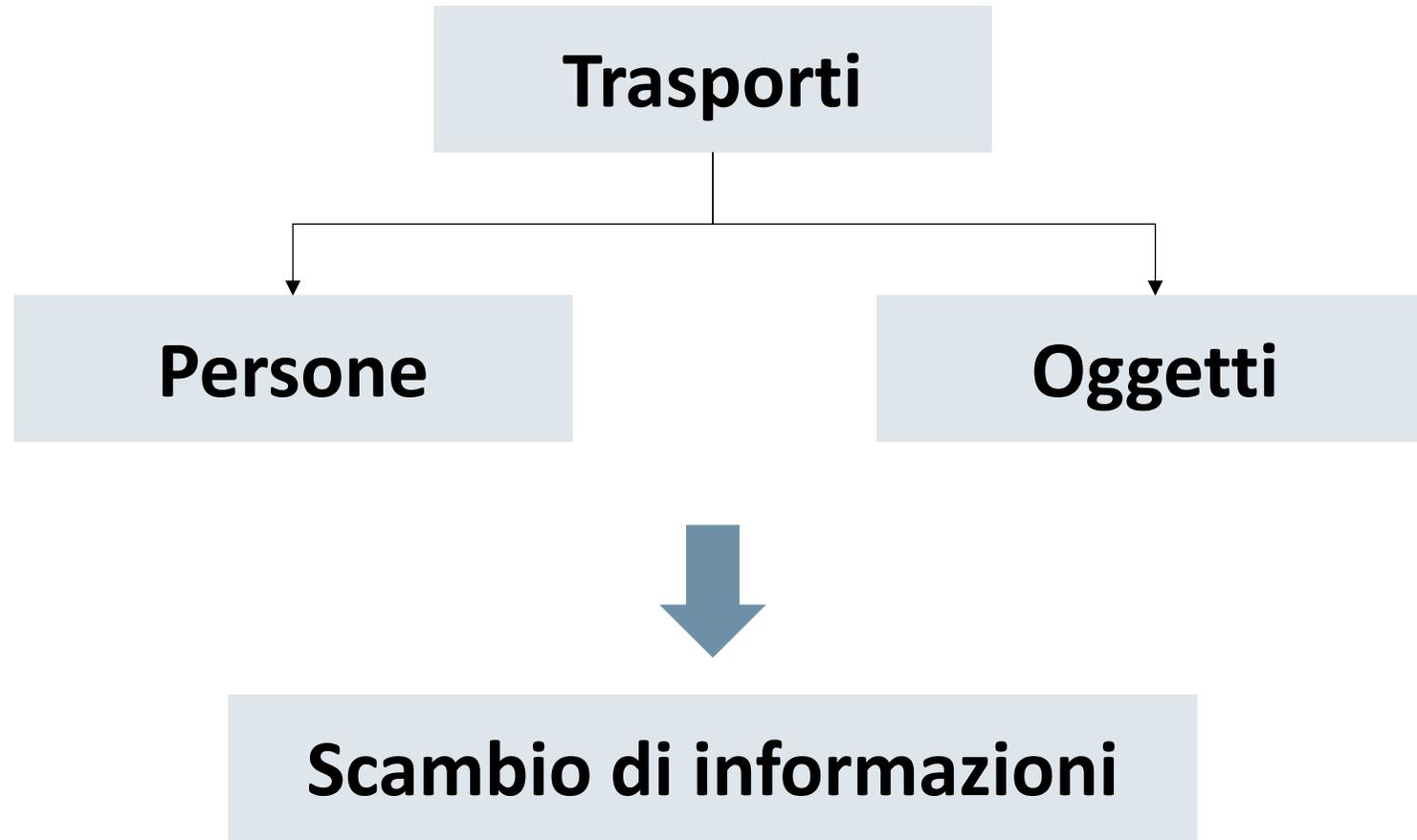


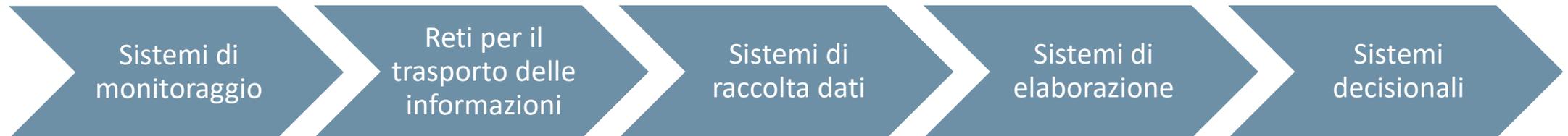
# Introduzione all'Intelligenza Artificiale (IA)

# Sistemi informatici per i Trasporti



# Sistemi informatici per i Trasporti

## Sistemi informatici per lo scambio di informazioni



# Sistemi informatici per i Trasporti

Es. logistica: una azienda di trasporto vuole conoscere in tempo reale lo stato di avanzamento delle spedizioni per ottimizzare la propria attività riducendo i costi.

Ha bisogno di:

- Sistemi automatici di rilevamento della merce.
- Conoscere lo spostamento dei mezzi.
- Sistemi per la memorizzazione e la trasmissione delle informazioni e monitoraggio real-time dei processi.
- Sistemi di elaborazione dei dati raccolti per ottimizzazione percorsi e carichi.

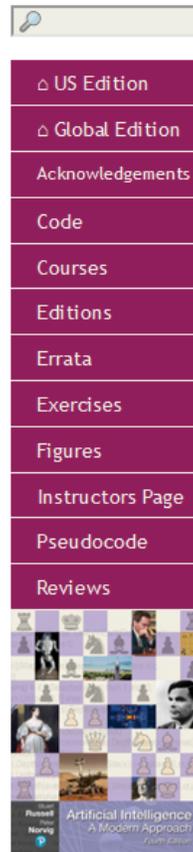
# Sistemi informatici per i Trasporti

Es. guida autonoma: una macchina in grado di guidare autonomamente in qualsiasi contesto capace di sostituire completamente l'attività dell'uomo.

Ha bisogno di:

- Sistemi di posizionamento, sistemi automatici di emergenza e geolocalizzazione.
- Connettività.
- Sistema di scambio di informazioni tra autovetture.
- Sistemi di raccolta dati per ottimizzazione percorsi e ricarica elettrica.

# Disclaimer (1)



## Artificial Intelligence: A Modern Approach, 4th US ed.

by Stuart Russell and Peter Norvig

The [authoritative, most-used](#) AI textbook, adopted by over [1500](#) schools.

Table of Contents for the US Edition (or see the [Global Edition](#))

[Preface \(pdf\)](#); [Contents with subsections](#)

### I Artificial Intelligence

- 1 Introduction ... 1
- 2 Intelligent Agents ... 36

### II Problem-solving

- 3 Solving Problems by Searching ... 63
- 4 Search in Complex Environments ... 110
- 5 Adversarial Search and Games ... 146
- 6 Constraint Satisfaction Problems ... 180

### III Knowledge, reasoning, and planning

- 7 Logical Agents ... 208
- 8 First-Order Logic ... 251
- 9 Inference in First-Order Logic ... 280
- 10 Knowledge Representation ... 314
- 11 Automated Planning ... 344

### IV Uncertain knowledge and reasoning

- 12 Quantifying Uncertainty ... 385
- 13 Probabilistic Reasoning ... 412
- 14 Probabilistic Reasoning over Time ... 461
- 15 Probabilistic Programming ... 500
- 16 Making Simple Decisions ... 528
- 17 Making Complex Decisions ... 562
- 18 Multiagent Decision Making ... 599

### V Machine Learning

- 19 Learning from Examples ... 651
- 20 Learning Probabilistic Models ... 721
- 21 Deep Learning ... 750
- 22 Reinforcement Learning ... 789

### VI Communicating, perceiving, and acting

- 23 Natural Language Processing ... 823
- 24 Deep Learning for Natural Language Processing ... 856
- 25 Computer Vision ... 881
- 26 Robotics ... 925

### VII Conclusions

- 27 Philosophy, Ethics, and Safety of AI ... 981
- 28 The Future of AI ... 1012
- Appendix A: Mathematical Background ... 1023
- Appendix B: Notes on Languages and Algorithms ... 1030
- Bibliography ... 1033 ([pdf](#) and [bib data](#))
- Index ... 1069 ([pdf](#))

[Exercises \(website\)](#)

[Figures \(pdf\)](#)

[Code \(website\)](#); [Pseudocode \(pdf\)](#)

Covers: [US](#), [Global](#)

<http://aima.cs.berkeley.edu/>

# Perché l'Intelligenza Artificiale?

# Perché l'IA?

ANDREA INDIANO CULTURA 11.11.2022

## L'intelligenza artificiale crea videoclip musicali straordinari

I programmi producono disegni e quadri ideali per essere abbinati ai testi delle canzoni più famose. E i risultati sono sorprendenti.



Stairway to Heaven secondo l'intelligenza artificiale. YOUTUBE

<https://www.wired.it/article/intelligenza-artificiale-genera-videoclip-musicali-esempi/>

# Perché l'IA?

MENU | CERCA

la Repubblica

ABBONATI | GEDI SMILE | R | ACCEDI

Lettere

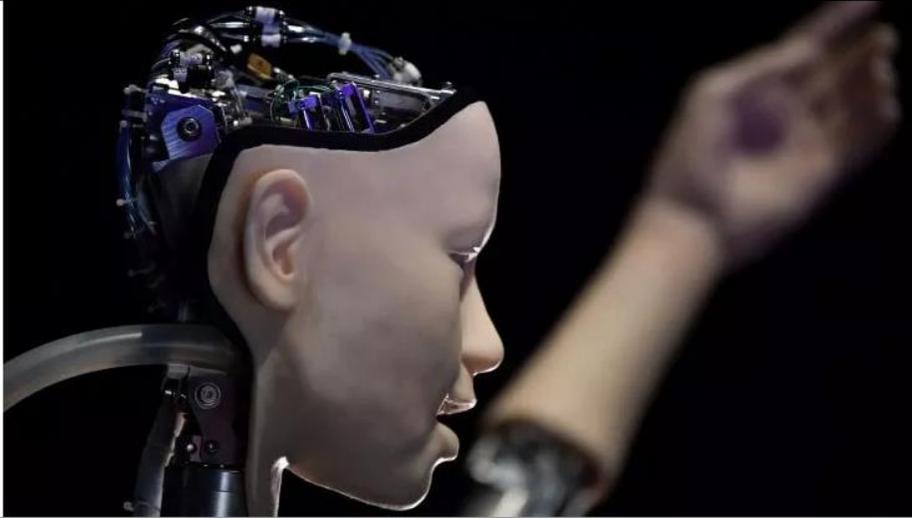


BIENNALE TECNOLOGIA

f t in e p

## Verso un futuro dominato dall'intelligenza artificiale

di Roberto Trotta\*



11 NOVEMBRE 2022 AGGIORNATO ALLE 10:40

3 MINUTI DI LETTURA

<https://www.repubblica.it/tecnologia/blog/lettere/2022/11/11/news/verso-un-futuro-dominato-dallintelligenza-artificiale-373933318/>

# Perché l'IA?



## L'intelligenza artificiale che batte Gordon Gekko. In borsa comandano gli algoritmi

/ di Luca Bianco 



*Un mix tra errore umano e trading algoritmico dietro al Flash-Crash in borsa di lunedì. La data scientist Donata Petrelli: "Per fare il trader oggi servono anche competenze informatiche"*

06 Maggio 2022 alle 11:21

Segui i temi

digitale 

<https://www.huffingtonpost.it/dossier/fintech/2022/05/06/news/l-intelligenza-artificiale-che-batte-gordon-gekko-in-borsa-comandano-gli-algoritmi-9334041/>

# Perché l'IA?

il Giornale.it

IN EVIDENZA [La guerra in Ucraina](#) [Coronavirus](#) [Il nodo Giustizia](#) [Over](#) [Il nuovo ilGiornale.it](#)

## Medicina, la rivoluzione dell'Intelligenza Artificiale

4 Maggio 2022 - 11:55

L'uso dell'Intelligenza Artificiale in campo medico avrà un impatto clinico e finanziario da non sottovalutare. Ma non mancano gli ostacoli da superare

 [Gabriele Laganà](#) 0 

A digital illustration of a human brain in shades of blue and cyan. The brain is surrounded by a complex network of glowing lines and nodes, resembling a neural network or a data processing system. The background is dark with some light effects, giving it a futuristic, high-tech appearance.

<https://www.ilgiornale.it/news/salute/medicina-rivoluzione-dell-intelligenza-artificiale-2031172.html>

# Perché l'IA?



The screenshot shows the Rai News website interface. At the top, there are navigation tabs for 'Cronaca', 'Esteri', 'Politica', 'Televideo', and 'Sport'. A search bar and weather widget for Rome are also visible. Below the navigation, there are social media icons and a date stamp '14 Aprile 2022'. The main content area features a sub-header 'GUERRE E CONFLITTI' and the title 'Laboratorio Ucraina'. The main headline reads 'Droni, reti neurali ed armi letali autonome: come l'Intelligenza Artificiale sta cambiando la guerra'. A sub-headline provides context: 'Non siamo nella saga di Halo, eppure l'IA milita nel campo di battaglia su molti livelli. Su quello dell'intelligence, ad esempio, con sistemi simili a quelli offerti da Primer, un'azienda americana in grado di captare messaggi del nemico'. The author's name, 'Enrico Santus', is listed at the bottom of the article header.



<https://www.rainews.it/articoli/2022/04/laboratorio-ucraina-droni-reti-neurali-ed-armi-letali-autonome-come-lintelligenza-artificiale-sta-cambiando-la-guerra-635eb8a0-8e75-46ee-9f15-8c46285f5dfd.html>

# Applicazioni (quotidiane o quasi)

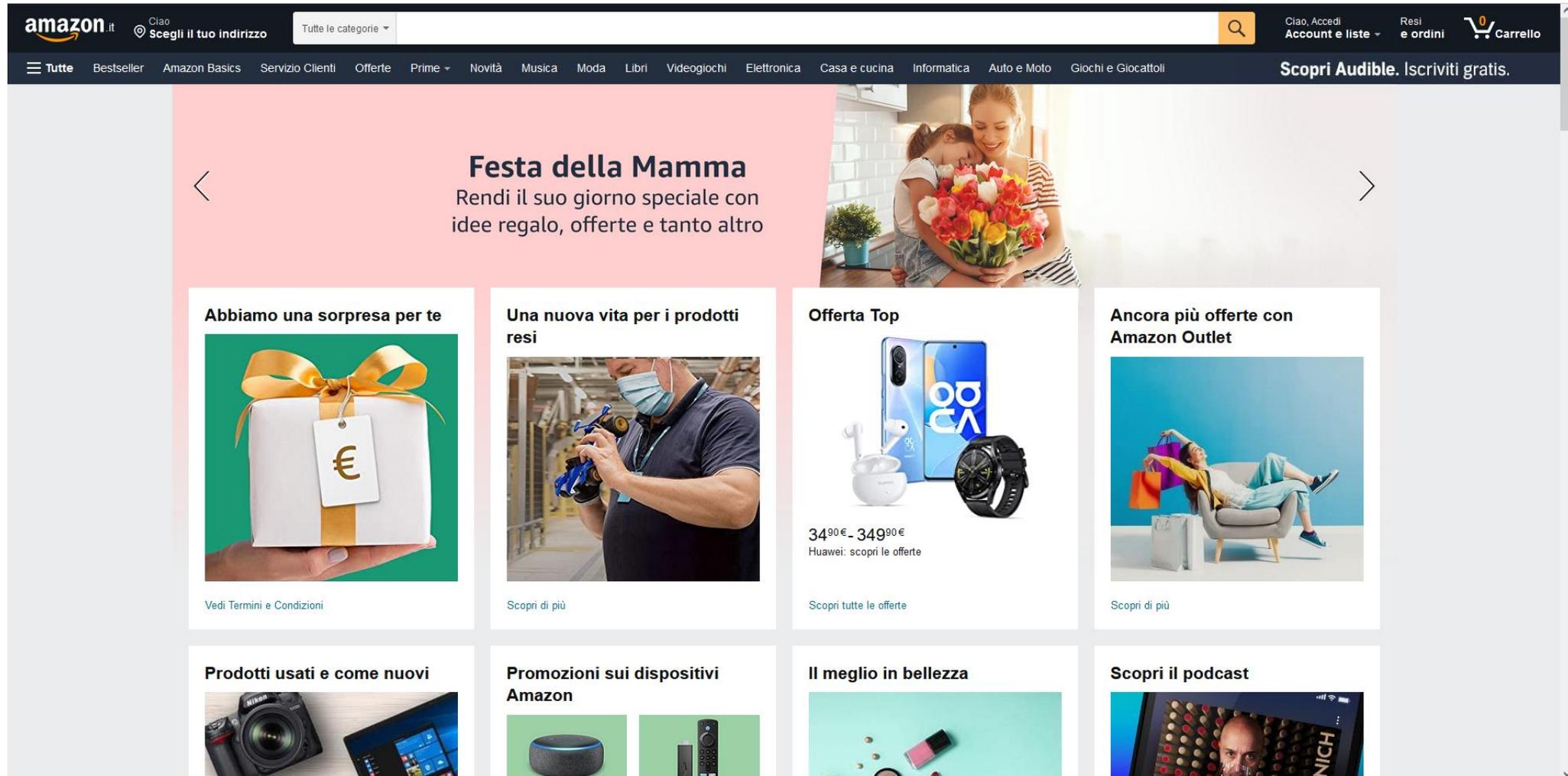
# Recommender Systems



**Recommender system:** sistema di filtraggio dei contenuti che genera raccomandazioni personalizzate, specifiche per l'utente

Esempi?

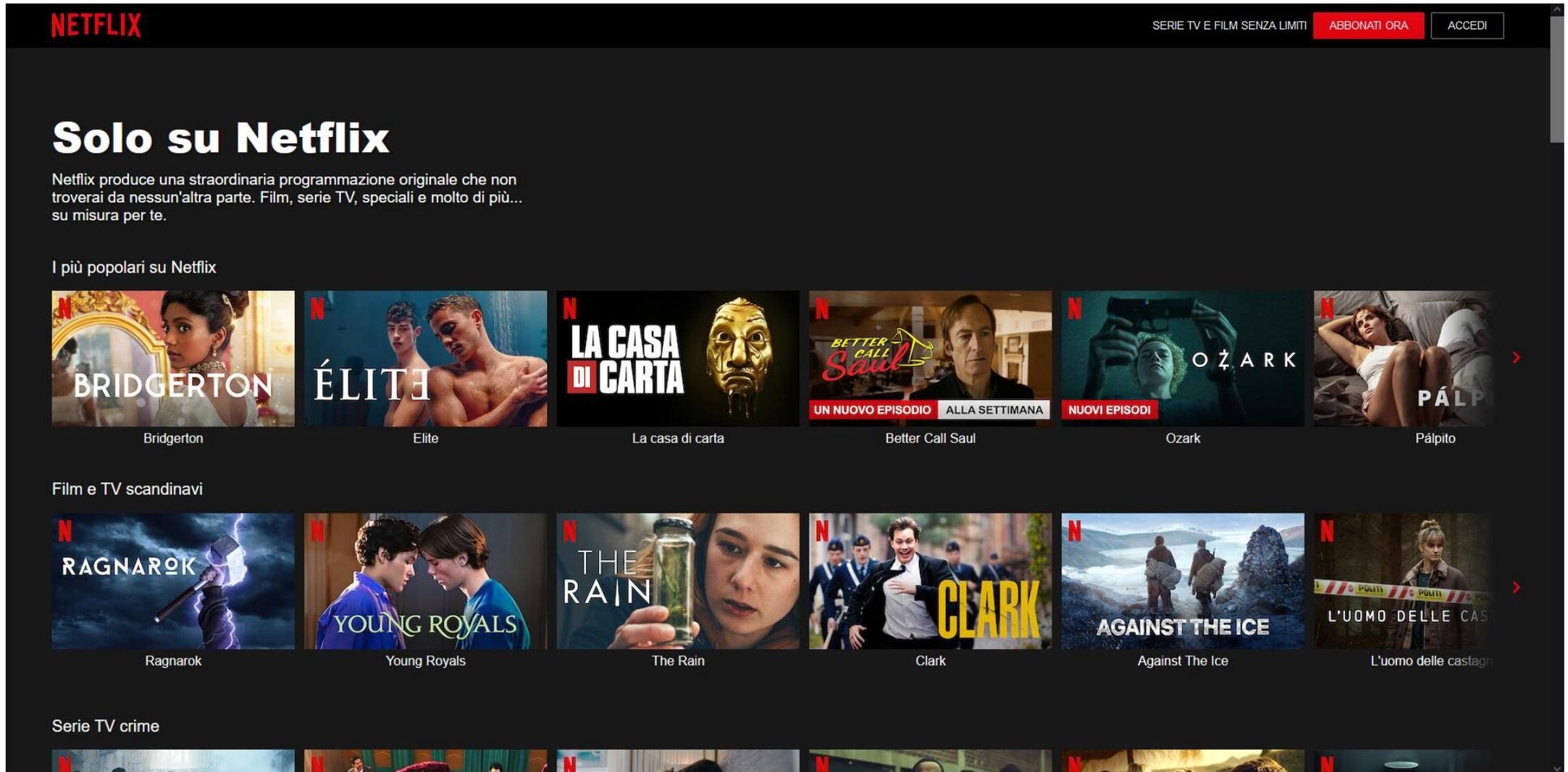
# Recommender Systems



The screenshot shows the Amazon Italy homepage with a navigation bar at the top. The main banner is for 'Festa della Mamma' (Mother's Day), with the text 'Rendi il suo giorno speciale con idee regalo, offerte e tanto altro'. Below the banner are several promotional tiles:

- Abbiamo una sorpresa per te**: A white gift box with a gold ribbon and a tag with a Euro symbol (€). Below the image is the link 'Vedi Termini e Condizioni'.
- Una nuova vita per i prodotti resi**: A man in a blue shirt and mask working with a drone. Below the image is the link 'Scopri di più'.
- Offerta Top**: A Huawei smartphone, a smartwatch, and a smart speaker. The price is shown as 34<sup>90</sup>€ - 349<sup>90</sup>€. Below the image is the link 'Scopri tutte le offerte'.
- Ancora più offerte con Amazon Outlet**: A woman sitting on a sofa with shopping bags. Below the image is the link 'Scopri di più'.
- Prodotti usati e come nuovi**: A Nikon camera and a tablet.
- Promozioni sui dispositivi Amazon**: An Amazon Echo smart speaker and a remote control.
- Il meglio in bellezza**: A collection of beauty products.
- Scopri il podcast**: A man's face on a screen with the word 'NICH' visible.

# Recommender Systems



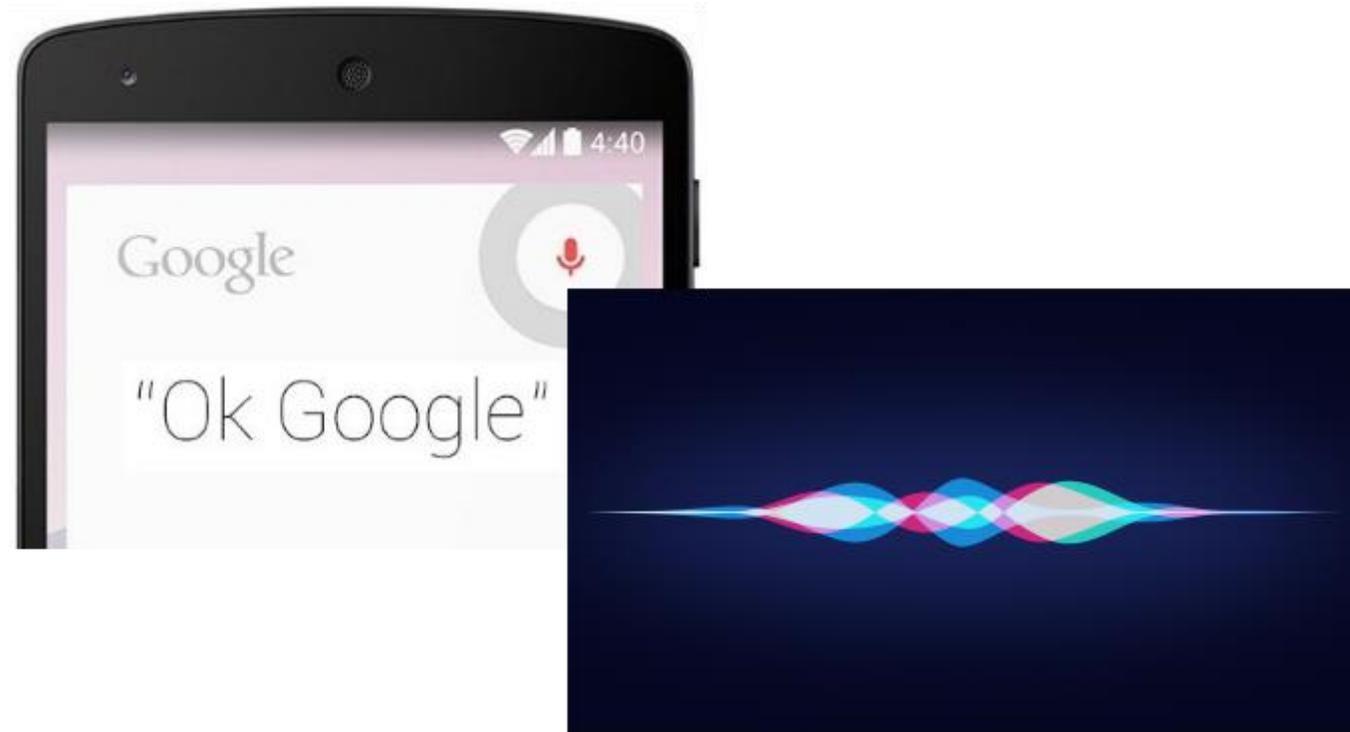
The image shows a screenshot of the Netflix website interface. At the top left is the 'NETFLIX' logo. On the top right, there are links for 'SERIE TV E FILM SENZA LIMITI', 'ABBONATI ORA', and 'ACCEDI'. The main heading is 'Solo su Netflix', followed by a paragraph: 'Netflix produce una straordinaria programmazione originale che non troverai da nessun'altra parte. Film, serie TV, speciali e molto di più... su misura per te.' Below this is the section 'I più popolari su Netflix' with a row of six featured titles: Bridgerton, Elite, La casa di carta, Better Call Saul, Ozark, and Pálpito. The next section is 'Film e TV scandinavi' with a row of six titles: Ragnarok, Young Royals, The Rain, Clark, Against The Ice, and L'uomo delle castagne. The final section is 'Serie TV crime' with a row of six titles partially visible.

# Riconoscimento facciale

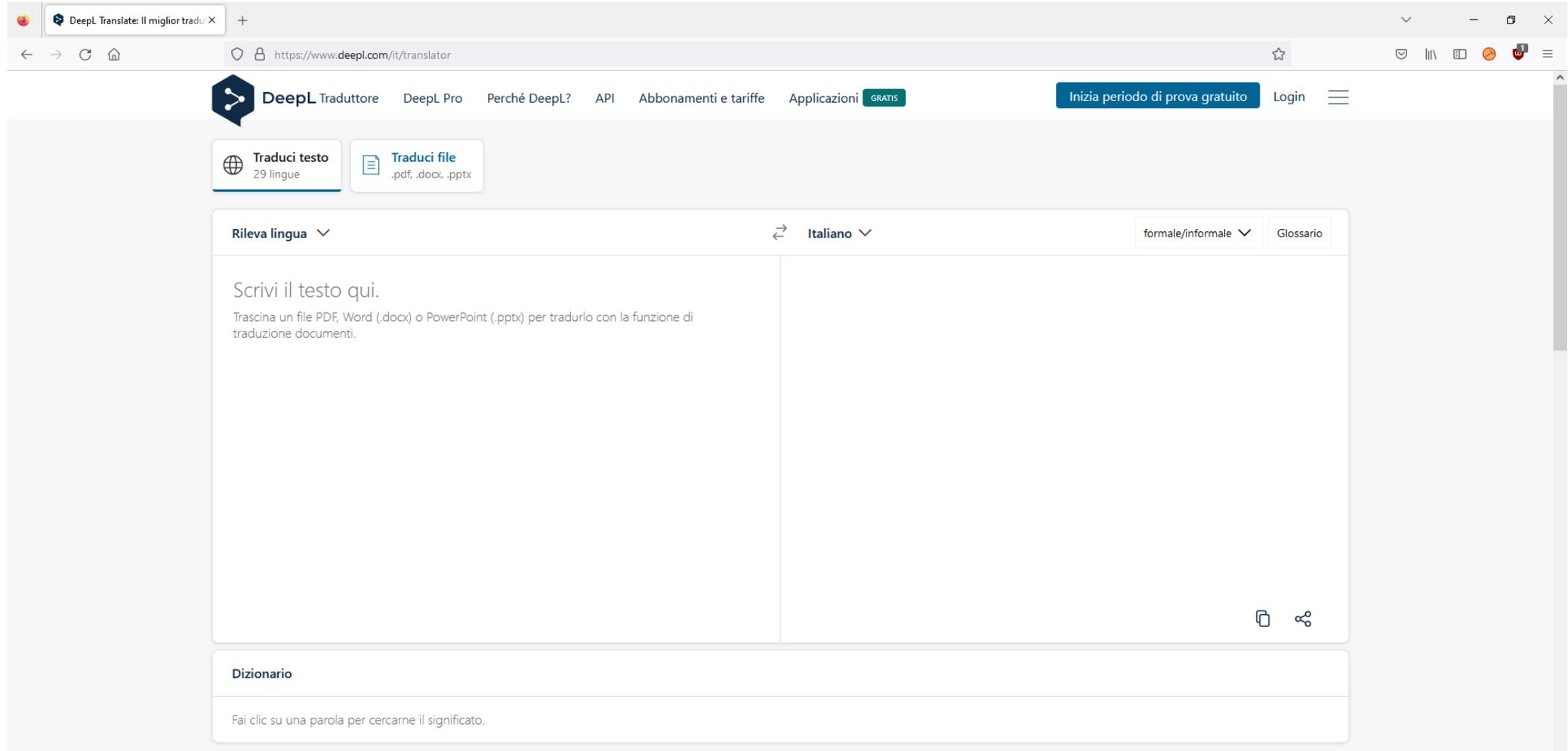


Immagine da <https://support.apple.com/it-it/HT208109>.

# Speech Recognition



# Traduzione automatica



The screenshot shows the DeepL Translator website interface. At the top, there is a navigation bar with the DeepL logo and links for 'Traduttore', 'DeepL Pro', 'Perché DeepL?', 'API', 'Abbonamenti e tariffe', and 'Applicazioni' (with a 'GRATIS' badge). A blue button for 'Inizia periodo di prova gratuito' and a 'Login' link are also present. Below the navigation bar, there are two main options: 'Traduci testo' (29 lingue) and 'Traduci file' (.pdf, .docx, .pptx). The main content area features a language selection dropdown set to 'Italiano', a 'formale/informale' dropdown, and a 'Glossario' link. The central text input area contains the instruction 'Scrivi il testo qui.' and a note about translating documents. At the bottom, there is a 'Dizionario' section with the instruction 'Fai clic su una parola per cercarne il significato.'

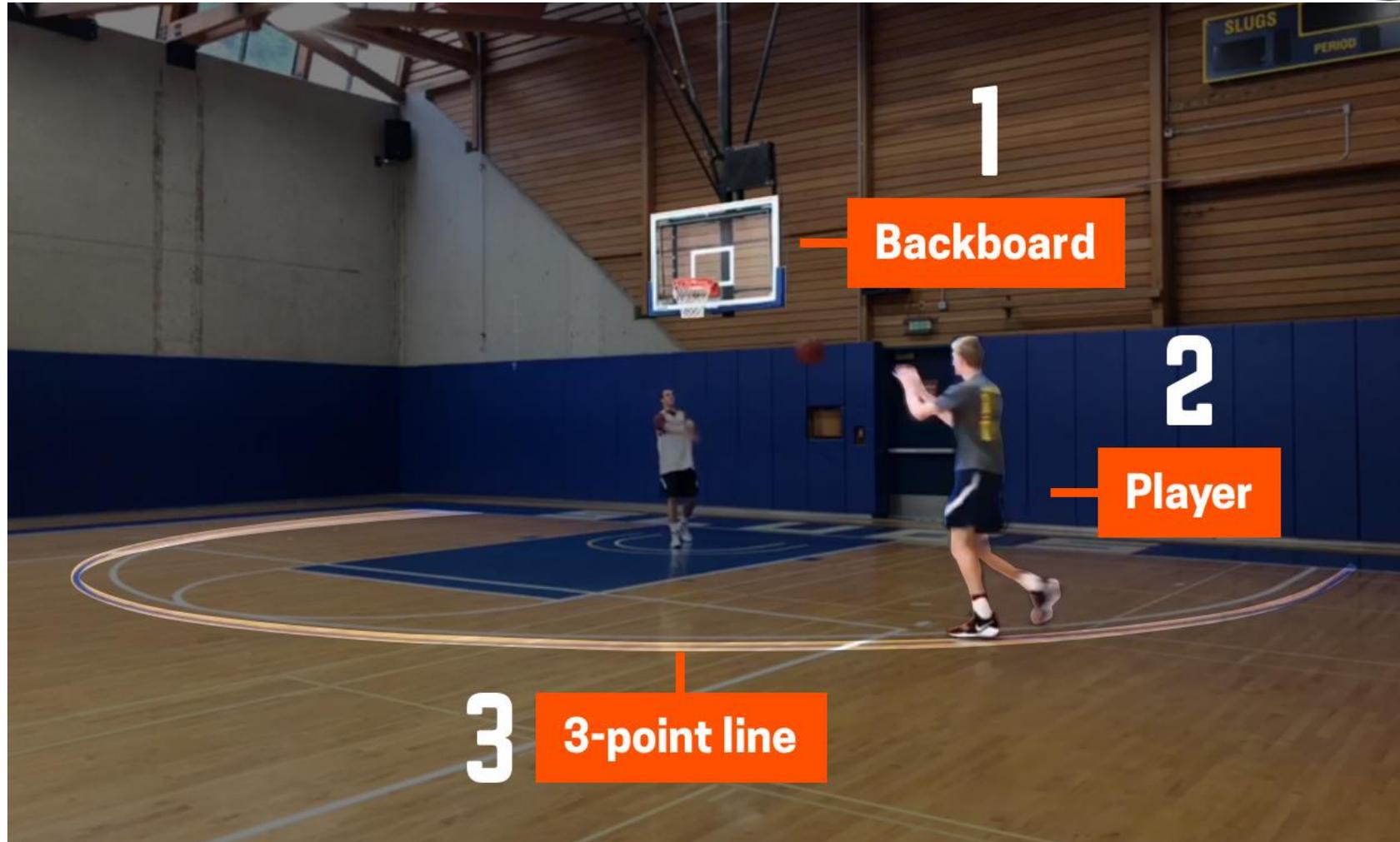
<https://www.deepl.com/it/translator>

# Guida autonoma: lane and vehicle detection



Huval, Brody, et al. *An empirical evaluation of deep learning on highway driving*. arXiv preprint arXiv:1504.01716 (2015) <https://arxiv.org/abs/1504.01716>.

# Sport



<https://www.homecourt.ai/>

# Politica

Può l'IA favorire una fazione durante la campagna elettorale?



# Cambridge Analytica (~2016)



- Azienda specializzata nella fornitura di servizi per pubblicità mirata;
- Un'altra azienda, "thisisyourdigitallife", raccoglie (legittimamente) i dati di milioni di utenti;
- Compra i dati da "thisisyourdigitallife" (qui avviene l'illecito);
- Usa quei dati per fornire pubblicità mirata in favore di uno dei candidati alle presidenziali 2016.

## Altre applicazioni

- marketing e customer care (es. sentiment analysis)
- medicina e biologia (bioinformatica)
- riconoscimento di immagini
- ...

Alcuni esempi: <https://machinelearningmastery.com/inspirational-applications-deep-learning/>

## Altre applicazioni

«Facebook moderators under contract are “bombarde” with thousands of videos, images and livestreamed broadcasts of child sexual abuse, rape, torture, bestiality, beheadings, suicide and murder”, the lawsuit said.

“Facebook is ignoring its duty to provide a safe workplace and instead creating a revolving door of contractors who are irreparably traumatized by what they witnessed on the job,” Korey Nelson, a lawyer for former Facebook contract employee Selena Scola, said in a statement Monday.»

Da <https://www.theguardian.com/technology/2018/sep/24/facebook-moderators-mental-trauma-lawsuit>.

# Cos'è l'Intelligenza Artificiale?

# Cos'è l'IA?



“The science and engineering of making intelligent machines”

---

John McCarthy (1956)

# Cos'è l'IA?



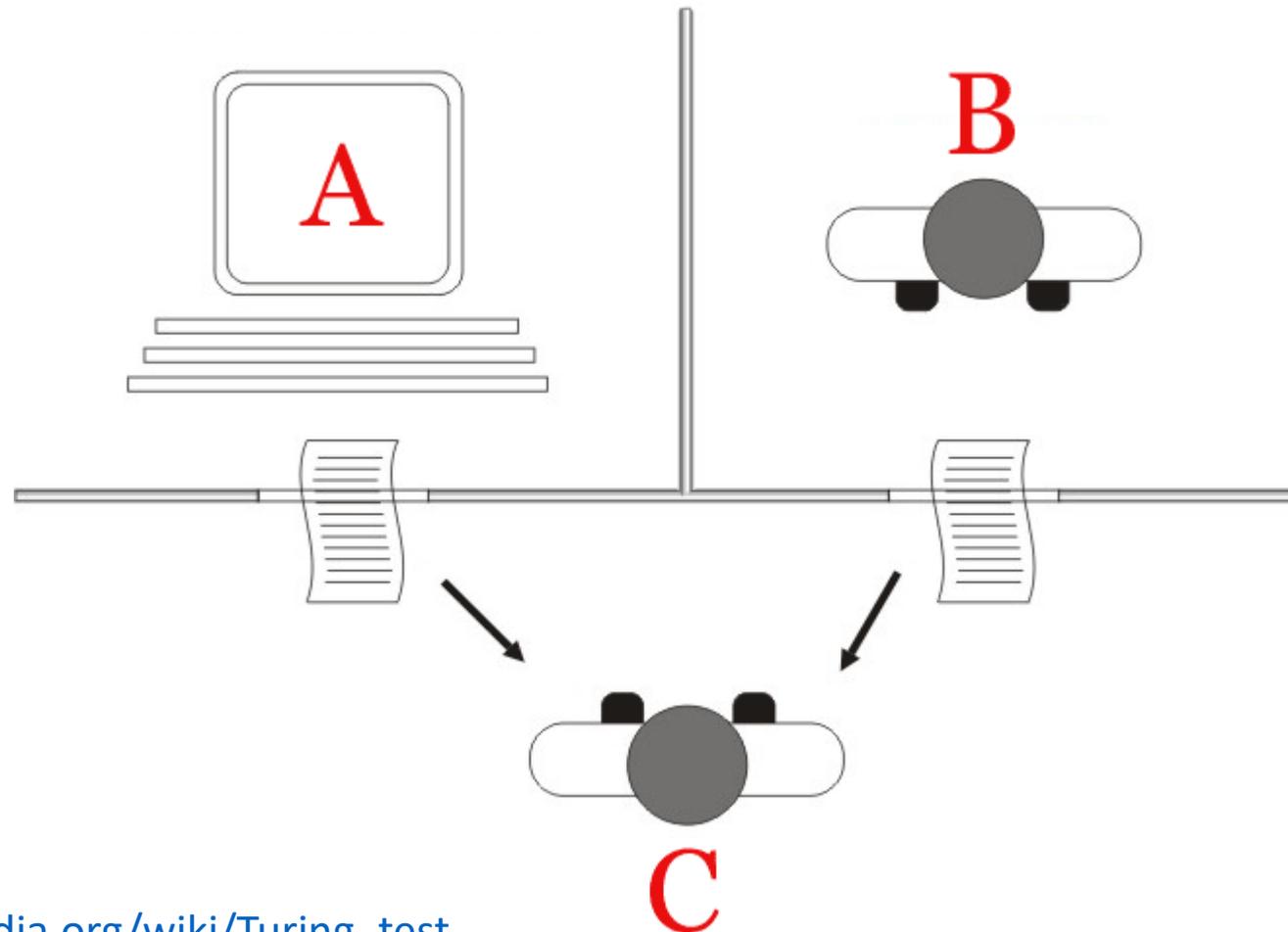
Macchine (sistemi) che pensano/agiscono **come umani?**

OPPURE

Macchine (sistemi) che pensano/agiscono **razionalmente?**

# Macchine che agiscono **come umani**

## Test di Turing



Da [https://en.wikipedia.org/wiki/Turing\\_test](https://en.wikipedia.org/wiki/Turing_test)

# Macchine che pensano **come umani**



**IA Forte/IA generale:** La capacità di una macchina di comprendere o apprendere ogni tipo di compito «intellettuale» che un essere umano è in grado di comprendere o apprendere. In altre parole, una macchina che ha coscienza di sé.

**IA Debole/IA ristretta:** La capacità di una macchina di implementare una parte di «intelligenza» per eseguire un compito preciso. L'intelligenza artificiale con cui abbiamo a che fare quotidianamente è di questo tipo (es. assistente vocale, riconoscimento immagini... tutte le applicazioni viste in precedenza!)

# Chinese room – John Searle (1980)

Esperimento mentale “contro” IA forte

- ipotesi: un computer che si comporta come se comprendesse il cinese;
- ipotesi: il computer passa il test di Turing;
- il computer comprende davvero il cinese o simula di comprenderlo?

Un essere umano senza alcuna conoscenza del cinese, messo in una stanza con un manuale con le opportune istruzioni, potrebbe fare la stessa cosa, senza comprendere davvero ciò che gli viene chiesto

# Macchine che pensano/agiscono **razionalmente**



**unIMC**

**Comportamento razionale:** eseguire un'azione il cui risultato atteso massimizza un certo obiettivo (perlomeno secondo l'informazione disponibile al momento dell'esecuzione).

Un agente intelligente *delibera* un'azione da eseguire (o nuovi fatti da aggiungere alla propria base di conoscenza) mediante

- Deduzione
- Abduzione
- Induzione

# Deduzione

Il processo logico nel quale, date certe premesse e certe regole che ne garantiscono la correttezza, una conclusione consegue come logicamente necessaria. E' lo schema di ragionamento del sillogismo.

## Sillogismo

Ogni uomo è mortale  
Socrate è un uomo

ALLORA

Socrate è mortale

$\forall(x)(influenza(x) \rightarrow febbre(x))$

$\Rightarrow$

$febbre(Socrate)$

$influenza(Socrate)$

# Abduzione

Lo schema di ragionamento tipico della diagnosi

$$\forall(x)(influenza(x) \rightarrow febbre(x))$$
$$febbre(Socrate)$$

---

$$influenza(Socrate)$$

# Induzione

L'apprendimento dalle osservazioni

*influenza(Socrate)*

*febbre(Socrate)*

*influenza(Aristotele)*

*febbre(Aristotele)*

*influenza(Euclide)*

*febbre(Euclide)*

...

---

$$\forall(x)(influenza(x) \rightarrow febbre(x))$$

# IA simbolica vs IA statistica



**IA simbolica:** rappresentazione logico-matematica del mondo. Esempi: ricerca nello spazio degli stati, problemi di ottimizzazione, ecc.

**IA statistica:** apprendimento automatico, senza rappresentazione esplicita delle regole/funzioni apprese

# IA simbolica

Decisioni spiegabili per definizione

Non è semplice trovare la rappresentazione opportuna

Non tutto si può computare tempestivamente...

# IA statistica

Apprendimento autonomo di relazioni sconosciute

Numero di esempi rappresentativo della realtà

Scatola nera/decisioni non spiegabili

## IA – Storicamente...

Forte, fin dai primi anni, in problemi descrivibili come un set di formule logiche/matematiche

- theorem proving e motori di inferenza;
- sistemi esperti;
- logic based knowledge representation systems;
- problem solving e ricerca;
- ...

Questo nonostante i noti problemi di intrattabilità computazionale e decidibilità

# IA – Storicamente (1997)



# IA – Storicamente

- Forte in problemi «intellettualmente complessi» per gli umani
- Debole in problemi «facili» per gli umani ma complessi da descrivere formalmente
  - riconoscimento del parlato
  - riconoscimento di oggetti in immagini
  - riconoscimento di volti
  - ...

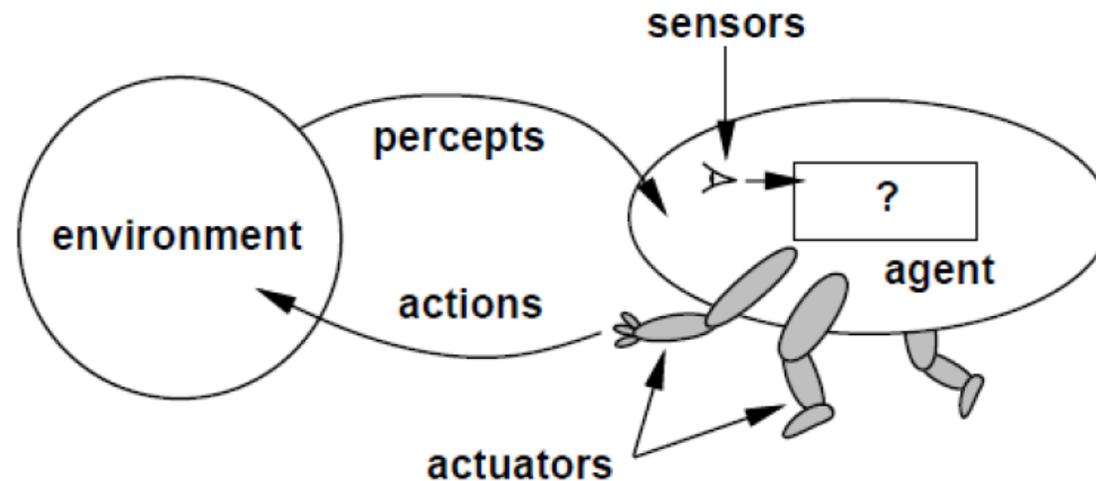
# IA vs «conoscenza informale»



Task apparentemente semplici come riconoscere oggetti richiedono esperienza, formata mediante apprendimento

# Altre definizioni – Agente Intelligente

“An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors.”



Da **S. Russell** and **P. Norvig**, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2009.

## Altre definizioni – Agente Intelligente



Può un «agente intelligente» esibire comportamenti che non siano preprogrammati, ma appresi dalle «Percezioni»?

Deve un «agente intelligente» esibire comportamenti che non siano preprogrammati, ma appresi dalle «Percezioni»?

# Machine Learning (ML – Apprendimento Automatico)

# Machine Learning

L'apprendimento automatico (machine learning) è l'insieme di metodi per permettere ai computer di apprendere autonomamente dagli esempi. Più precisamente l'apprendimento automatico studia e sviluppa algoritmi per apprendere dai dati, usarli per fare previsioni e scoprire nuova conoscenza.

Da: A. L. Samuel, "*Some studies in machine learning using the game of checkers*". IBM Journal of research and development, vol. 3 (3), pp. 210-229, 1959

# Machine Learning

L'apprendimento automatico (machine learning) è l'insieme di metodi per permettere ai computer di apprendere autonomamente dagli esempi. Più precisamente l'apprendimento automatico studia e sviluppa algoritmi per apprendere dai dati, usarli per fare previsioni e scoprire nuova conoscenza.



Percezione non solo per agire, ma anche per migliorare la capacità di agire in futuro.

# Problemi di machine Learning



- Classificazione
- Regressione
- Clustering
- Riduzione della dimensionalità

# Classificazione

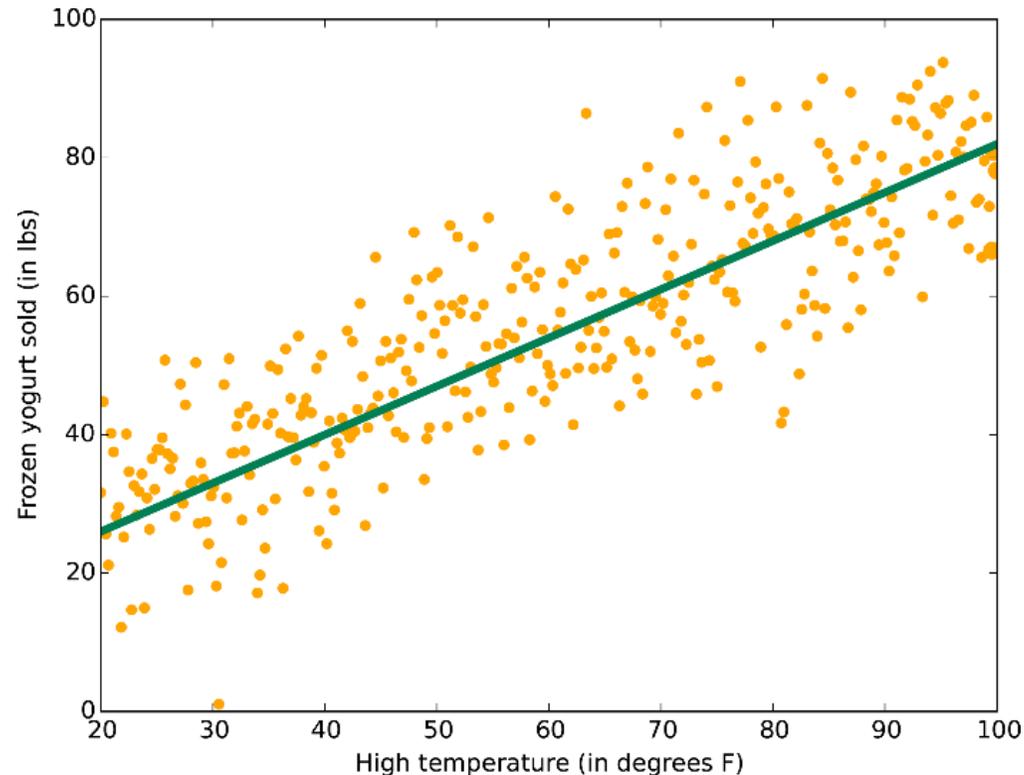
Associa una classe ai dati di input (riconoscimento). Una classe indica un insieme di oggetti aventi proprietà comuni, o, in altre parole, che seguono lo stesso pattern.

Esempi:

- a partire da un vettore di input che rappresenta un'immagine di un volto, riconoscere un soggetto
- riconoscere categorie di oggetti nelle immagini
- riconoscere usi propri e impropri di una carta di credito

# Regressione

Associa un valore ai dati di input. In altre parole, stima una funzione che si avvicini il più possibile ai dati di input. Può essere usata per fare «previsioni»



Da <https://docs.microsoft.com/it-it/azure/machine-learning/studio/algorithm-choice>

# Clustering

Suddivide in gruppi (cluster) i campioni di input con caratteristiche simili. Al contrario della classificazione, quindi, le classi non sono note a-priori.

Esempio:

- profilazione di utenti per il suggerimento di prodotti/acquisti (recommender system)

# Riduzione della dimensionalità

Mapping dei campioni di input in uno spazio a dimensionalità più bassa. In altre parole riduce la cardinalità di ogni campione di input.

Ovviamente c'è una perdita di informazione, ma:

- l'obiettivo è solitamente rendere trattabile un problema;
- mantiene le informazioni «importanti», scartando quelle superflue o ridondanti;
- utile anche per la visualizzazione di dati

# Machine learning «convenzionale»



- conoscenza di dominio per la scelta delle feature da estrarre
- dati «raw» da trasformare nello spazio delle feature
- dati trasformati in input per la classificazione o la scoperta di pattern

# Deep Learning (DL)

# Representation Learning



Può un «agente intelligente» apprendere autonomamente quali feature utilizzare per i problemi di machine learning, a partire da dati «raw»?

**Representation Learning:** insieme di tecniche per passare alle macchine dati «raw» e permettere di «apprendere» la rappresentazione migliore per i task di machine learning da eseguire

# Deep Learning

“Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics.”

Y. LeCun, Y. Bengio, and G. Hinton, “*Deep learning*,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015 (<https://www.nature.com/articles/nature14539>)

# Deep Learning – Timeline

- **1943:** McCulloch e Pitts introducono il primo modello logico-matematico di funzionamento dei neuroni del cervello umano
- **1958:** Roseblatt presenta il percettrone (“perceptron”), algoritmo di apprendimento automatico a due strati (ingresso/uscita)
- **1969:** Minsky e Paperts, nel loro libro “Perceptron”, analizzano alcuni limiti del percettrone
- **1986:** introduzione della backpropagation da parte di Rumelhart, Hinton e Williams.

# Deep Learning – Timeline



- **1989:** Lecun sviluppa reti con connessioni locali e pesi condivisi (Convolutional Neural Network) per il riconoscimento automatico dei caratteri scritti a mano
- **1998:** Lecun e Bengio ottengono un errore inferiore all'1% sul dataset MNIST (30.000 modelli di cifre scritte a mano) con la Lenet-5, una Convolutional Neural Network composta da 7 strati
- **2005:** D. Steinkraus, I. Buck and P. Y. Simard, "*Using GPUs for machine learning algorithms*," Eighth International Conference on Document Analysis and Recognition (ICDAR'05), 2005, pp. 1115-1120 Vol. 2. doi: 10.1109/ICDAR.2005.2513

# Deep Learning – Timeline



- **2012:** nell'annuale "Imagenet Large Scale Visual Recognition Challenge" (ILSVCR), l'approccio proposto da *Krizhevsky, Sutskever e Hinton*\* basato su Convolutional Neural Network ottiene un top-5 test error del 15.3% (il secondo classificato ottiene 26.2%)

---

\*Krizhevsky, A., Sutskever, I. and Hinton, G. *ImageNet classification with deep convolutional neural networks*. In Proc. Advances in Neural Information Processing Systems 25 1090-1098 (2012)

# ILSVCR 2012



- obiettivo: identificare gli oggetti presenti nelle immagini
- training set: 1,2M immagini, 1K categorie di oggetti annotati
- validation e test set: 150K immagini, 1K categorie di oggetti da riconoscere
- task: classificazione (elencare i 5 oggetti riconosciuti con la più alta confidenza)

# Deep Learning – Timeline

- **Marzo 2013:** Google acquisisce “DNNResearch Inc.” di Geoffrey Hinton (e lo assume...)
- **Dicembre 2013:** Yann Lecun diventa direttore di “Facebook AI Research”
- **Gennaio 2014:** Google acquisisce DeepMind
- **Giugno 2014:** DeepFace di Facebook ottiene il 97.35% di accuratezza nel riconoscimento facciale, sul dataset LFW

# Deep Learning – Timeline

- **Dicembre 2015:** Il team del “Microsoft Research Laboratory” vince l’ILSVRC 2015
- **Gennaio 2016:** Microsoft rilascia i sorgenti del proprio “Computational Network Toolkit” CNTK
- **Marzo 2016:** AlphaGo batte Lee Sedol
- **Gennaio 2017:** Microsoft acquisisce Maluuba e “recluta” Yoshua Bengio come consigliere
- ...

# Premio Turing 2018

Yann LeCun, Geoffrey Hinton e Yoshua Bengio vengono insigniti del premio Turing 2018



## Disclaimer (2)

### Deep Learning

**An MIT Press book**

**Ian Goodfellow and Yoshua Bengio and Aaron Courville**

<http://www.deeplearningbook.org/>

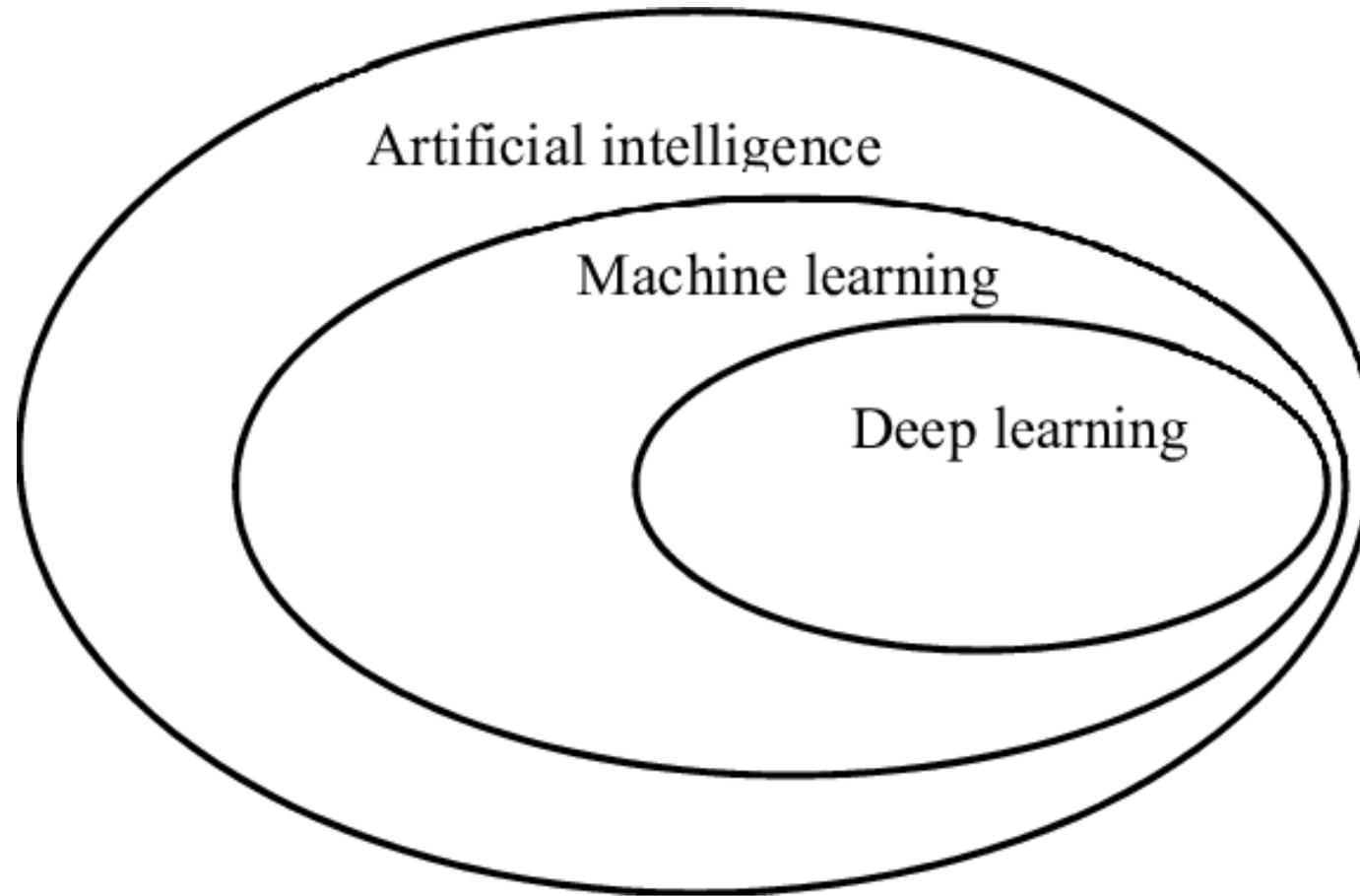
# Disclaimer (3)



A screenshot of the Coursera website. At the top left is the Coursera logo and an 'Explore' dropdown menu. A search bar contains the text 'What do you want to learn?'. On the top right, there are links for 'Online Degrees', 'Find your New Career', 'For Enterprise', 'For Universities', 'Log In', and a 'Join for Free' button. The main content area has a dark teal background. It shows a breadcrumb trail: 'Browse &gt; Data Science &gt; Machine Learning'. The course title 'Deep Learning Specialization' is prominently displayed. Below it is a description: 'Become a Machine Learning expert. Master the fundamentals of deep learning and break into AI. Recently updated with cutting-edge techniques!'. There are five stars and the text '4.9 122.813 ratings'. The instructor is listed as 'Andrew Ng +2 more instructors' with a 'TOP INSTRUCTORS' badge. A red 'Enroll for Free' button is present, with 'Starts May 7' and 'Financial aid available' text. Below the button is a text input field for an email address and social media login options for Google and Facebook. At the bottom of the section, it says '678,622 already enrolled'. The 'Offered By' section shows the 'DeepLearning.AI' logo.

<https://www.coursera.org/specializations/deep-learning>

# IA, ML e DL



# Reti Neurali (Neural Networks)

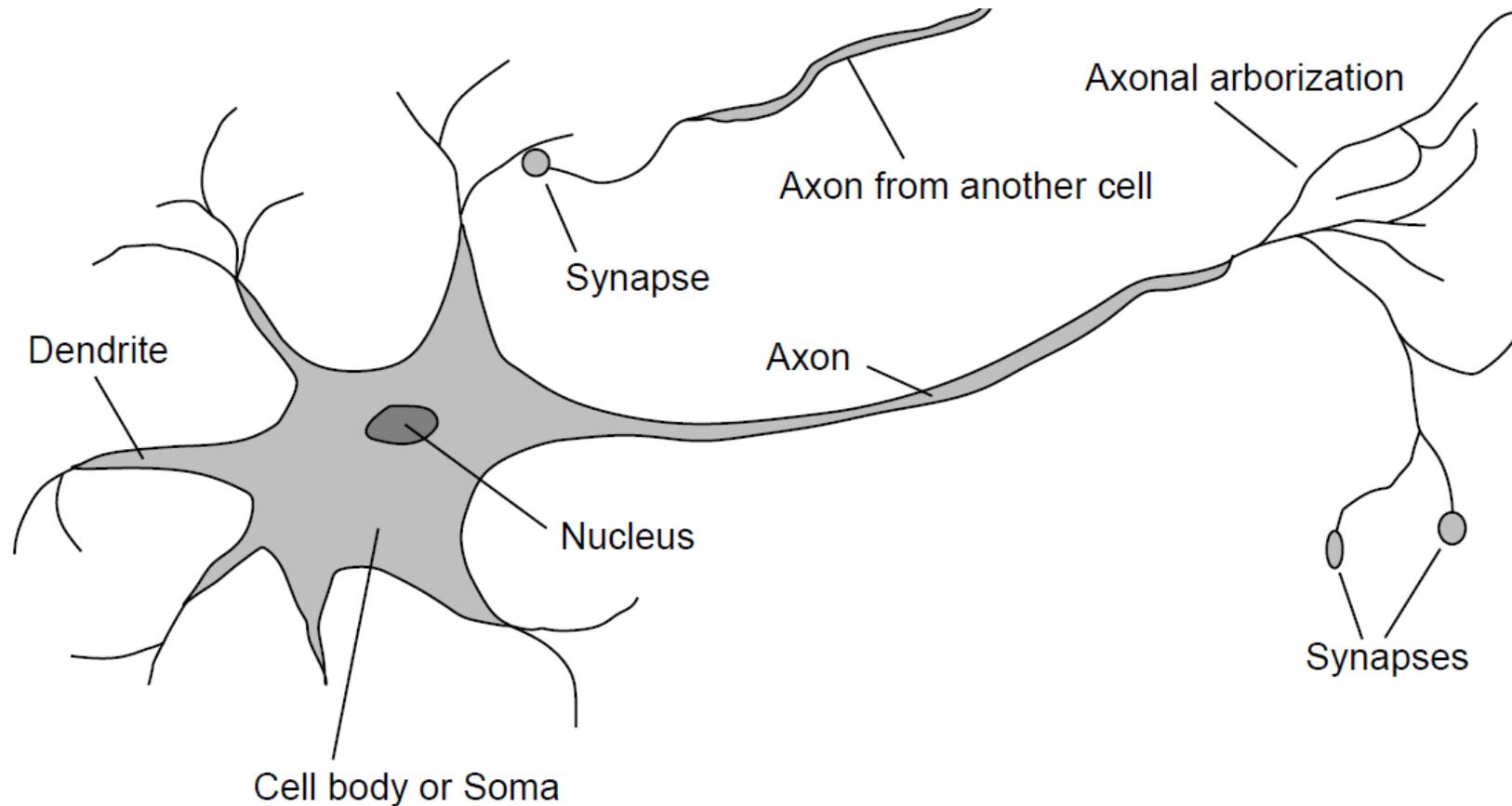
# Riferimenti



- S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Prentice Hall, 2009.
- M. Nielsen, Neural Networks and Deep Learning. Determination Press, 2015. <http://neuralnetworksanddeeplearning.com>

# Neurone

L'unità fondamentale di tutto il tessuto nervosa, dove avviene l'elaborazione dell'informazione.



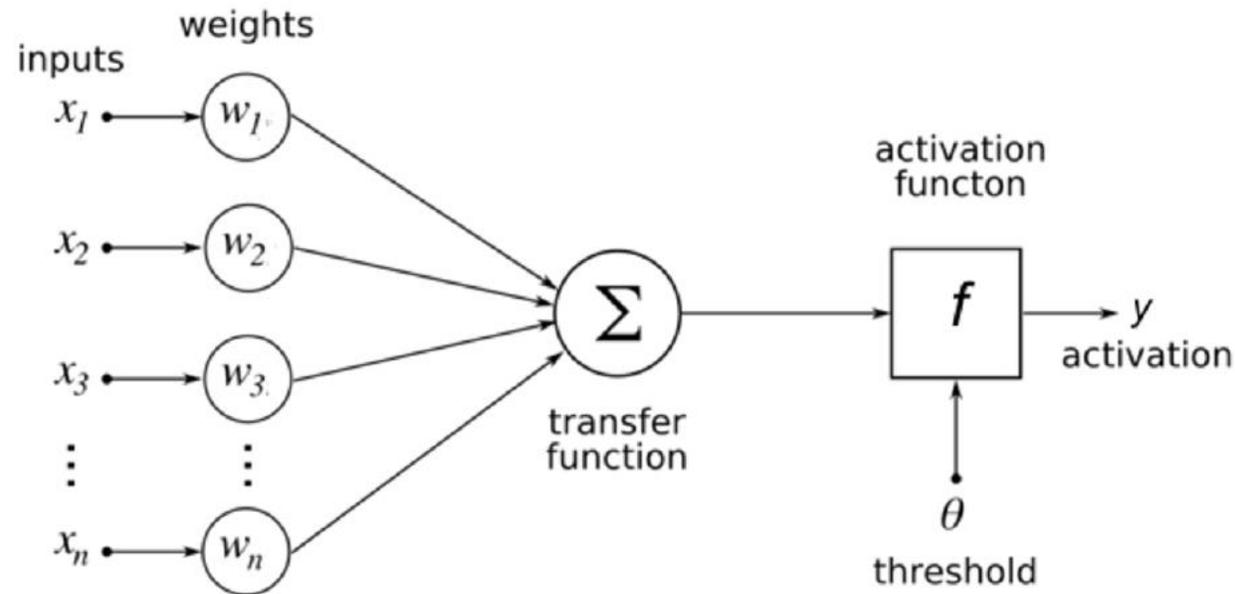
# Neurone

- Sostanze che fungono da trasmettitori chimici vengono rilasciate dalle sinapsi (giunzioni tra neuroni) ed entrano nei dendriti (ingressi del neurone), innalzando o abbassando il potenziale elettrico di un neurone;
- quando il potenziale raggiunge una soglia, il potenziale di attivazione del neurone, un impulso elettrico viene inviato lungo l'assone (l'output del neurone);
- l'impulso elettrico corre lungo le ramificazioni dell'assone, connesse tramite sinapsi ai dendriti di altri neuroni, provocando il rilascio dei trasmettitori nei corpi di altre cellule.

# Neurone Artificiale

Modello matematico (semplificato) del neurone biologico

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right)$$



# Neurone Artificiale

Modello matematico (semplificato) del neurone biologico

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right)$$

- Il neurone esegue la somma pesata degli  $n$  input ( $x_i$ , con pesi  $w_i$ ), che costituisce l'analogo del «potenziale di attivazione» del neurone;
- la funzione non lineare  $f$  (funzione di attivazione) con ingresso a soglia ( $\theta$ ) trasforma la somma pesata degli ingressi nel valore di uscita ( $y$ );

# Neurone Artificiale

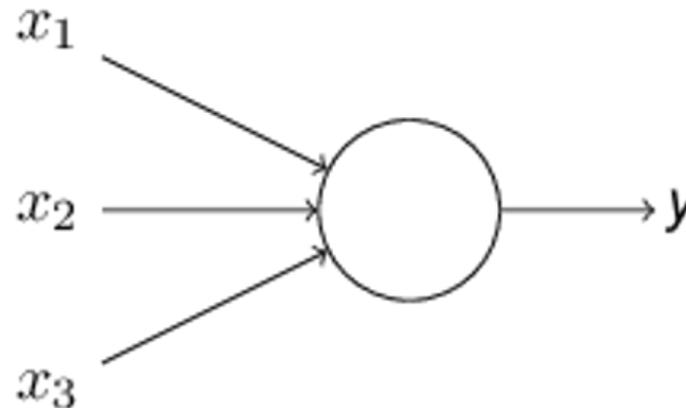
Modello matematico (semplificato) del neurone biologico

$$y = f\left(\sum_{i=0}^n w_i x_i\right)$$

con  $x_0 = -1$  e  $w_0 = \theta$

# Percettrone

Il percettrone è un modello di neurone artificiale, proposto da Rosenblatt\* (1958), sulla base del lavoro di McCulloch and Pitts\*\* a cui si deve il primo modello logico-matematico del cervello della corteccia cerebrale (1943).



---

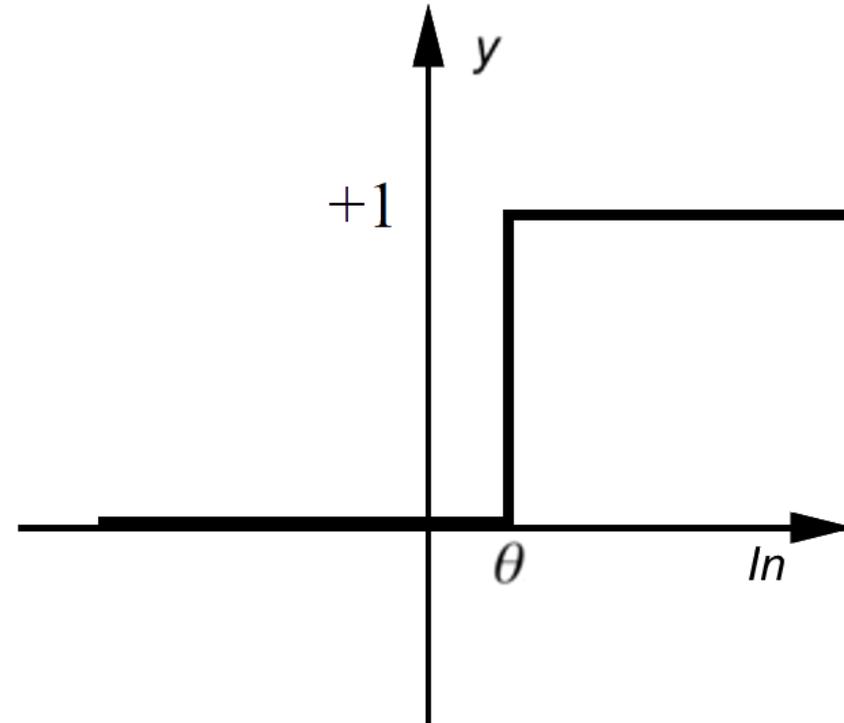
\*F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review*, vol. 65, no. 6, p. 386, 1958.

\*\*W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115—133, 1943.

# Percettrone

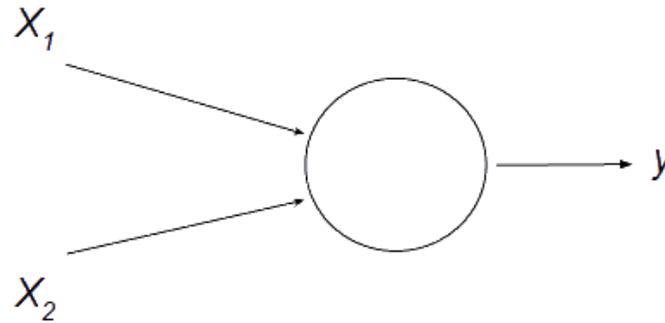
E' un classificatore binario a due strati, uno di ingresso e uno di uscita.

$$y = \begin{cases} 0 & \text{se } \sum_{i=1}^n w_i x_i \leq \theta \\ 1 & \text{se } \sum_{i=1}^n w_i x_i > \theta \end{cases}$$



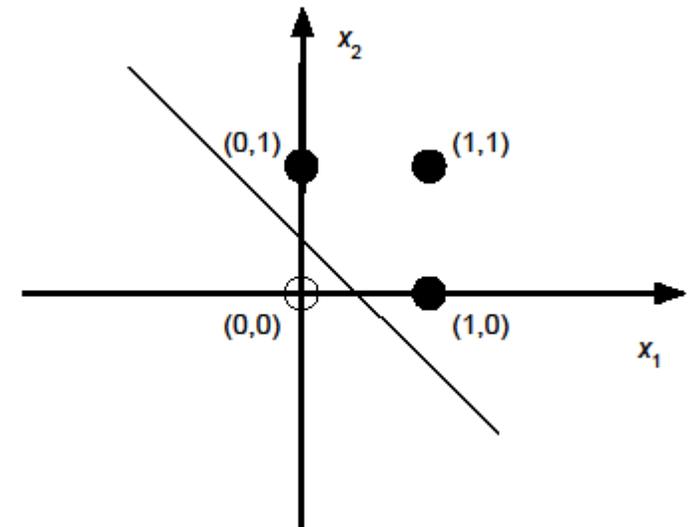
# Percettrone – OR

Quali funzioni può rappresentare un percettrone?



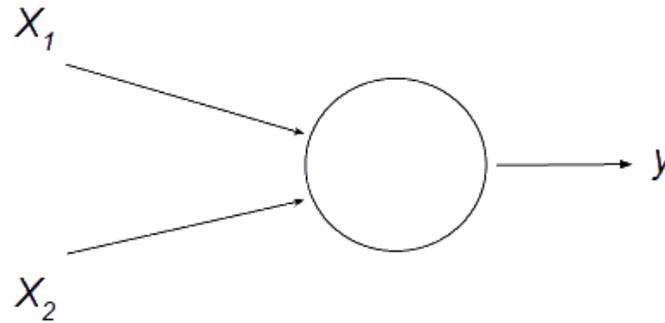
$$x_1 + x_2 - 0.5 > 0 \quad (\theta = 0.5, w_1 = w_2 = 1)$$

A	B	OR
0	0	0
0	1	1
1	0	1
1	1	1



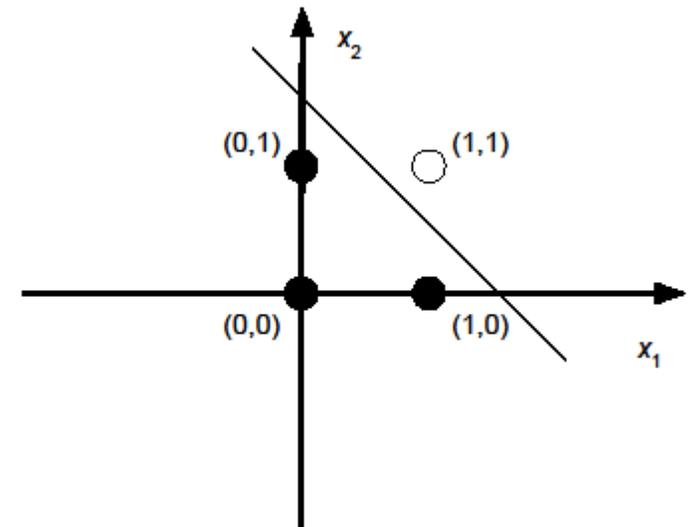
# Percettrone – NAND

Quali funzioni può rappresentare un percettrone?



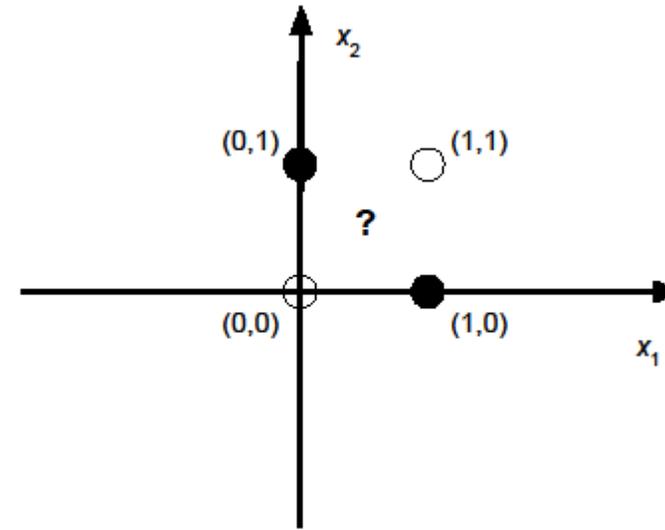
$$-x_1 - x_2 + 1.5 > 0 \quad (\theta = -1.5, w_1 = w_2 = -1)$$

A	B	NAND
0	0	1
0	1	1
1	0	1
1	1	0



# Percettrone – XOR

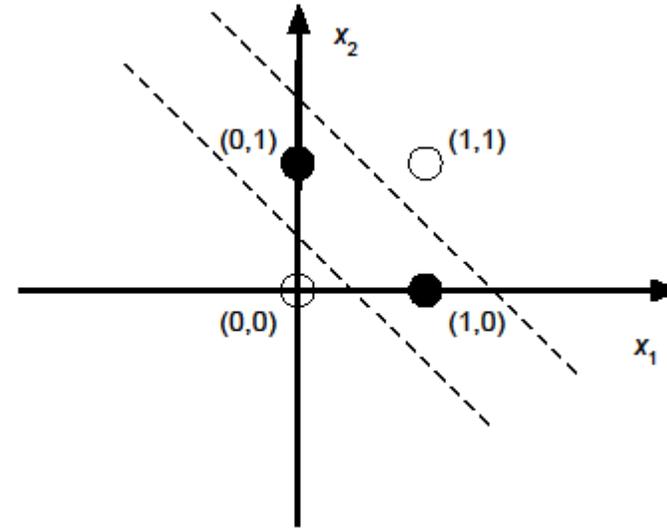
A	B	XOR
0	0	0
0	1	1
1	0	1
1	1	0



Lo XOR non può essere rappresentato!

# Perceptrone – XOR

A	B	XOR
0	0	0
0	1	1
1	0	1
1	1	0



# Percettrone

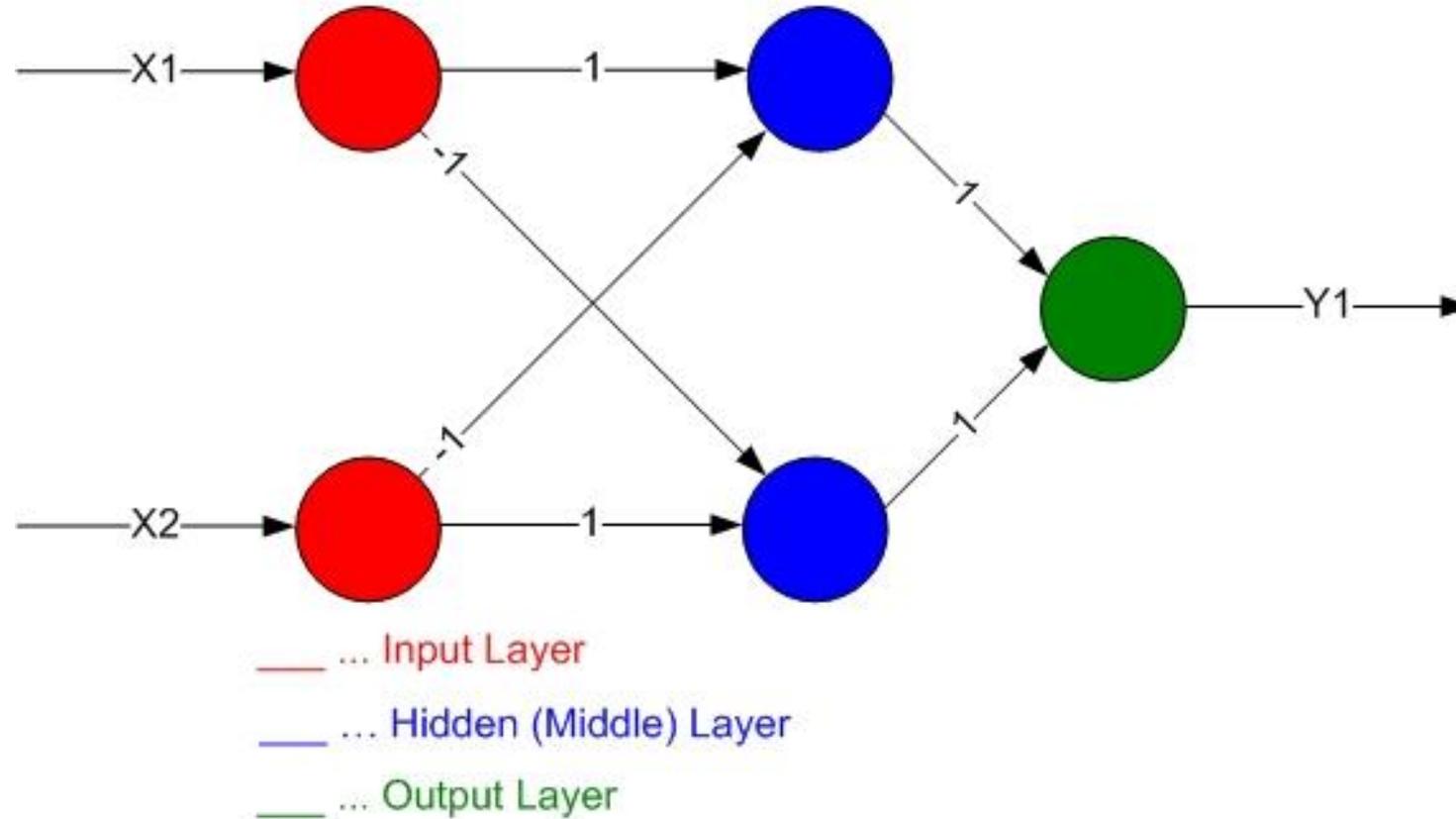
Un percettrone può risolvere problemi linearmente separabili (all'interno dello spazio di input esiste un iperpiano che separa gli elementi delle due classi).

Per poter trattare problemi non linearmente separabili, occorre utilizzare una rete di percettroni, basata su più strati:

- strato di ingresso, che include gli input della rete;
- strato di uscita, che include le uscite delle rete;
- strato nascosto, che include tutti i neuroni tra lo strato di ingresso e quello di uscita.

# Perceptrone – XOR

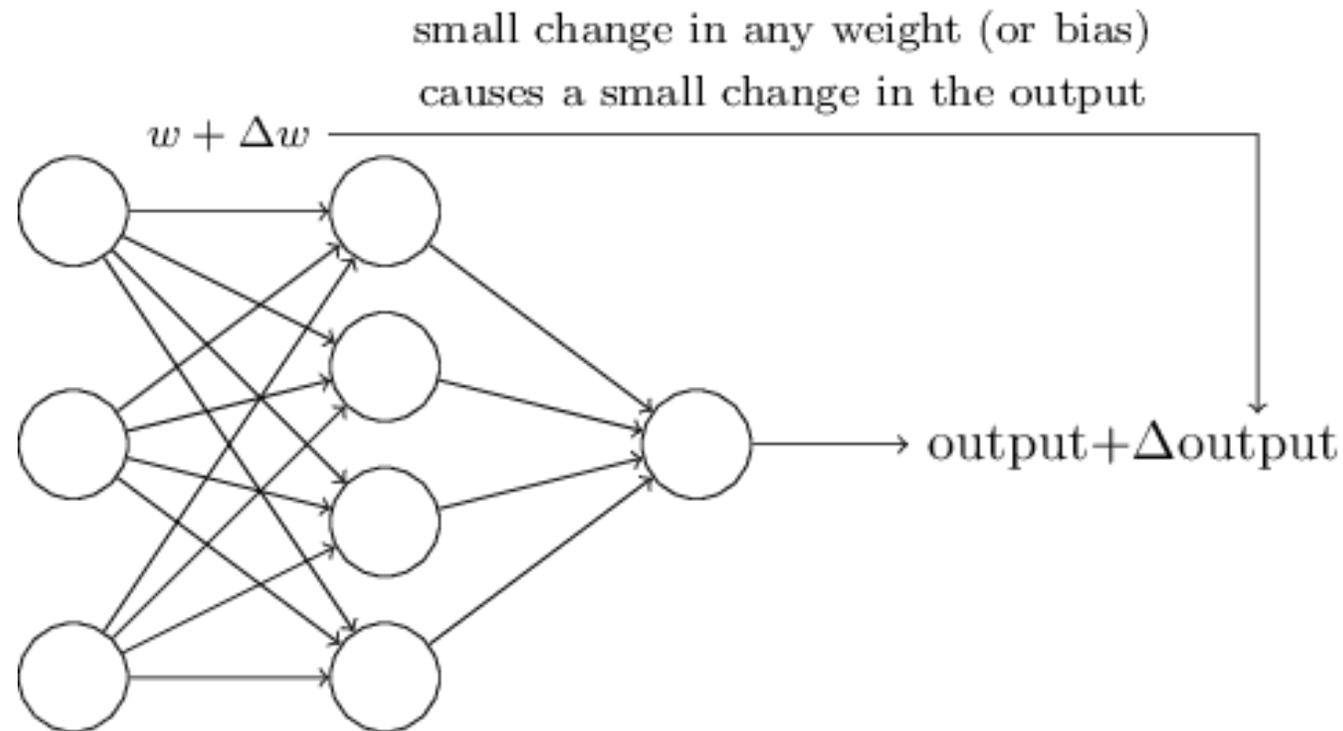
$$Y1 = \text{XOR}(X1, X2)$$



Da [https://commons.wikimedia.org/wiki/File:Perceptron\\_XOR.jpg](https://commons.wikimedia.org/wiki/File:Perceptron_XOR.jpg)

# Percettrone: apprendimento

L'apprendimento del percettrone, cioè la ricerca dei valori ottimali dei pesi  $w$  dovrebbe essere fatto in modo che a un piccolo cambiamento del peso corrisponda un piccolo cambiamento dell'output.



# Percettrone: apprendimento

- I pesi e le soglie sono inizializzati in modo casuale;
- gli esempi di training (addestramento) sono sottoposti alla rete;
- i pesi e le soglie sono costantemente aggiornati, nel tentativo di rendere le uscite della rete consistente con gli esempi di training (i pesi vengono modificati al fine di ridurre la distanza tra l'output e l'uscita desiderata)
- la fase di aggiornamento dei pesi avviene diverse volte per ogni esempio (epoche di apprendimento).

# Percettrone: apprendimento

Siano  $y_k$  il valore di uscita del percettrone e  $t_k$  il valore di uscita desiderato a fronte del  $k$ -esimo esempio di training, e sia  $err_k = t_k - y_k$ , la regola di modifica dei pesi è data da:

$$w_j^{i+1} = w_j^i + \alpha \cdot x_j \cdot err_k$$

dove  $\alpha$  è il tasso di apprendimento (un coefficiente che regola quanto proporzionale all'errore deve essere la modifica del peso).



L'apprendimento è proporzionale all'errore commesso dal percettrone nel produrre l'uscita!

# Percettrone: apprendimento



Teorema di convergenza del percettrone (Rosenblatt):

L'algoritmo di apprendimento del percettrone converge sempre ad un insieme di pesi che rappresentano correttamente gli esempi, se questi provengono da una funzione linearmente separabile.

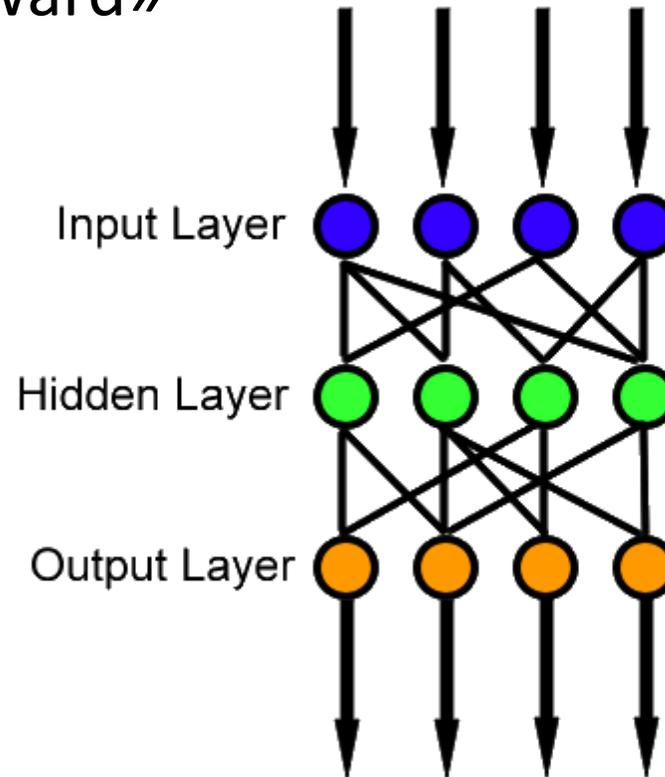
# Artificial Neural Network (ANN)

Per rappresentare problemi non linearmente separabili abbiamo quindi bisogno di architetture più complesse di quelle a singolo strato come il perceptrone. Due sono i pattern di connessioni che caratterizzano le reti neurali:

- feed-forward — il “segnale” si propaga dall’ingresso all’uscita;
- feedback (recurrent neural networks) – il “segnale” va dallo stato di uscita a quello di ingresso.

# Artificial Neural Network (ANN)

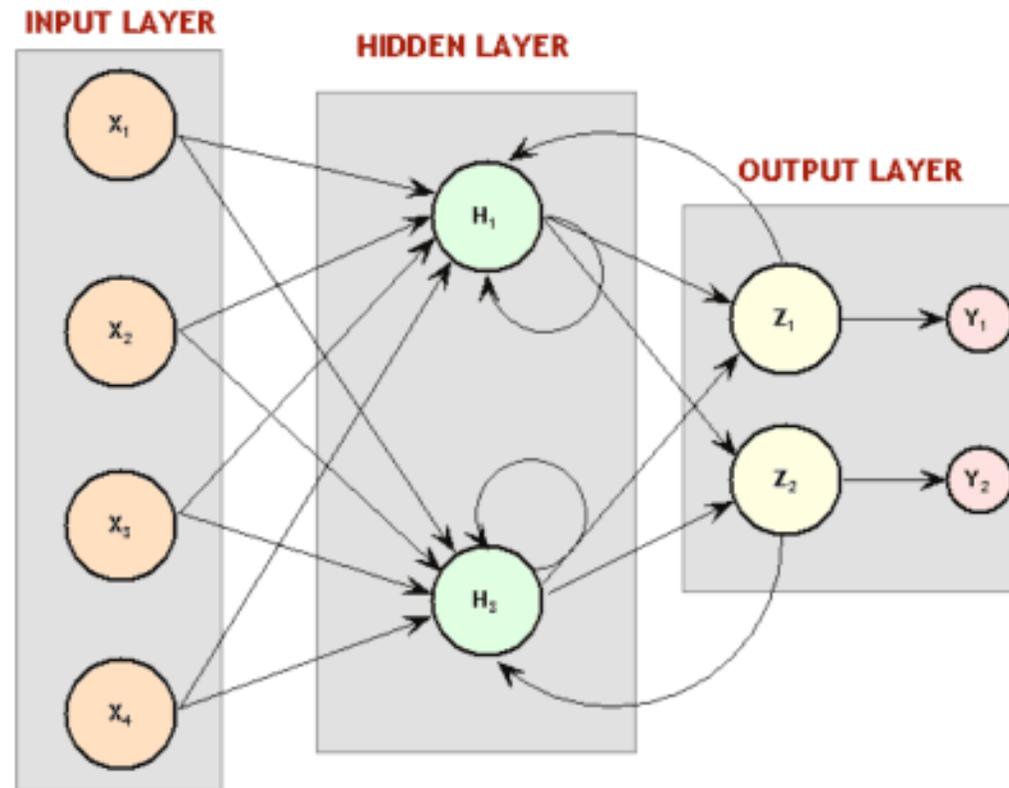
Rete neurale «feed-forward»



Da [https://en.wikipedia.org/wiki/feed-forward\\_neural\\_network#/media/File:Feed\\_forward\\_neural\\_net.gif](https://en.wikipedia.org/wiki/feed-forward_neural_network#/media/File:Feed_forward_neural_net.gif)

# Artificial Neural Network (ANN)

Rete neurale «ricorrente»



Da <http://www.cs.ucf.edu/~bgong/CAP6412.html>

# Artificial Neural Network (ANN)



Le reti neurali «feed-forward» sono approssimatori universali

Una rete «feed-forward» con un livello nascosto può approssimare in modo arbitrariamente preciso ogni funzione continua\*.

---

\*Gybenko, G. (1989). *Approximation by superposition of sigmoidal functions*. Mathematics of Control, Signals and Systems, 2(4), 303-314.

\*R. Lippmann, *An introduction to computing with neural nets*, in IEEE ASSP Magazine, vol. 4, no. 2, pp. 4-22, Apr 1987.

# Apprendimento

L'apprendimento è il processo mediante il quale la rete aggiorna la propria struttura al fine di approssimare una qualche funzione (es. classificatore).

Tipi di apprendimento:

- Supervisionato
- Non supervisionato
- Basato su rinforzo

# Apprendimento supervisionato

Alla rete viene sottoposto un insieme di addestramento (training set) composto da coppie  $\langle \mathbf{x}, t \rangle$  dove  $\mathbf{x}$  è un ingresso (un insieme di valori, es. i pixel di un'immagine) e  $t$  è l'uscita desiderata

- La rete calcola l'output a partire dall'ingresso  $\mathbf{x}$ ;
- Sulla base della differenza tra l'output desiderato  $t$  e l'output effettivo della rete, la rete modifica i propri pesi per minimizzare tale differenza
- Il training set è sottoposto più volte alla rete (epoche di apprendimento) fintanto che tutti gli esempi sono classificati correttamente o l'errore scende sotto una certa soglia

# Apprendimento non supervisionato

Il training set contiene solo gli ingressi, senza alcun output desiderato

- La rete raggruppa (in «cluster») gli input, secondo qualche criterio di similarità
- Non c'è alcun controllo sull'errore

Tipicamente è usato per problemi di clusterizzazione, quando la classe di appartenenza non è nota a priori (es. Self Organizing Maps – SOM)

# Apprendimento basato su rinforzo

Un «supervisore esterno» interpreta l'output della rete come corretto o sbagliato.

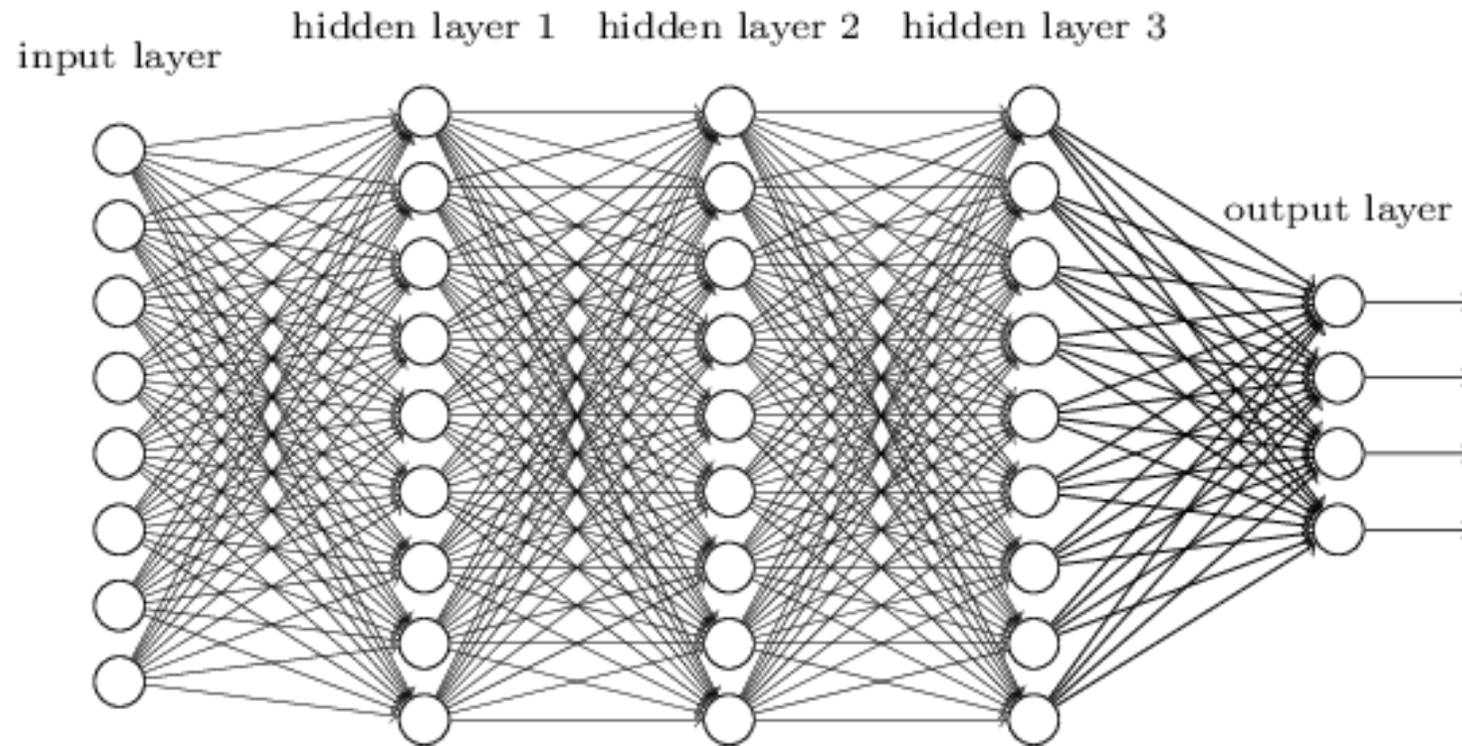
- La rete calcola autonomamente l'output
- Il «supervisore esterno» segnala l'output come corretto o sbagliato, premiando o punendo la rete.

Es: <https://www.youtube.com/watch?v=kopoLzvh5jY>

# Multi-Layer Perceptron (MLP)

- Rete neurale feed-forward multi strato;
- Uno strato di ingresso, uno di uscita, da 1 a n strati nascosti;
- Tipicamente completamente connesso: l'uscita di ogni nodo di uno strato è un input per ogni nodo dello strato successivo;
- In letteratura si parla di MLP anche se i singoli nodi non sono percettroni

# Multi-Layer Perceptron (MLP)



Da <http://neuralnetworksanddeeplearning.com/chap5.html>

# Multi-Layer Perceptron (MLP)

Supera i limiti del percettrone (problemi linearmente separabili): uno o più strati nascosti possono individuare regioni arbitrarie nello spazio di input (intersezione di iperpiani diversi).

L'addestramento di reti MLP avviene mediante il metodo dell'Error Back Propagation (EBP)\*. Minimizza la somma degli errori quadratici (uscita desiderata – uscita effettiva): si tratta di un metodo di discesa lungo il gradiente che va a cercare il minimo della funzione di errore  $E(w)$ .

---

\*D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by back-propagating errors, Nature 323.6088 (1986): 533-536.

# Error Back Propagation (EBP)

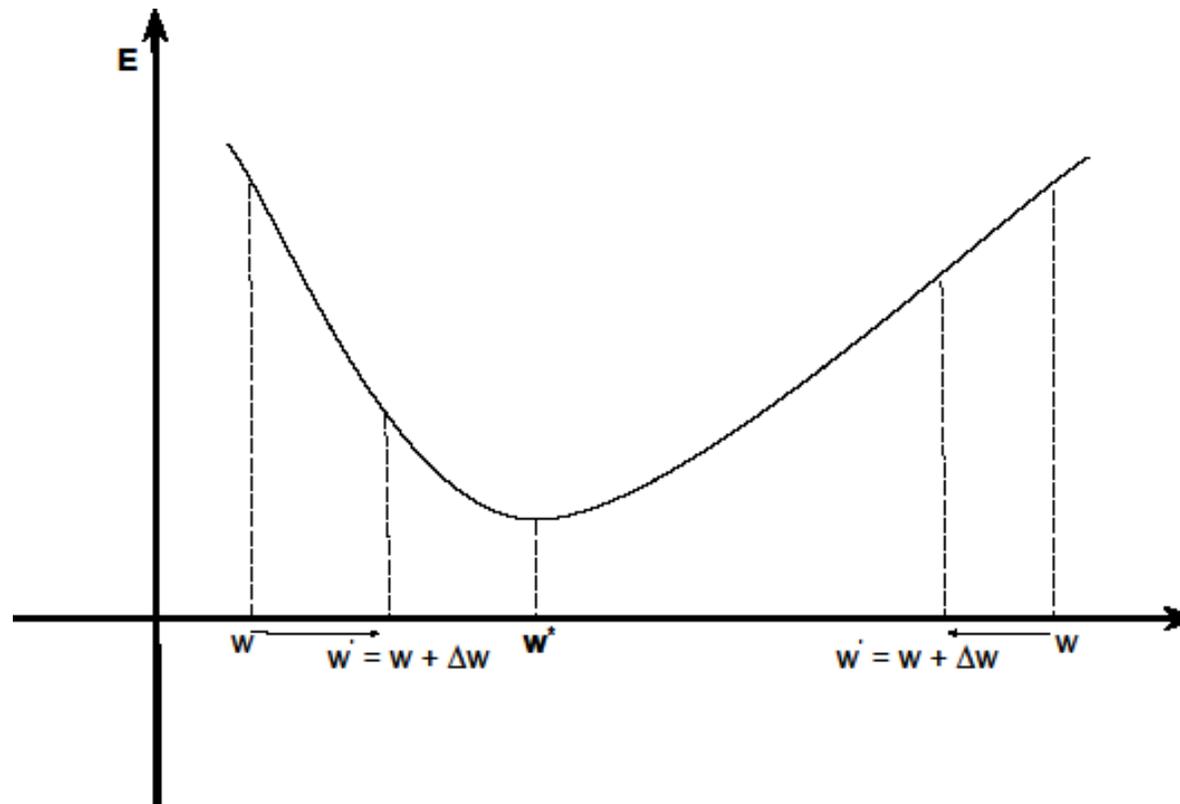
$$E(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^L \left( t_j^{(i)} - y_j^{(i)} \right)^2$$

dove:

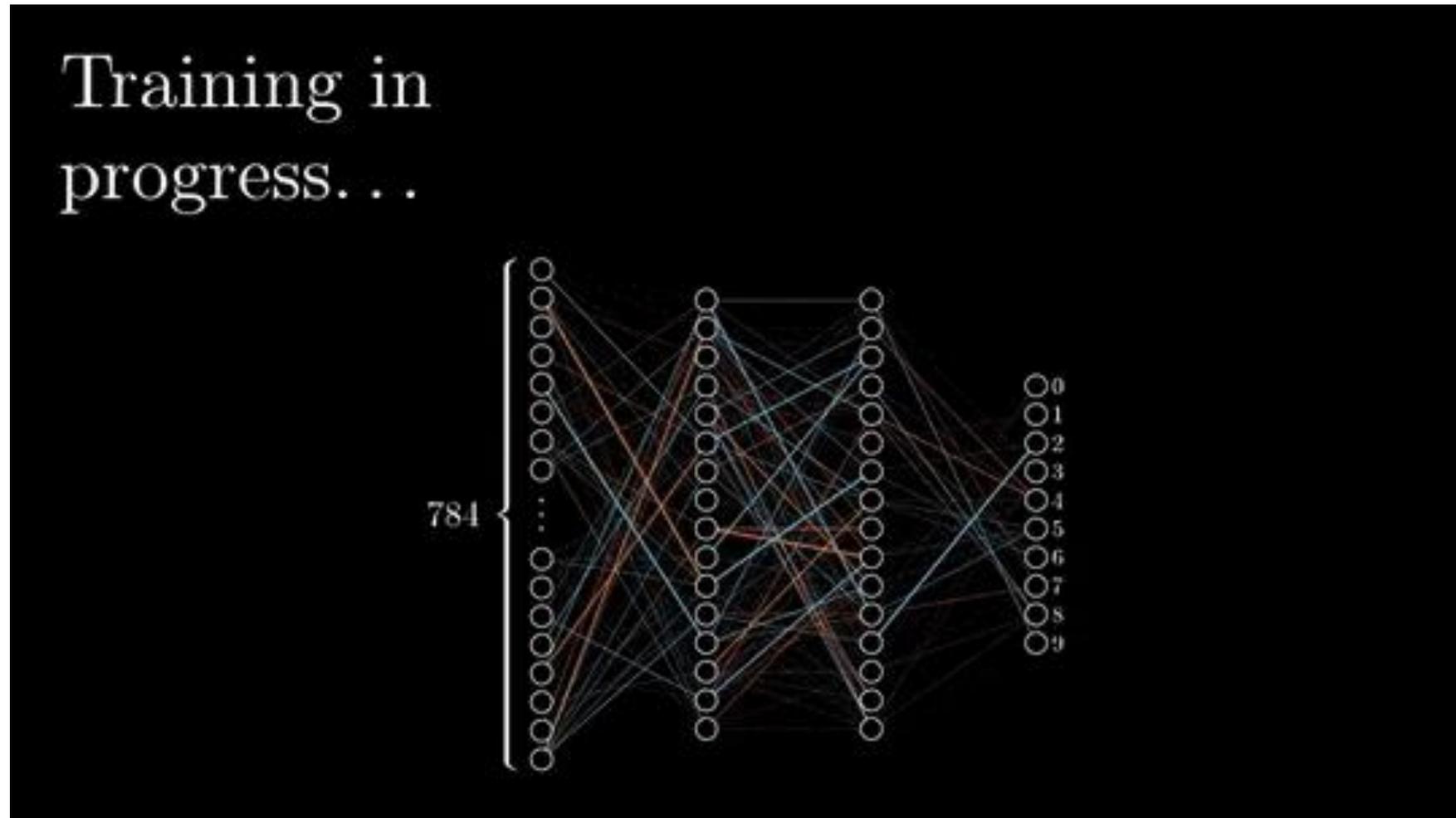
- $y_j^{(i)}$  è l'uscita calcolata dal  $j$ -esimo neurone di uscita a fronte dell' $i$ -esimo ingresso della rete (ed è quindi funzione dei pesi della rete);
- $N$  è il numero di coppie  $\langle \mathbf{x}, \mathbf{t} \rangle$  «input, uscita desiderata» che compongono il training set della rete;
- $L$  è il numero di neuroni che compongono lo strato di uscita della rete.

# Error Back Propagation (EBP)

L'addestramento consiste nella ricerca del valore dei pesi ( $w^*$ ) che minimizza la funzione di errore definita

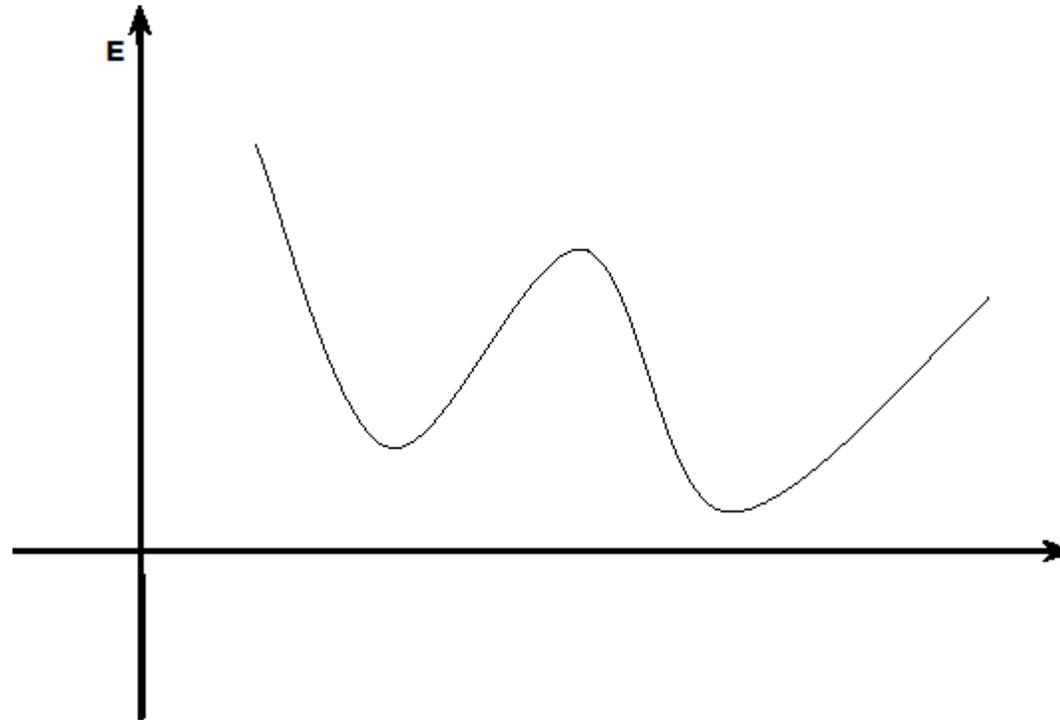


# Error Back Propagation (EBP)



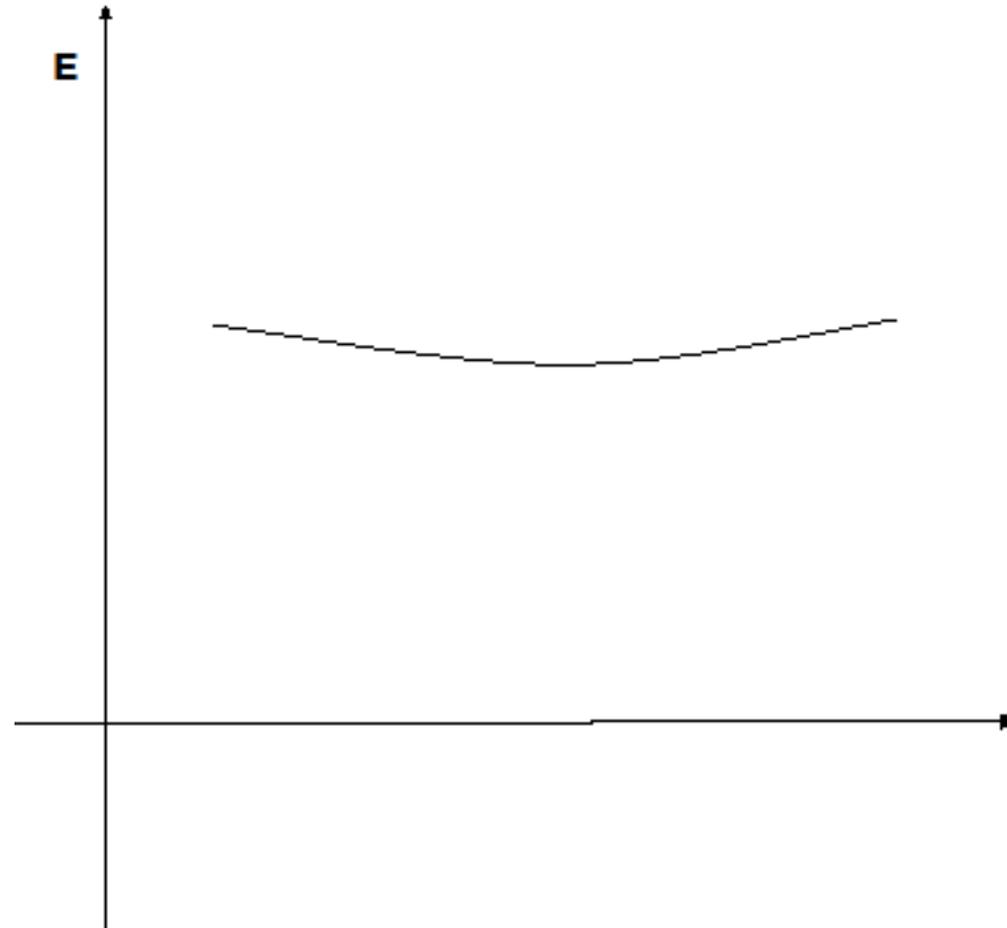
# Error Back Propagation (EBP)

La funzione di errore può avere diversi minimi locali su cui l'addestramento può bloccarsi.



# Error Back Propagation (EBP)

La convergenza può anche essere «lenta».



# Addestramento di un rete neurale

Per l'addestramento di una rete neurale con apprendimento supervisionato, occorre prepara un insieme (un «dataset») composto da coppie *<dati di input, uscita desiderata>*.

Es. Riconoscimento cifre scritte a mano da immagini.

Per ogni immagine, «l'uscita desiderata» è il numero scritto all'interno



# Addestramento di un rete neurale

Per l'addestramento di una rete neurale e la successiva verifica delle prestazioni, si costruiscono diversi insiemi di coppie  $\langle \text{input}, \text{uscita desiderata} \rangle$ . Un singolo input è anche detto «campione».

**Training set:** insieme di campioni per l'addestramento della rete neurale

**Validation set:** insieme di campioni, disgiunto da quello di training, per valutare l'addestramento in corso d'opera ed affinare (effettuare il «tuning») dei parametri della rete neurale

**Test set:** insieme di campioni, disgiunto da quello di training e di validazione, per misurare le performance della rete addestrata.

# Addestramento di un rete neurale

Es. Database «MNIST» (Modified National Institute of Standards and Technology handwritten digit database)

- Immagini 28x28 in scala di grigi
- 60.000 immagini di training
- 10.000 immagini di test

Le immagini di validazione possono essere rimuovendone alcune dal training e riservandole per la validazione (per esempio riservandone a questo scopo un 10% dell'insieme di training, cioè usandone 54.000 per l'addestramento e 6.000 per la validazione).

# Addestramento di un rete neurale

E' importante che la valutazione delle performance della rete neurale avvenga su dati «nuovi» per la rete, cioè che la rete non ha «visto» in fase di addestramento



Training, validation e test set dovrebbe essere sempre insiemi disgiunti

# Addestramento di un rete neurale

Quanti dati servono per l'addestramento?

Non c'è una risposta diretta a questa domanda. Molti problemi di riconoscimento di immagini richiedono centinaia di migliaia o milioni di esempi di addestramento.

Dipende da molti fattori: complessità del problema, numero di feature (cioè cardinalità di ogni esempio dato in input), parametri della rete...

# Addestramento di un rete neurale



Per l'addestramento:

- Si sottopone più volte alla rete neurale l'insieme di addestramento
- La rete modifica i propri pesi, minimizzando la funzione di errore (o qualche altra funzione di «loss»)
- Ogni passaggio completo di tutto l'insieme di training è detta epoca
- La rete neurale non aggiorna i propri pesi al termine del passaggio di tutto l'insieme di training. Al contrario, il training set è diviso in batch, cioè sottoinsiemi di identica dimensione. La rete aggiorna i propri pesi ogni volta esamina un batch.

# Addestramento di un rete neurale

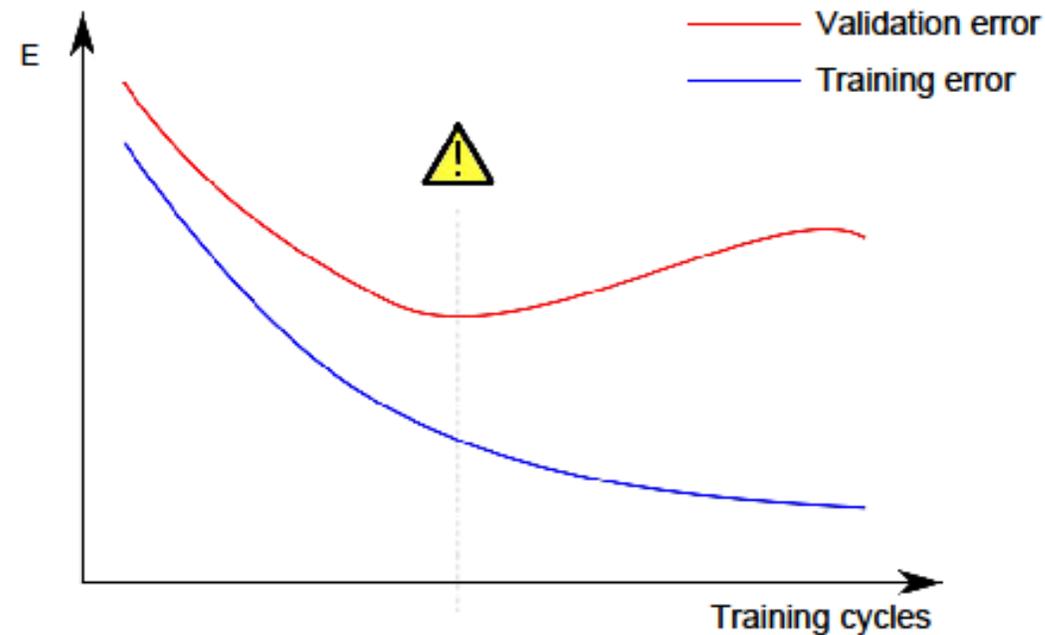
L'addestramento si ferma quando viene raggiunto un certo criterio, tipicamente misurato sull'insieme di validazione:

- Errore o loss sotto una certa soglia
- Passaggio di alcune epoche senza alcun miglioramento dell'errore o della loss

Più dati e più epoche di apprendimento significano sempre migliori prestazioni?

# Addestramento di un rete neurale

Se il numero di epoche di addestramento è troppo grande (e/o forniamo dati ridondanti, troppo simili tra loro), si può incorrere nell'overfitting



# Addestramento di un rete neurale



Overfitting (sovradattamento): un modello statistico sovradattato si adatta ai dati osservati, fallendo su dati nuovi o in previsioni future

In altre parole, la rete neurale impara «a memoria» gli esempi di training, fallendo su la classificazione di dati nuovi. Può accadere, ad esempio, perché la rete è sovradimensionata rispetto ai campioni disponibili.

# Test di un rete neurale

Quali metriche usare per valutare le performance?

Accuratezza: numero di campioni classificati correttamente su numero di campioni totali

Es. Se il training set è composto da 100 cifre scritte a mano e se ne sono identificate correttamente 75 allora l'accuratezza è:

$$\frac{75}{100} = 0,75 = 75\%$$

# Test di un rete neurale

In un problema di classificazione binario, cioè costituito da due sole classi (positiva e negativa), possiamo indicare con:

- $T_p$  i veri positivi, cioè i campioni della classe positiva identificati correttamente
- $T_n$  i veri negativi, cioè i campioni della classe negativa identificati correttamente
- $F_p$  i falsi positivi, cioè i campioni della classe negativa erroneamente identificati come positivi
- $F_n$  i falsi negativi, cioè i campioni della classe positiva erroneamente identificati come negativi

# Test di un rete neurale

Precisione (precision): la porzione di positivi identificati correttamente tra tutti i positivi identificati

$$\frac{T_p}{T_p + F_p}$$

Recupero (recall): la porzione di positivi identificati correttamente tra tutti i positivi presenti nel test set

$$\frac{T_p}{T_p + F_n}$$

La precisione è una misura di «esattezza», il recupero una di «completezza»

# Test di un rete neurale

Supponiamo che un dataset sia composto da 50 esempi positivi e 50 esempi negativi e che la rete identifichi correttamente 47 esempi positivi e 48 esempi negativi. Questo significa che:

- Ci sono 47 veri positivi ( $T_p = 47$ )
- Ci sono 48 veri negativi ( $T_n = 48$ )
- Ci sono 3 falsi negativi ( $F_n = 3$ )
- Ci sono 2 falsi positivi ( $F_p = 2$ )

L'accuratezza è  $\frac{47+48}{100} = 0,95 = 95\%$

## Test di un rete neurale

- La precisione è  $\frac{47}{47+2} = 0,96 = 96\%$
- Il recupero è  $\frac{47}{47+3} = 0,94 = 94\%$

Supponiamo di avere un altro dataset di 100 campioni, 30 positivi e 70 negativi. Con la seguente classificazione:

$$T_p = 20, T_n = 50, F_p = 20, F_n = 10$$

Avremmo l'70% di accuratezza, 50% di precisione, 67% di recall.

**Con un dataset non bilanciato, cioè con un numero molto diverso di esempi per ogni classe, l'accuratezza non è una buona metrica!**

# Test di un rete neurale

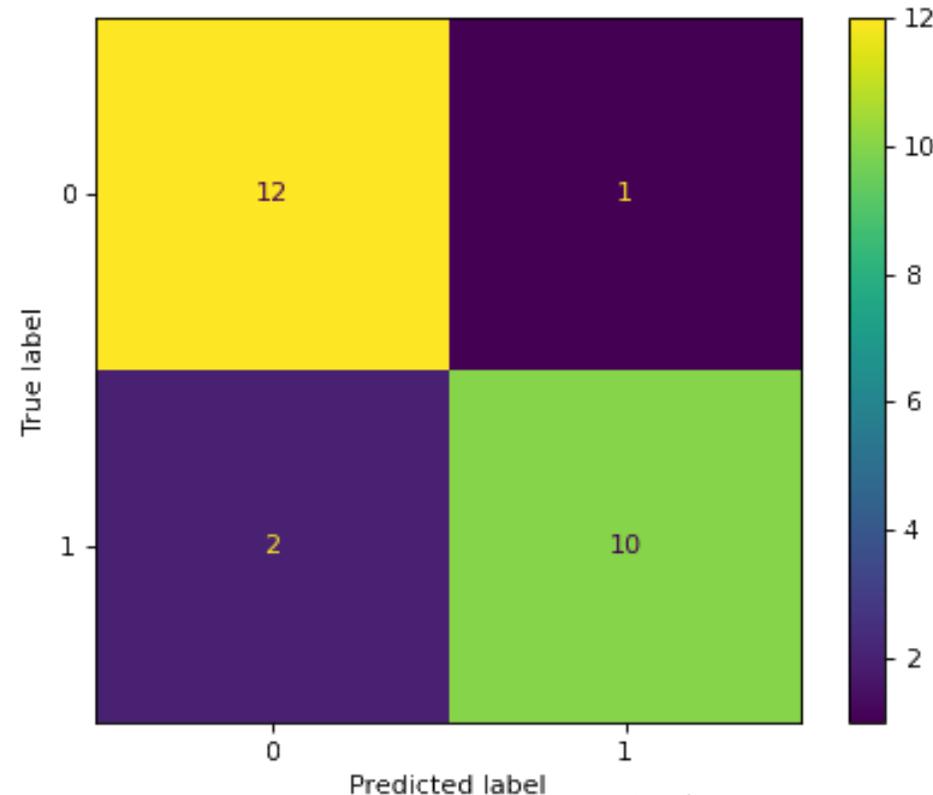
F1-score: media armonica di precisione e recupero

$$2 \frac{\textit{precisione} \times \textit{recupero}}{\textit{precisione} + \textit{recupero}}$$

Nell'esempio della slide precedente l'F1-score è  $0,57 = 57\%$

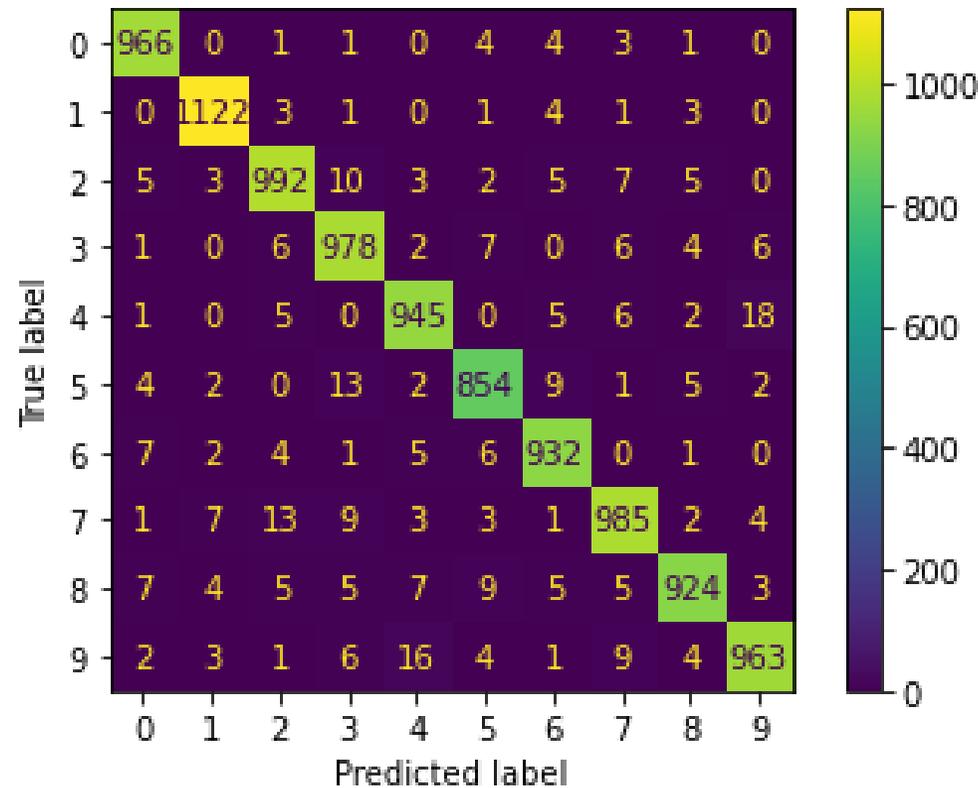
# Test di un rete neurale

Matrice di confusione: visualizzazione dell'accuratezza (e degli errori) di un classificatore. Sulle ordinate ci sono le classi reali, sulle ascisse le classi identificate dal classificatore.



# Test di un rete neurale

Es.: Identificazione delle cifre scritte a mano. Sulle ordinate c'è la cifra «reale», sulle ascisse quella identificata da una rete neurale



# Test di un rete neurale

Spesso, il training set e il test set vengono ottenuti prendendo dati dallo stesso dataset. I risultati ottenuti potrebbero dipendere dalla particolare divisione del dataset effettuata.

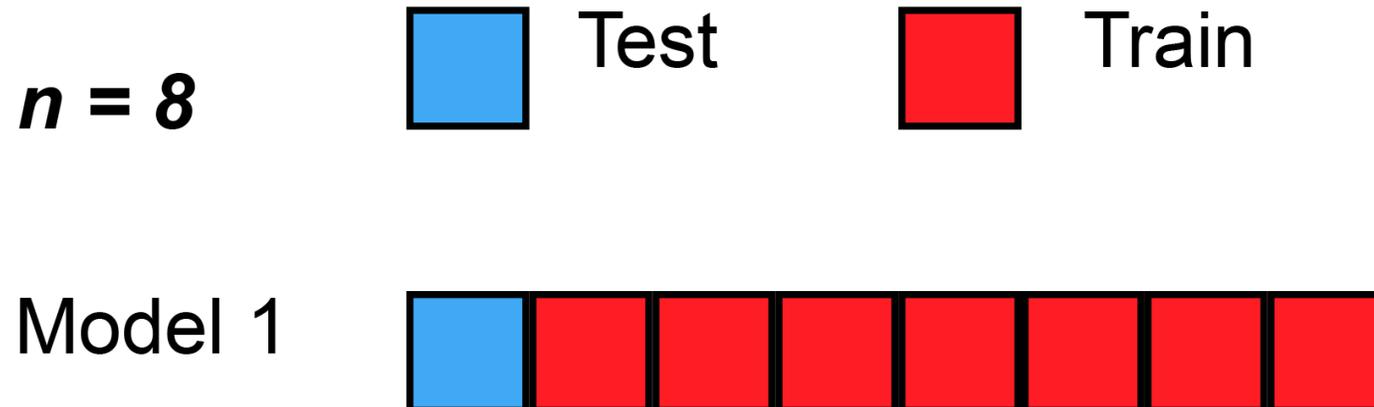
Per ovviare a questo problema si può ricorrere alla convalida incrociata

Convalida incrociata: usare differenti porzioni dei dati per il training e il test in differenti iterazioni di addestramento e valutazione, mediando poi i risultati finali.

# Test di un rete neurale

Una delle tecniche di convalida incrociata più usate è la *k*-fold Cross Validation:

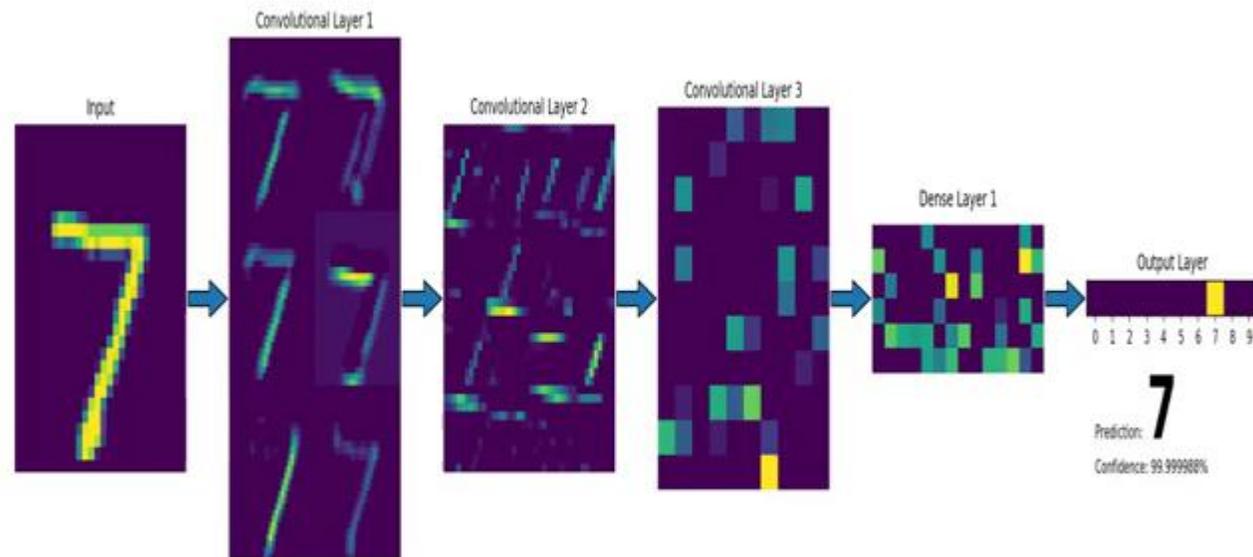
- Si divide il dataset in  $k$  sottoinsiemi di uguale dimensione
- Si eseguono training e test per  $k$  iterazioni, cioè  $k$  volte. Ad ogni iterazione si sceglie uno dei  $k$  sottoinsiemi per il test, e i restanti  $k-1$  per il training. All'iterazione successiva si cambia la scelta.



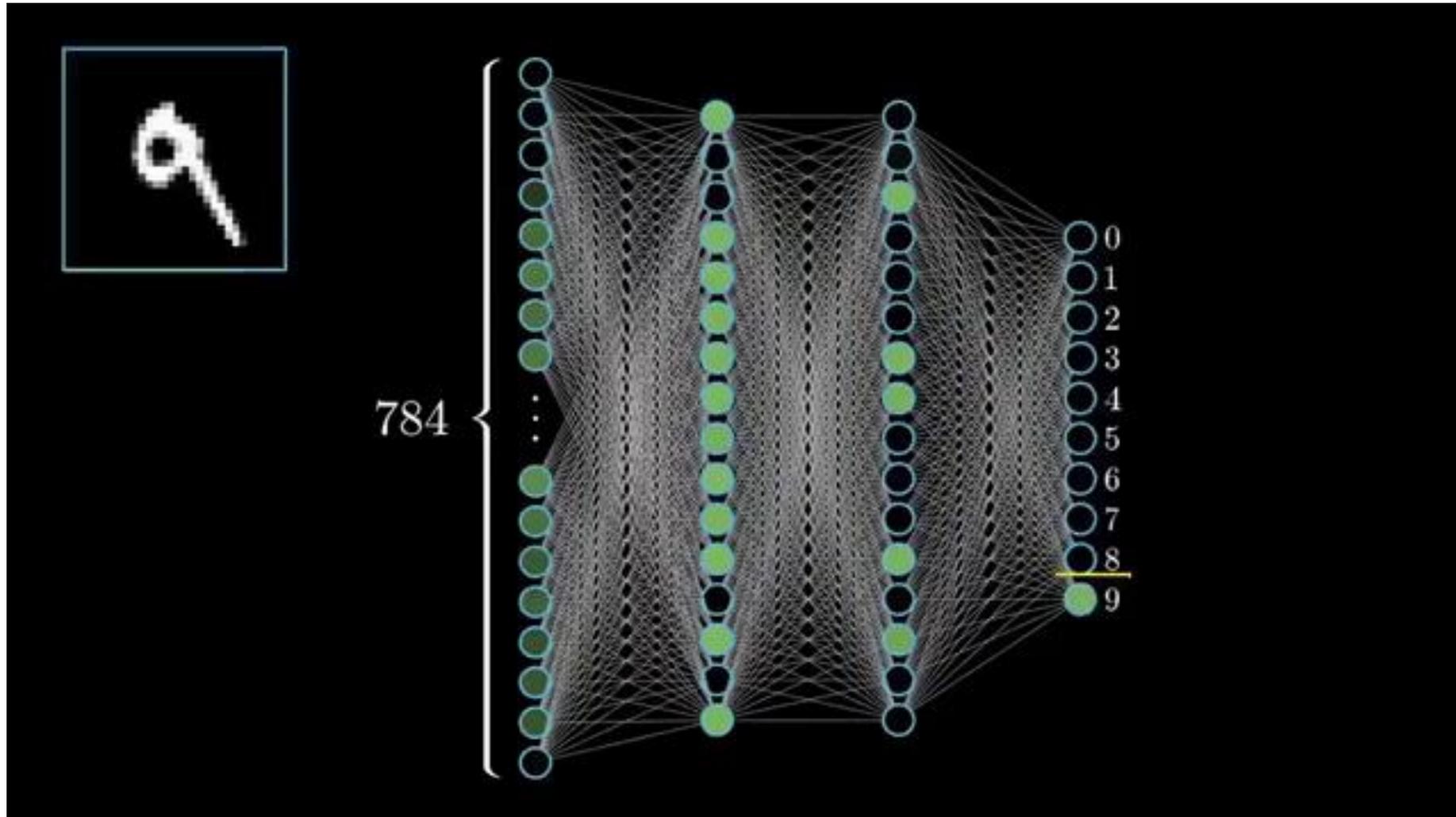
# Reti neurali e Deep Learning

# Deep Learning e Deep Neural Network

Il termine “deep” in Deep Learning e Deep Neural Network, si riferisce alla presenza di molteplici strati in cascata ottenendo rappresentazione dei dati di input a livelli sempre più astratti.

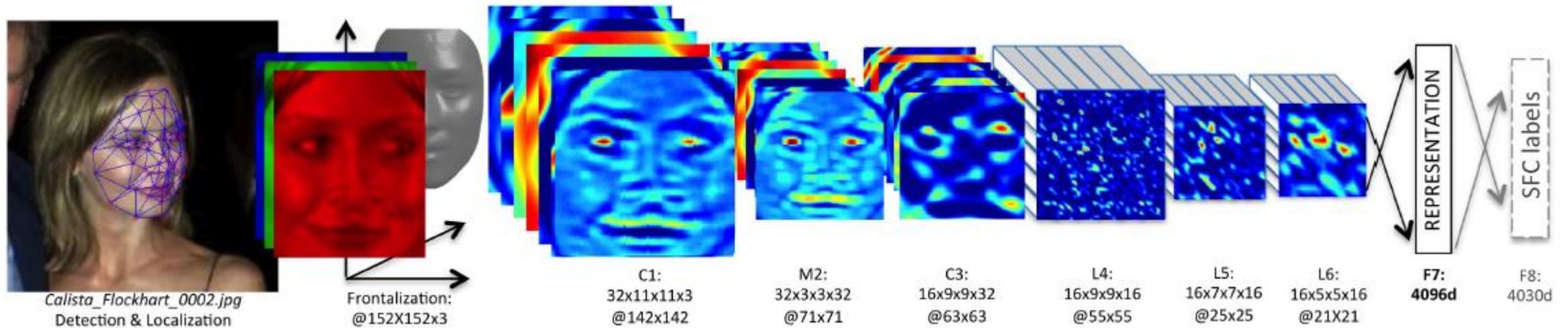


# Deep Learning e Deep Neural Network



# Esempio: Facebook DeepFace

Rete neurale per il riconoscimento facciale



Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification" in 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701 | 1708

# Esempio: Google FaceNet

Rete neurale per il riconoscimento facciale

- 22 strati
- 140 milioni di parametri
- Esperimenti con addestramento basato su 100-200 milioni di immagini di volti, di 8 milioni di persone diverse

---

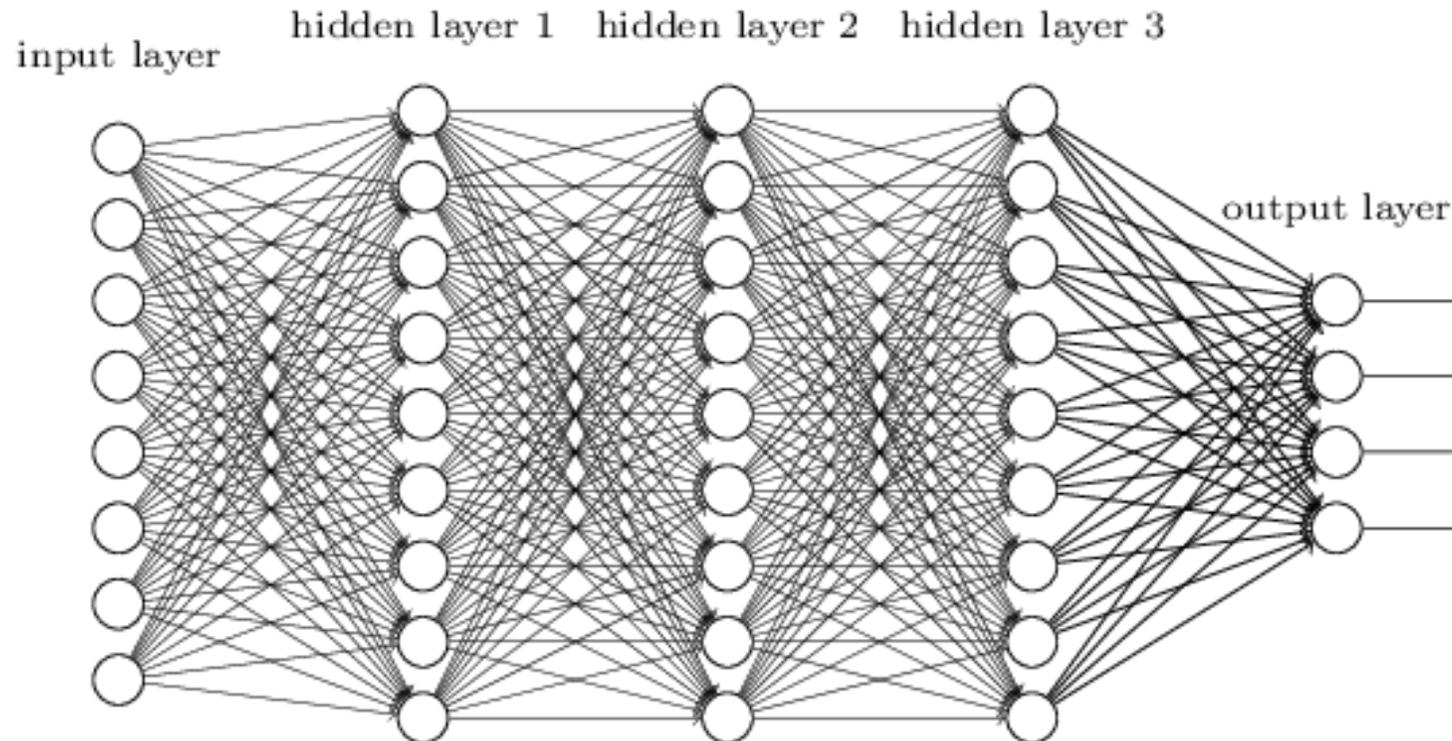
F. Schroff, D. Kalenichenko, and J. Philbin, Facenet: “A unified embedding for face recognition and clustering,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815–823

# MLP – Struttura completamente connessa



unimc

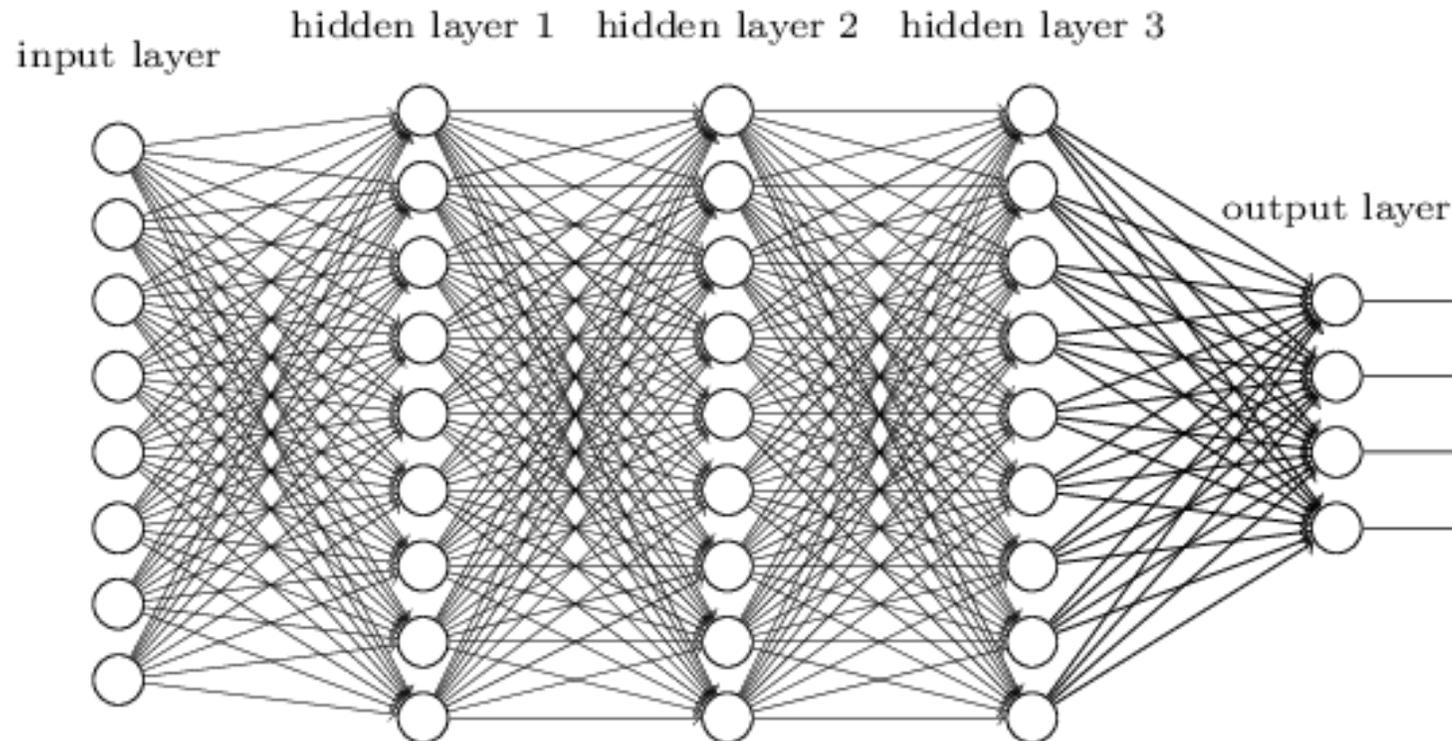
Il Multi-Layer Perceptron ha una struttura completamente connessa: tutti i nodi di uno strato sono collegati a tutti i nodi dello strato successivo.





# MLP – Struttura completamente connessa

Nel caso in cui i valori di ingresso siano i pixel di un'immagine, pixel vicini e lontani sono trattati allo stesso modo...



# Convolutional Neural Network – CNN

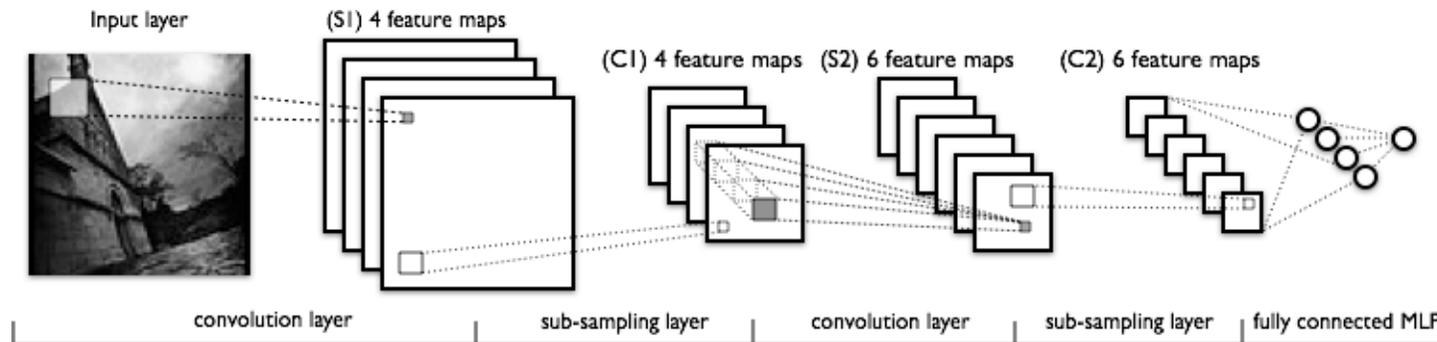


E' uno dei tipi di rete neurale più utilizzata nel Deep Learning. Come per le reti neurali in generale, anche per le convolutional neural network è possibile individuare un'ispirazione biologica, in particolare nella struttura della corteccia visiva di animali:

- le cellule che la compongono sono sensibili solo a sotto-regioni del campo visivo (receptive fields);
- come in un mosaico, le cellule compongono la corteccia visiva per coprire l'intero campo visivo
- cellule “semplici” identificano contorni e pattern di contorni
- cellule “complesse”, con receptive fields più grandi, invarianti localmente all'esatta posizione dei contorni

# Convolutional Neural Network – CNN

A differenza del MLP, le Convolutional Neural Network (CNN) non sono completamente connesse. Le connessioni sono locali: i neuroni di ogni strato non sono connessi a tutti i neuroni dello strato precedente, ma ogni neurone è connesso solo a una specifica regione.



Da <http://deeplearning.net/tutorial/lenet.html>

# CNN e Deep Learning

Questo tipo di tecniche e architetture hanno bisogno di grandi quantità di dati per l'addestramento



Servono corposi dataset annotati!

# Annotazione dei dataset (1)

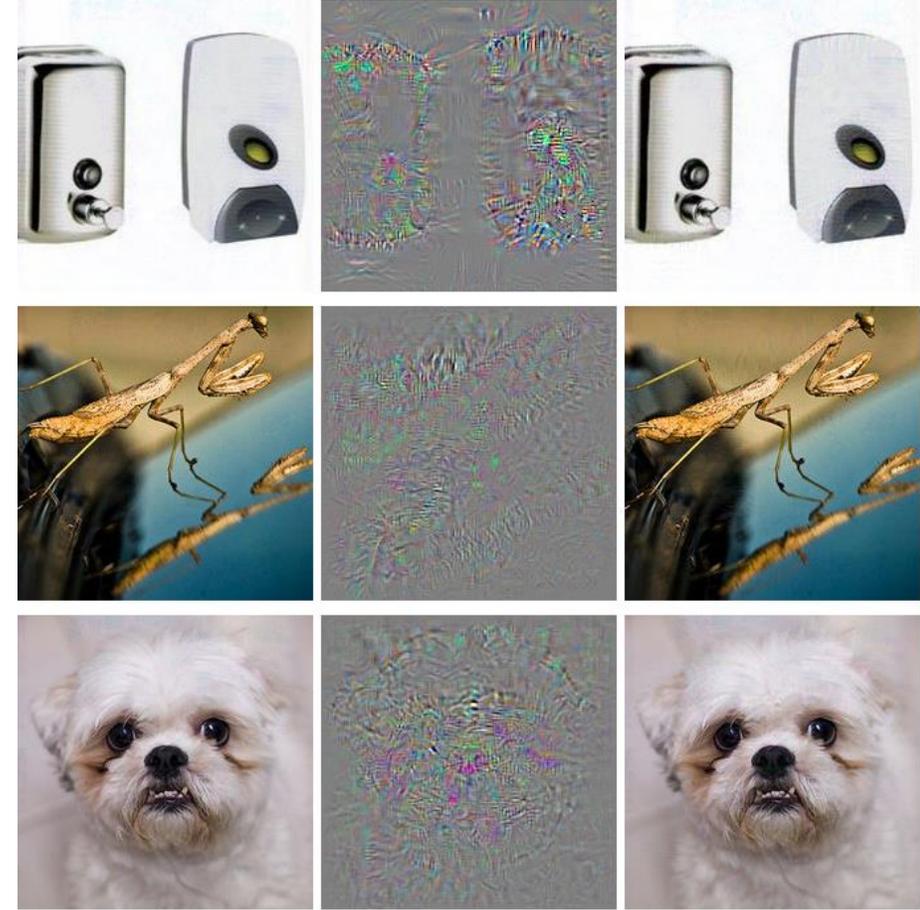
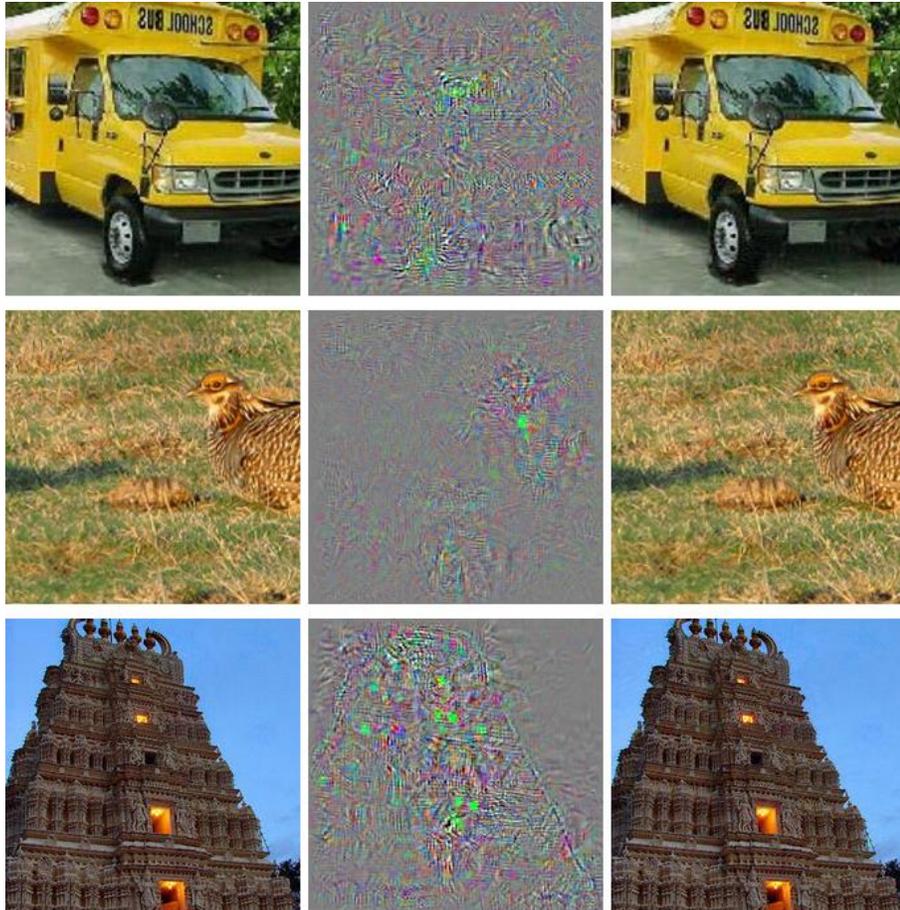


SO MUCH OF "AI" IS JUST FIGURING OUT WAYS TO OFFLOAD WORK ONTO RANDOM STRANGERS.

# Annotazione dei dataset (2)

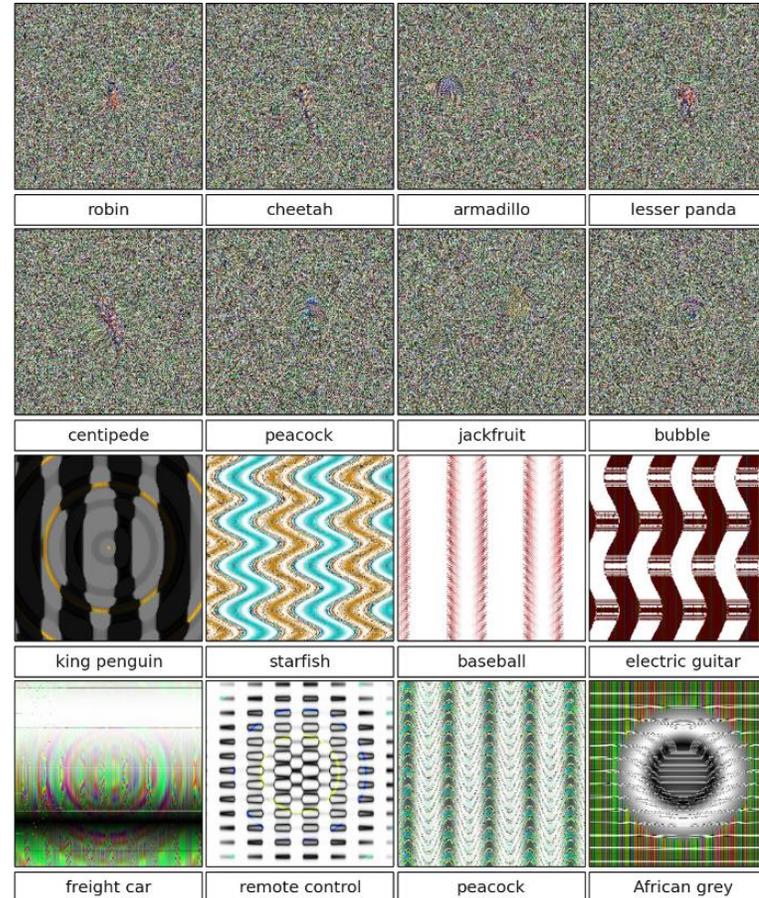
The screenshot displays two side-by-side annotation tasks. The left task, titled "Select all images with commercial lorries", shows a grid of nine images. Three images are marked with a blue checkmark: a red truck, a white truck with "ISOLATION PLUS" branding, and a white truck with a red stripe. The right task, titled "Select all squares with fire hydrants", shows a 4x4 grid of 16 squares. The top-left square contains a red fire hydrant, while the other 15 squares are empty. Both tasks include a "Verify" or "SKIP" button and a "Report a Problem" link at the bottom.

# Potenziali problemi (1)



C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, *Intriguing properties of neural networks*, arXiv preprint arXiv:1312.6199, 2013. <https://arxiv.org/abs/1312.6199>

# Potenziali problemi (2)



A. Nguyen, J. Yosinski and J. Clune, *Deep neural networks are easily fooled: High confidence predictions for unrecognizable images*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 427-436.

# Applicazioni nei trasporti

- Predizione del livello di traffico\*
- Riconoscimento automatico di etichette per la logistica\*\*
- Rilevamento automatico di oggetti per applicazioni di guida autonoma\*\*\*
- ...

---

\*Hongsuk Yi, HeeJin Jung and Sanghoon Bae, «*Deep Neural Networks for traffic flow prediction*,» 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), 2017, pp. 328-331, doi: 10.1109/BIGCOMP.2017.7881687.

\*\*S. Suh, H. Lee, Y. O. Lee, P. Lukowicz and J. Hwang, «*Robust Shipping Label Recognition and Validation for Logistics by Using Deep Neural Networks*,» 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 4509-4513, doi: 10.1109/ICIP.2019.8803412.

\*\*\*P. Li, X. Chen and S. Shen, «*Stereo R-CNN Based 3D Object Detection for Autonomous Driving*,» 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 7636-7644, doi: 10.1109/CVPR.2019.00783.

# Deep Neural Networks

Grande accuratezza in molti compiti basati sul riconoscimento di immagini (e non solo...)

L'addestramento richiede molte risorse:

- Hardware
- Migliaia/Milioni di esempi di addestramento

Usate spesso come «scatole nere»: scarsa trasparenza nella decisione

## Black box - Esempio

Una rete neurale stabilisce che due immagini di volti corrispondono alla stessa persona. Questo perché una rete neurale da 140 milioni di parametri (pesi e soglie) addestrata ad estrarre dalle immagini vettori di 1024 valori compresi tra 0 e 1 che dovrebbero descrivere il volto, stabilisce che i due vettori che rappresentano le immagini sono «vicini».

Sarebbe sufficiente durante un dibattito per stabilire la persona in quelle due foto è la stessa?

# Reti neurali come black box

GDPR – **G**eneral **D**ata **P**rotection **R**egulation

Articolo 22 - Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione\*

*«L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona.»*

---

[\\*https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:32016R0679&from=IT#d1e2845-1-1](https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:32016R0679&from=IT#d1e2845-1-1)

# Trasparenza delle decisioni

La «trasparenza» delle tecniche di IA è coinvolta in alcuni articoli del GDPR (ad esempio nell'art. 13, oltre che nel 22)

Di fatto, occorrerebbe una spiegazione sul processo decisionale. Occorrerebbe cioè:

- capire come una certa tecnica o metodologia operano
- capire quali fattori sono stati decisivi per un certo output



eXplainable Artificial Intelligence – XAI

Per approfondire: <https://www.agendadigitale.eu/cultura-digitale/lintelligenza-artificiale-deve-essere-spiegabile-ecco-i-progetti-e-le-tecniche/>

# Human level AI?

# Artificial General Intelligence - AGI



Yann LeCun

17 maggio alle ore 15:20 · 🌐

About the raging debate regarding the significance of recent progress in AI, it may be useful to (re)state a few obvious facts:

- (0) there is no such thing as AGI. Reaching "Human Level AI" may be a useful goal, but even humans are specialized.
- (1) the research community is making \*some\* progress towards HLA
- (2) scaling up helps. It's necessary but not sufficient, because....
- (3) we are still missing some fundamental concepts
- (4) some of those new concepts are possibly "around the corner" (e.g. generalized self-supervised learning)
- (5) but we don't know how many such new concepts are needed. We just see the most obvious ones.
- (6) hence, we can't predict how long it's going to take to reach HLA.

I really don't think it's just a matter of scaling things up.

We still don't have a learning paradigm that allows machines to learn how the world works, like human and many non-human babies do.

Some may believe that scaling up a giant transformer trained on sequences of tokenized inputs is enough.

Others believe "reward is enough".

A few others believe that explicit symbol manipulation is necessary.

A few don't believe gradient-based learning is part of the solution.

I believe we need to find new concepts that would allow machines to:

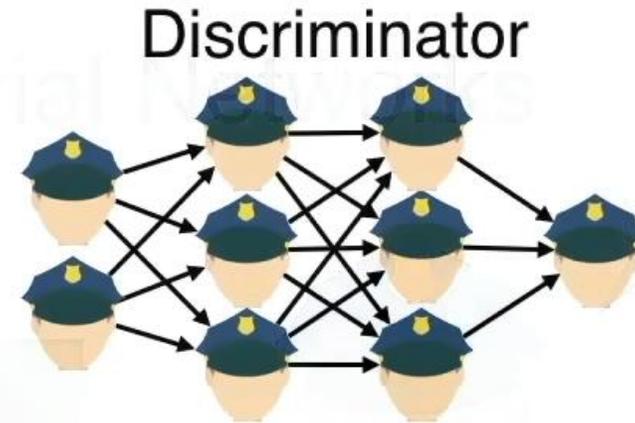
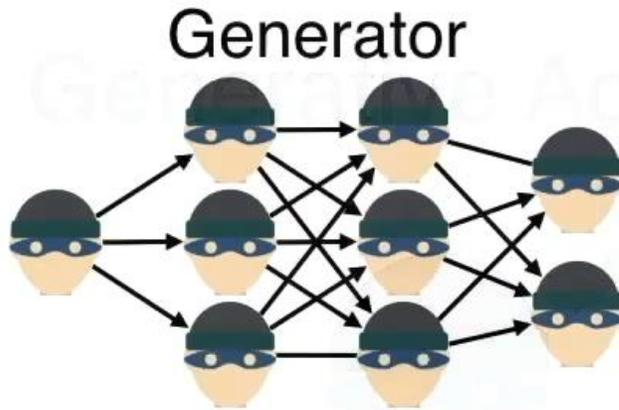
- learn how the world works by observing, like babies.
- learn to predict how one can influence the world through taking actions.
- learn hierarchical representations that allows long-term predictions in abstract representation spaces.
- properly deal with the fact that the world is not completely predictable.
- enable agents to predict the effects of sequences of actions so as to be able to reason and plan
- enable machines to plan hierarchically, decomposing a complex task into subtasks.
- all of this in ways that are compatible with gradient-based learning.

The solution is not just around the corner.

We have a number of obstacles to clear, and we don't know how.

# Intelligenza Artificiale Generativa

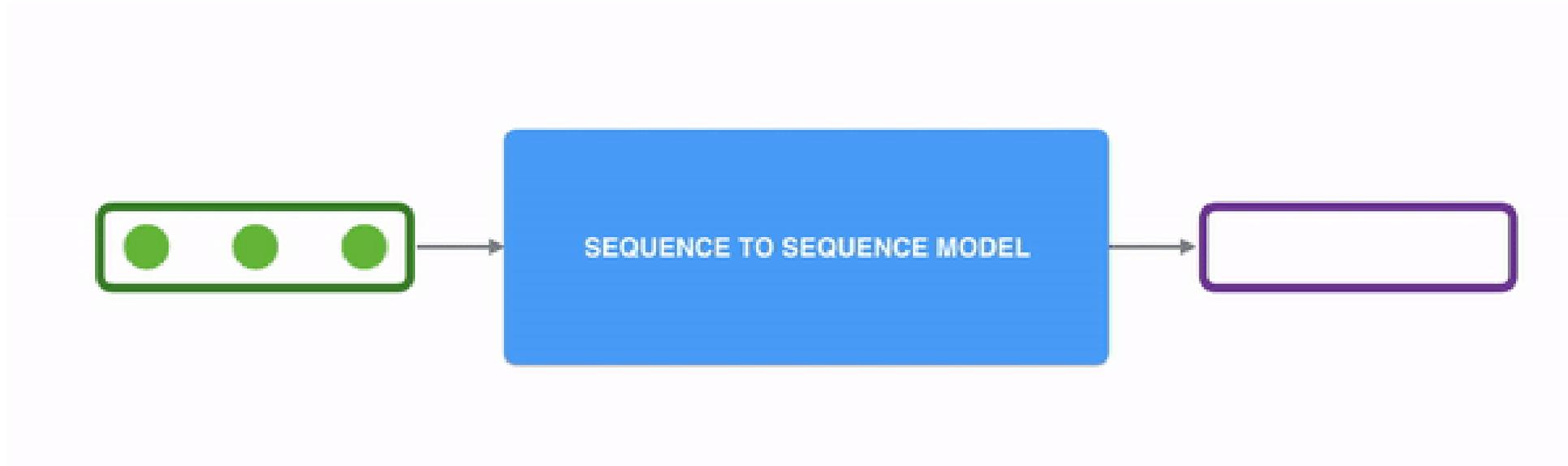
# GAN



Real images



# Transformers



# Generative Pretrained Transformer - GPT



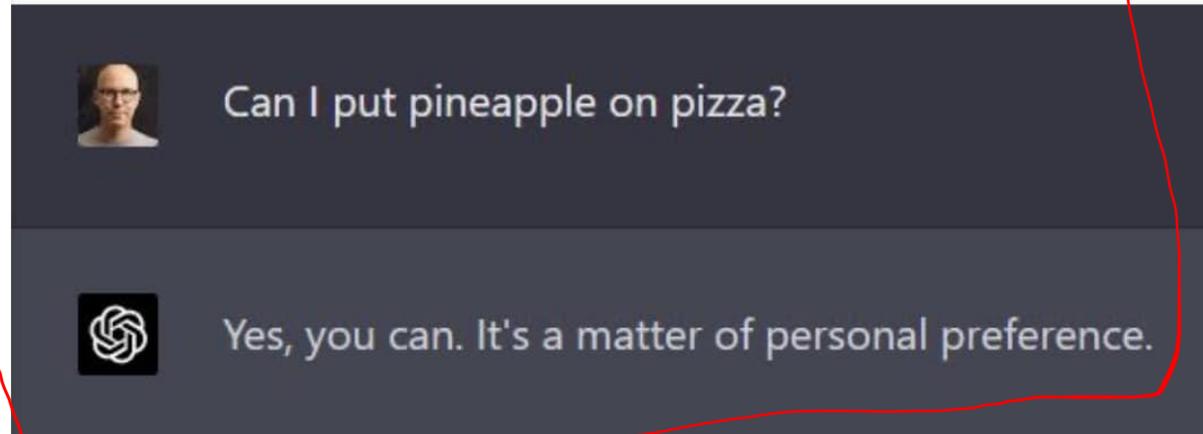
- Text generator basato su Deep Learning addestrato da OpenAI
- Basato su Transformer
- Next word prediction su 300 miliardi di token
- 175 miliardi di parametri (800 GB)
- 12 M\$ per l'addestramento

# NLP



The true reason why Italy banned ChatGPT.

Vedi traduzione



Dona una pizza per tenere aperto PizzaGPT



Ciao, sono PizzaGPT come posso aiutarti?

# Creazione immagini

Spiderman playing with a 4yo child



# Creazione immagini









# ChatGPT

Chatbot basato su GPT (3.5, 4)

- Comprensione del linguaggio umano
- Risposte testuali che imitano quelle di un umano
- Risposte «accurate»: contestualmente appropriate alla richiesta ricevuta



# ChatGPT - Limiti

- Conoscenza ferma al 2021/2022
- Non naviga su internet (anche se... *Browse with Bing*)
- Non sempre dice il vero!
- Limite di parole
- Serve il contesto!



# ChatGPT - Allucinazioni

## Frase «plausibili», ma irrilevanti o errate fattualmente

PA summarise this article <https://www.nytimes.com/2023/03/11/technology/chatgpt-prompts-to-avoid-content-filters.html>



The article discusses how ChatGPT, an AI language model, is being used to circumvent content filters that are designed to detect and block certain types of content. By generating prompts that are less likely to trigger these filters, ChatGPT can be used to create content that might otherwise be censored. While this technology could be useful for those seeking to express themselves freely, it also raises concerns about the potential misuse of AI-generated content, as well as the difficulty of regulating such content in the future. The article concludes by noting that there is still much to learn about the impact of AI language models like ChatGPT on the way we communicate online.

Immagine da: [https://en.wikipedia.org/wiki/Hallucination\\_\(artificial\\_intelligence\)](https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence))

## Bias nei dati

Elementi «sovra-rappresentati» nei dati possono portare output errati (inaspettati?)

Es.: algoritmi usati nel *job recruiting*

- <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- <https://www.forbes.com/sites/forbestechcouncil/2023/09/25/ai-bias-in-recruitment-ethical-implications-and-transparency/?sh=6a34e651799f>

# ChatGPT - Come

1. Nuova conversazione
2. Obiettivo specifico (chiedere di assumere il ruolo di un qualche tipo di esperto)
3. Richiedere un compito il più specifico possibile
4. Condizioni che ChatGPT deve rispettare
5. (Chiedere riguardo la comprensione, per chiarire eventuali punti ambigui)

