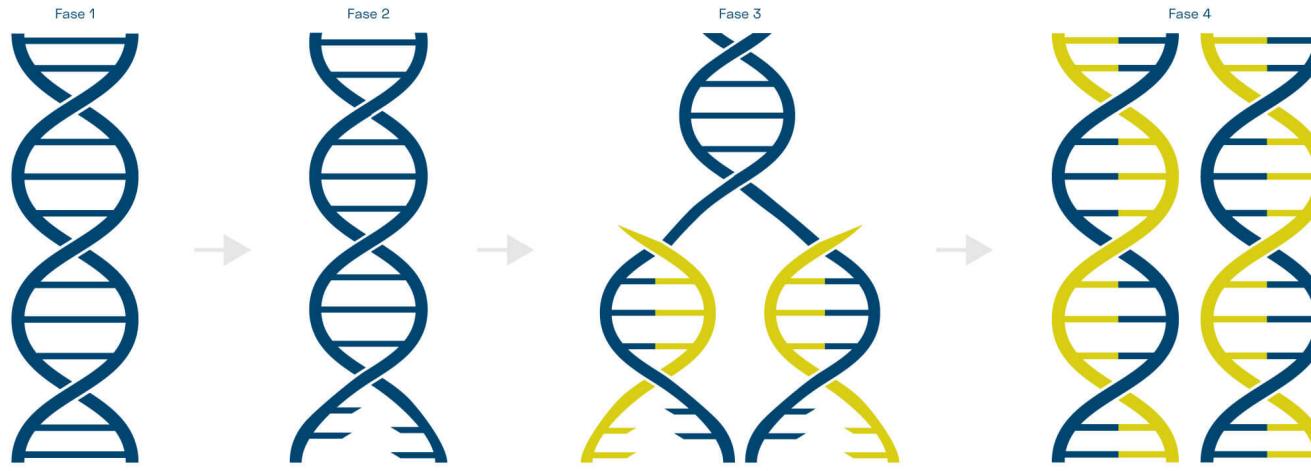
- Tensorflow e Keras: occupano un sacco di spazio sul PC
- Upload del notebook su Colab
- Esercizio: tunare il modello e studiare l'andamento dell'accuracy
- Notebook

Complex systems



What is Life?

- “*For otherwise the number of co-operating particles would be too small, the 'law' too inaccurate. The particularly exigent demand is the square root. For though a million is a reasonably large number, an accuracy of just 1 in 1,000 is not overwhelmingly good, If a thing claims the dignity of being a 'Law of Nature'"*



DNA
Processo di duplicazione semiconservativo



Ordine di Scala

- Concetti dei gas perfetti alla **larga scala**
- Particelle alla **microscala**
- Cosa succede quando passo da una all'altra?



Comportamento Emergente



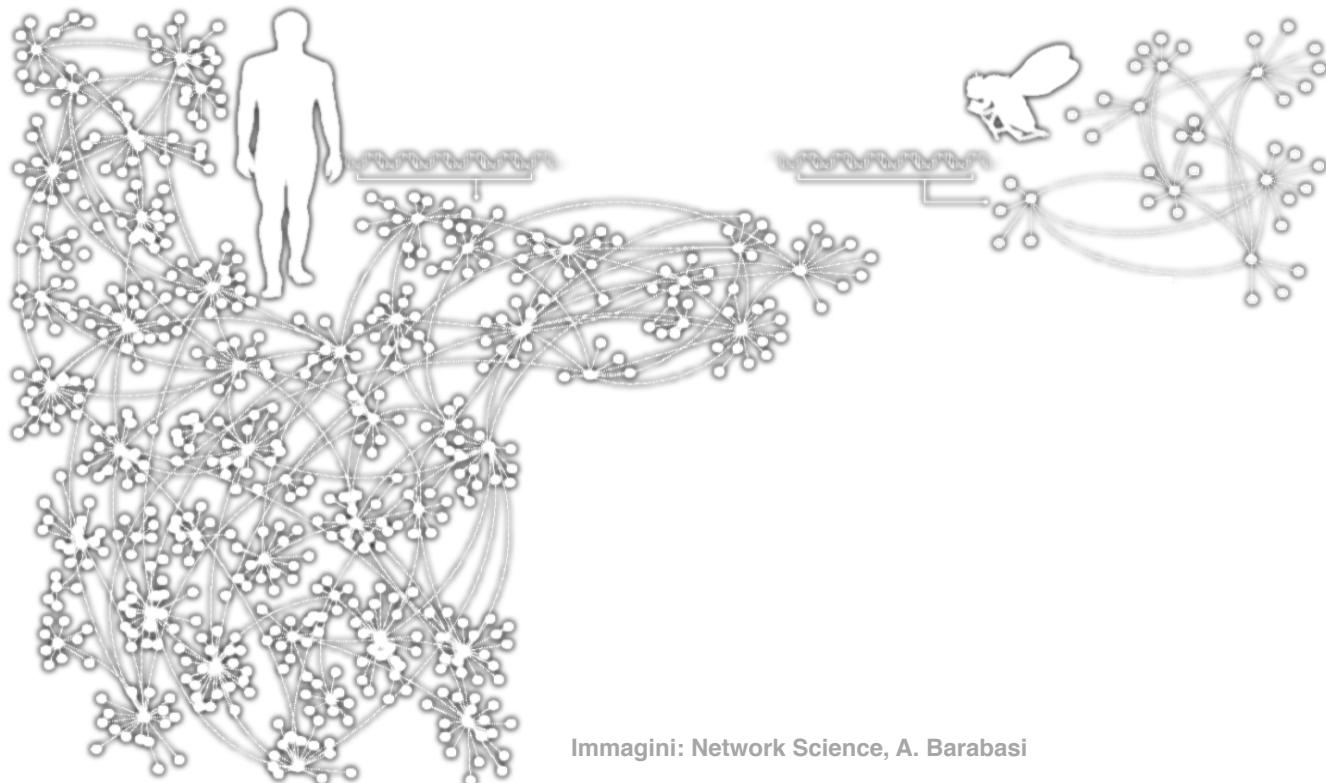
Alessandro Cossard

Interazioni non lineari



Reti

- Dietro ogni sistema complesso c'è una **rete** che ne descrive il comportamento

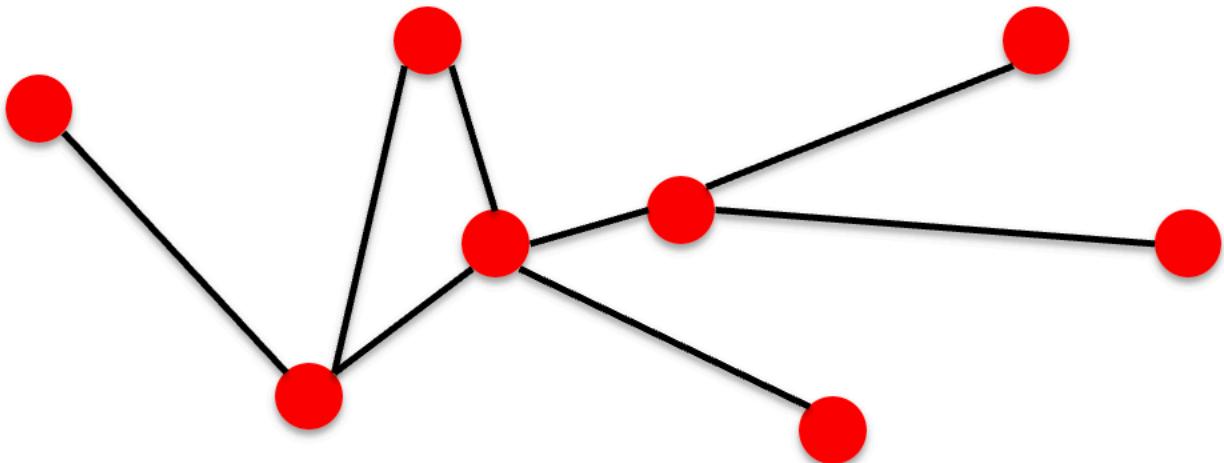


Immagini: Network Science, A. Barabasi



Reti

- Definizioni di base



- **components:** nodes, vertices

N

- **interactions:** links, edges

L

- **system:** network, graph

(N,L)

Network e Grafi

Le reti spesso si riferiscono a sistemi reali refers to real systems.

- www e internet
- social network
- Reti di protein e metaboliche

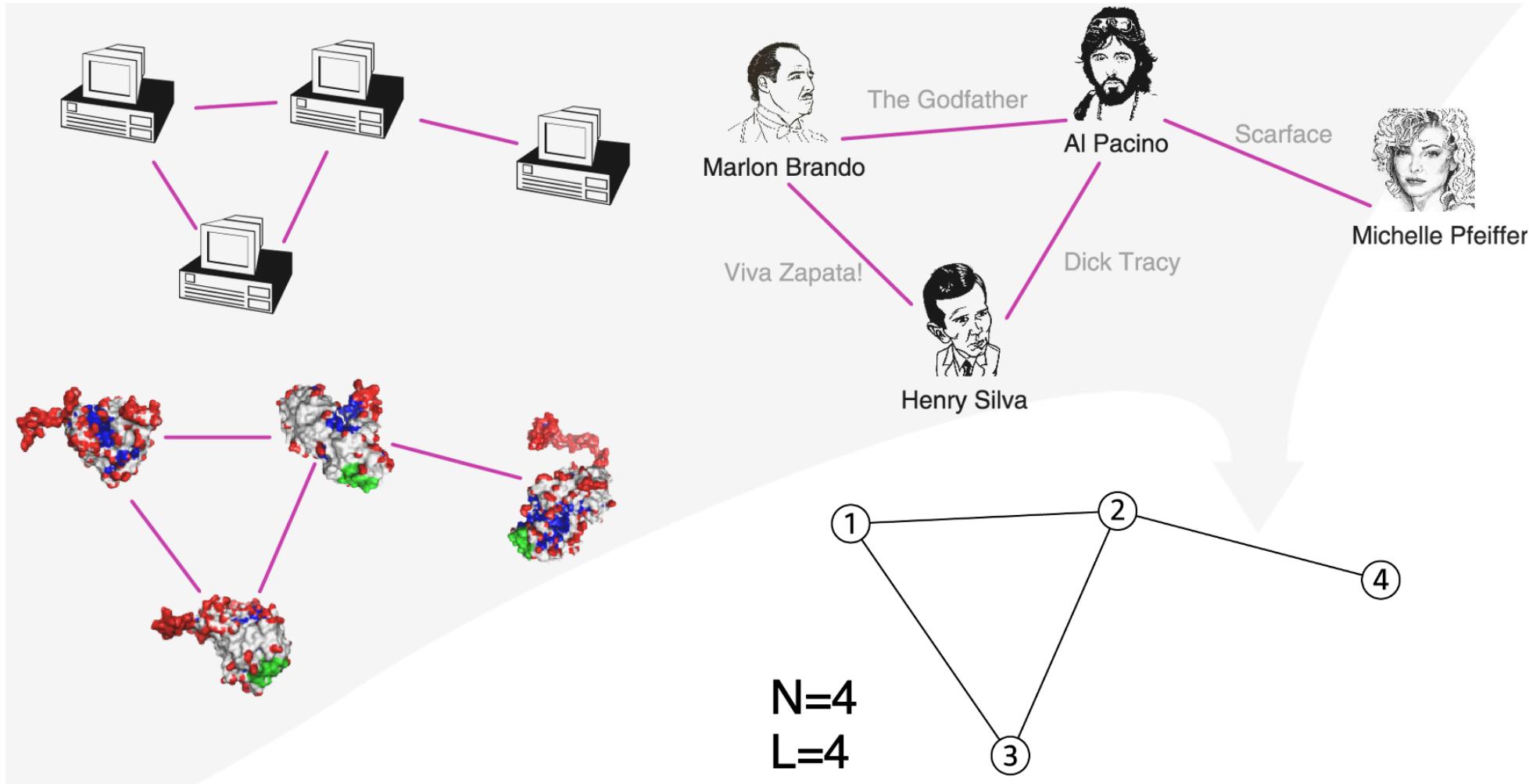
Termini: (Network, node, link)

I grafi sono la rappresentazione matematica di una rete.

- web graph

Termini: (Graph, vertex, edge)

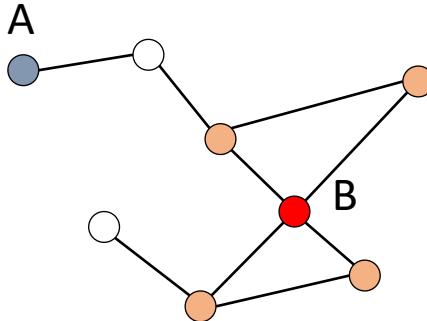
Uno stesso linguaggio



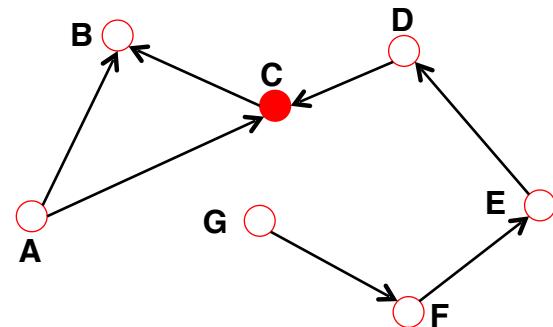
Definizioni di base

Grado di un nodo: il numero di links connessi a tale nodo.

Indiretta



Diretta

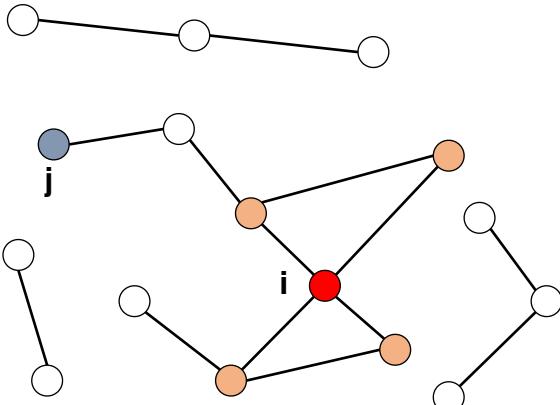


Nelle reti dirette definisco l'**in-degree** e l' **out-degree**. Il grado (totale) sarà la somma di in e out degrees.



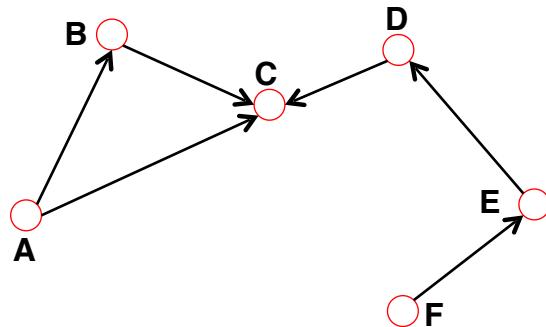
Definizioni di base

Indiretta



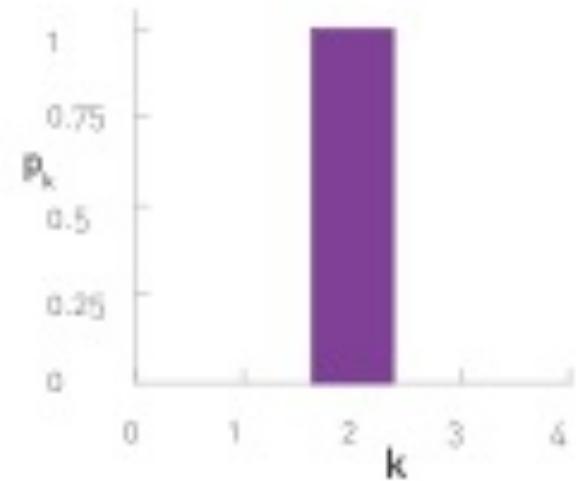
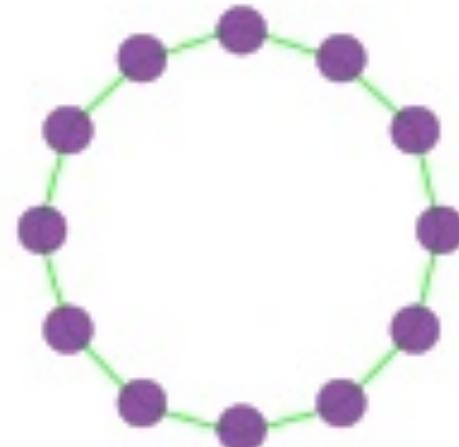
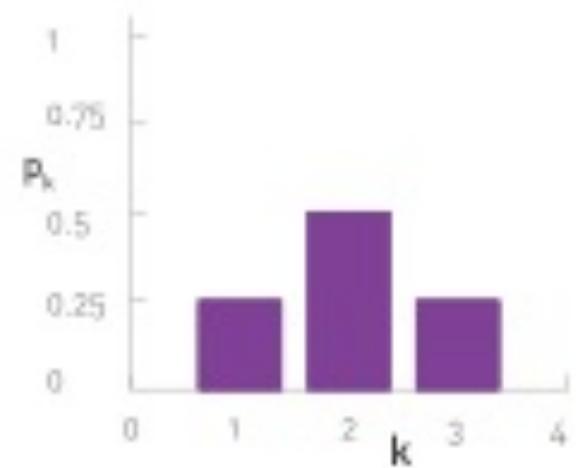
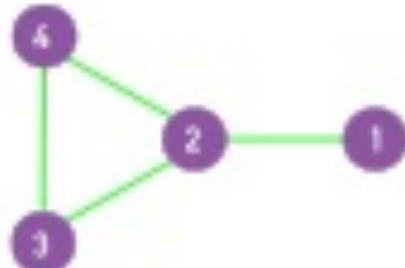
N – the number of nodes in the graph

Diretta



Distribuzione di Grado

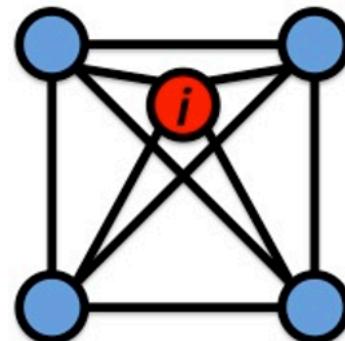
- **Degree Distribution:** in ascissa il grado, in ordinata il numero di nodi con quel grado
- Normalizzando per il numero di nodi ottengo una distribuzione di probabilità



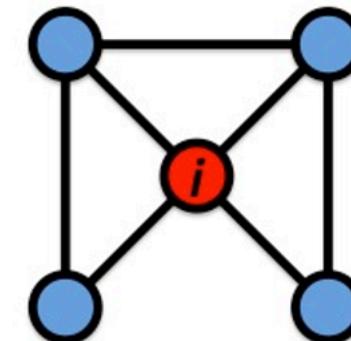
Coefficiente di Cluster

- Fornisce un'indicazione di quanto sono legati i **vicini** del nodo

$$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

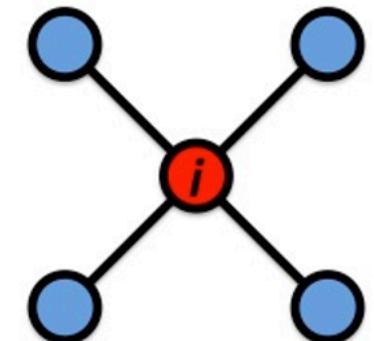


$$C_i = 1$$



$$C_i = 1/2$$

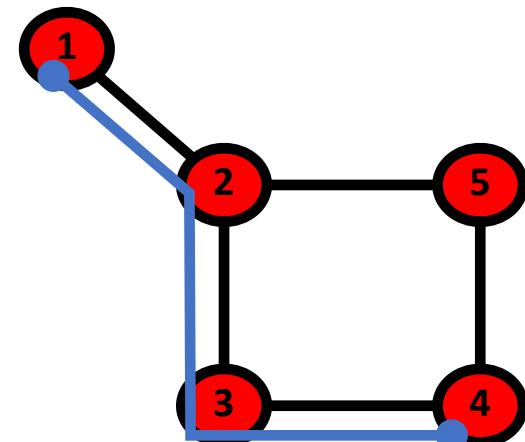
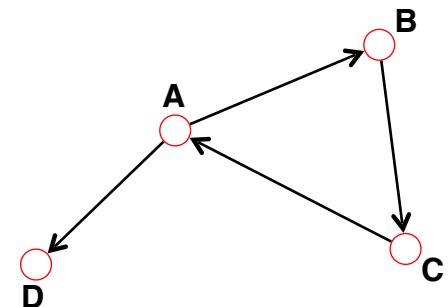
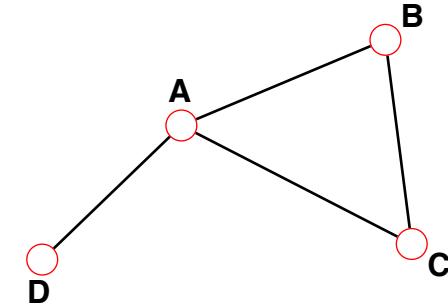
$$(2*3) / (4*3)$$



$$C_i = 0$$

Distanze nelle reti

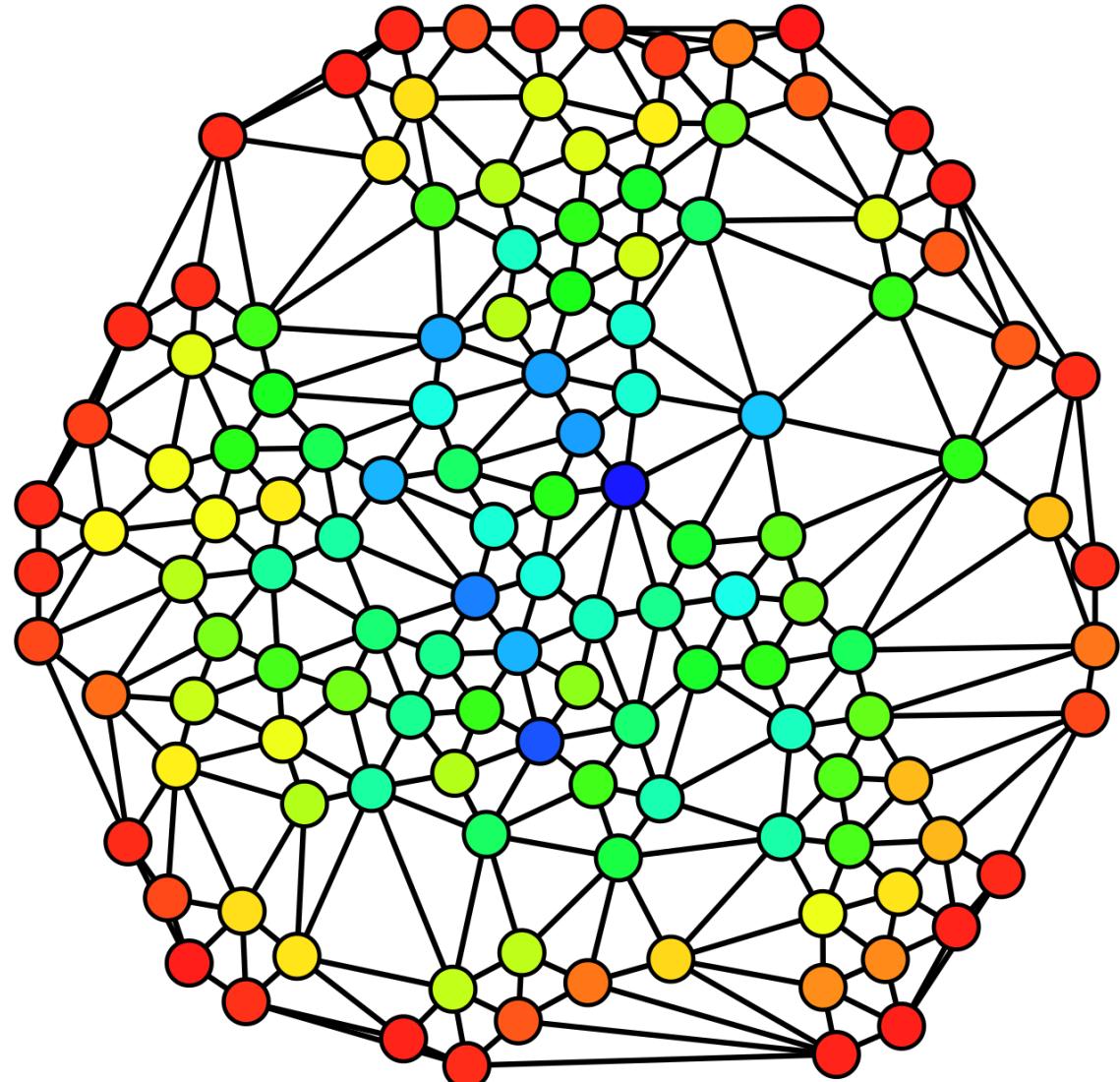
- La **distanza** tra due nodi è definita come il numero di links lungo il cammino più breve che li unisce. Se due nodi sono scollegati, la distanza è infinita.
- Nei grafi **diretti**, il cammino segue il verso della freccia. Quindi la distanza tra due nodi A e B può essere diversa dalla distanza tra B e A
- Il **diametro** è la distanza massima tra due nodi qualsiasi del grafo



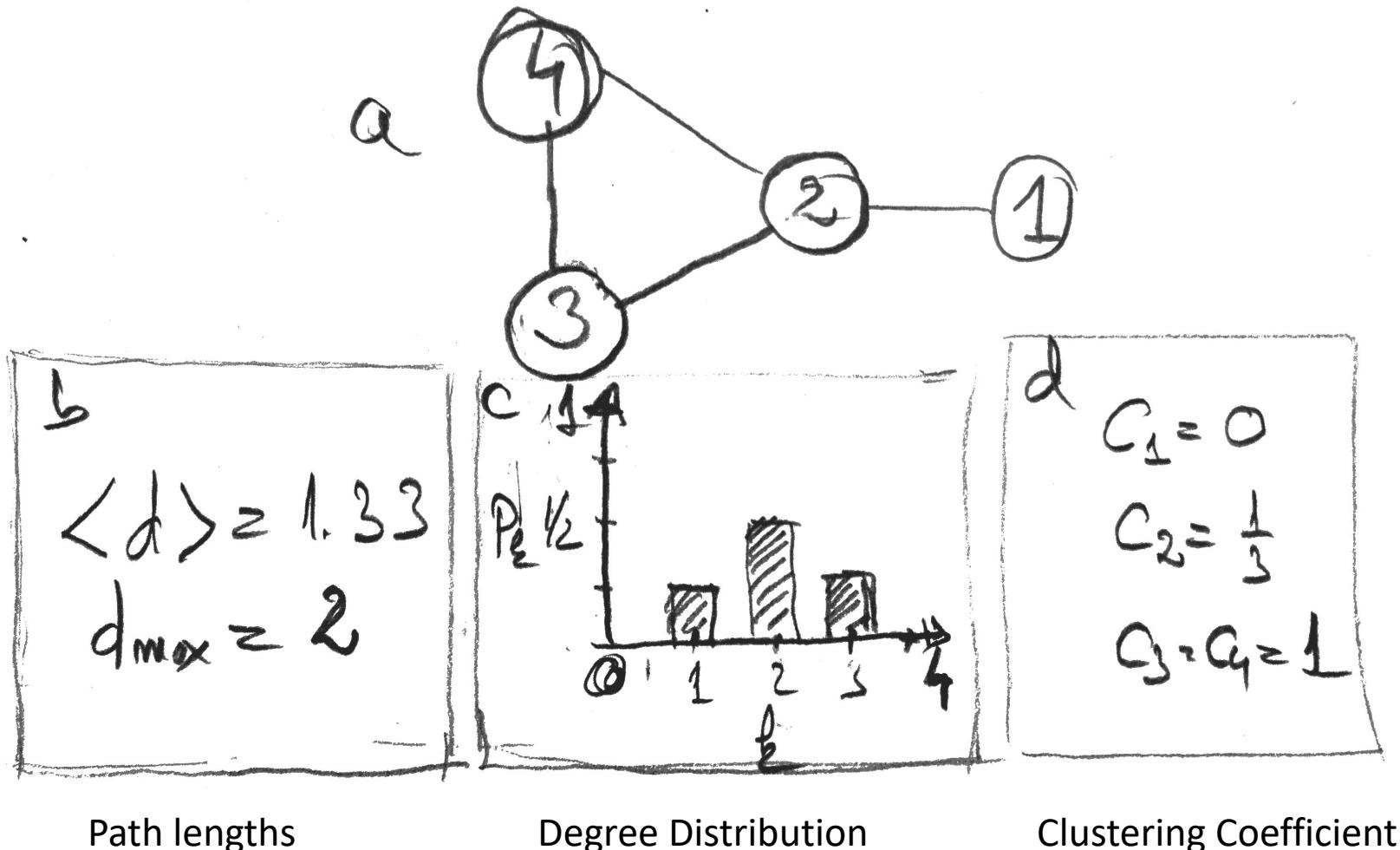
Betweenness centrality

- Un modo per indicare quanto un nodo è **centrale** nella rete

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$



Tre quantità fondamentali



Path lengths

Degree Distribution

Clustering Coefficient

Network Science AND ML

- Estraggo le **features** dalla rete
- Utilizzo Support Vector Machine

