

Esempi di Deep Learning ed applicazioni in Python

Cristina Muschitiello

Data Scientist, Statistician , R Developer

@infocamere - @faoun





Agenda di oggi

- 1 INTRODUZIONE: il fenomeno “Big Data”
- 2 FACCIAMO ORDINE: AI, Machine Learning, Neural Networks & Deep Learning
- 3 RETI NEURALI: Una semplificazione
 1. Definizione
 2. Elementi caratterizzanti le reti neurali
- 4 ESEMPI: applicazioni di Deep Neural Networks
- 5 APPLICAZIONE IN PYTHON: costruiamo la nostra prima rete neurale

INTRODUZIONE

Il fenomeno “Big Data”



La maggior parte delle nostre attività quotidiane, oggi, crea dei dati, che possono essere raccolti, analizzati e monetizzati.





Ognuno di noi genera più di 12 GB di dati ogni giorno e il numero è in costante crescita.

2000 AC

Abaco/Biblioteca di Alessandria

XIX secolo Business Intelligence

Analisi e collezione strutturata di dati di business per ottenere vantaggi. (Herman Hollerith, fondatore di IBM, crea il primo sistema di computazione automatica)

XX secolo: Database Relazionali

Nel 1965 nasce il primo data center in USA e soltanto pochi anni dopo uno sviluppatore in IBM crea il primo framework per un database relazionale.

XXI secolo: Big Data
Nel 1999 si legge per la prima volta la parola "Big Data".

....



La rete 5G di Vodafone

Luglio 2019

Dall'evoluzione della Giga Network 4.5G nasce la **Giga Network 5G**, la rete che mette a disposizione dei clienti sempre più capacità e potenza di calcolo, supportando i servizi di nuova generazione e la navigazione in tempo reale. Una rete ancora più flessibile che si adatta alle esigenze dei clienti. La rete che oggi può raggiungere il massimo della performance a Milano, Bologna, Torino, Roma e Napoli con l'obiettivo di estenderlo presto a 100 città italiane.

La velocità potrà arrivare fino a **10 Gbps**, ma soprattutto con una latenza nell'ordine dei **5 ms**, contro i 30 delle reti 4G. Questo aprirà le porte a tutta una serie di nuovi servizi, non solo legati all'ambito gaming. Vodafone ha ribadito infatti la crescita esponenziale non solo del traffico dati, che nel 2022 dovrebbe essere il triplo di quello attuale, ma anche dei sensori IoT, che nel 2025 arriveranno a 50 miliardi su scala globale.



Il fenomeno dei “Big Data” è, quindi, il fenomeno di immagazzinare, gestire e analizzare grandi quantità di dati, tipicamente DESTRUTTURATI



Un dato DESTRUTTURATO è un dato eterogeneo che rispecchia la realtà.



Leva per le aziende nella costruzione di servizi e generazione di profitto.



1 testo

2 Immagini

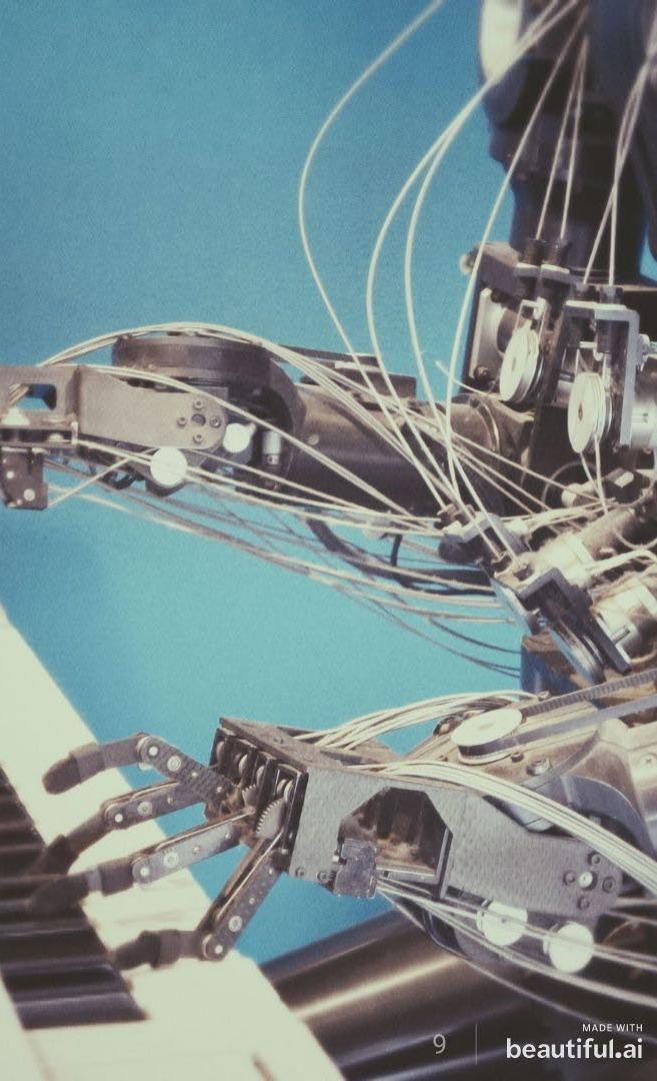
3 Video

4 Audio

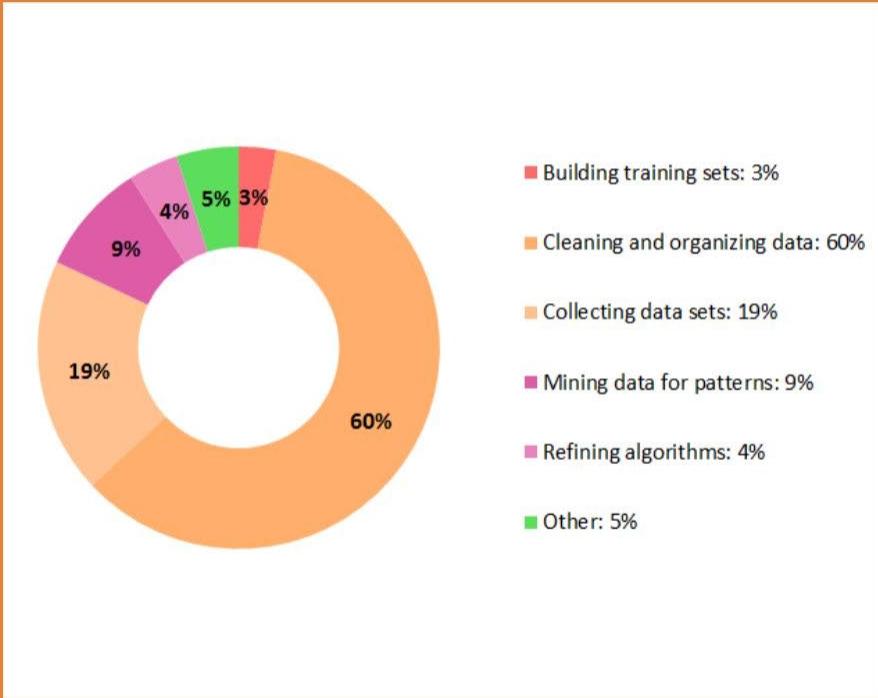
5 Elementi di calcolo

...

Con l'evoluzione del dato sono evolute anche le tecniche di analisi: Machine Learning e Deep Learning.



Aspetti negativi dei BIG DATA



Time consuming

Aspetti negativi dei BIG DATA



Data Storage

Aspetti negativi dei BIG DATA



Privacy

Aspetti negativi dei BIG DATA



Data Security

Aspetti negativi dei BIG DATA



Legislazione

FACCIAMO ORDINE

AI, Machine Learning, Neural Networks & Deep Learning

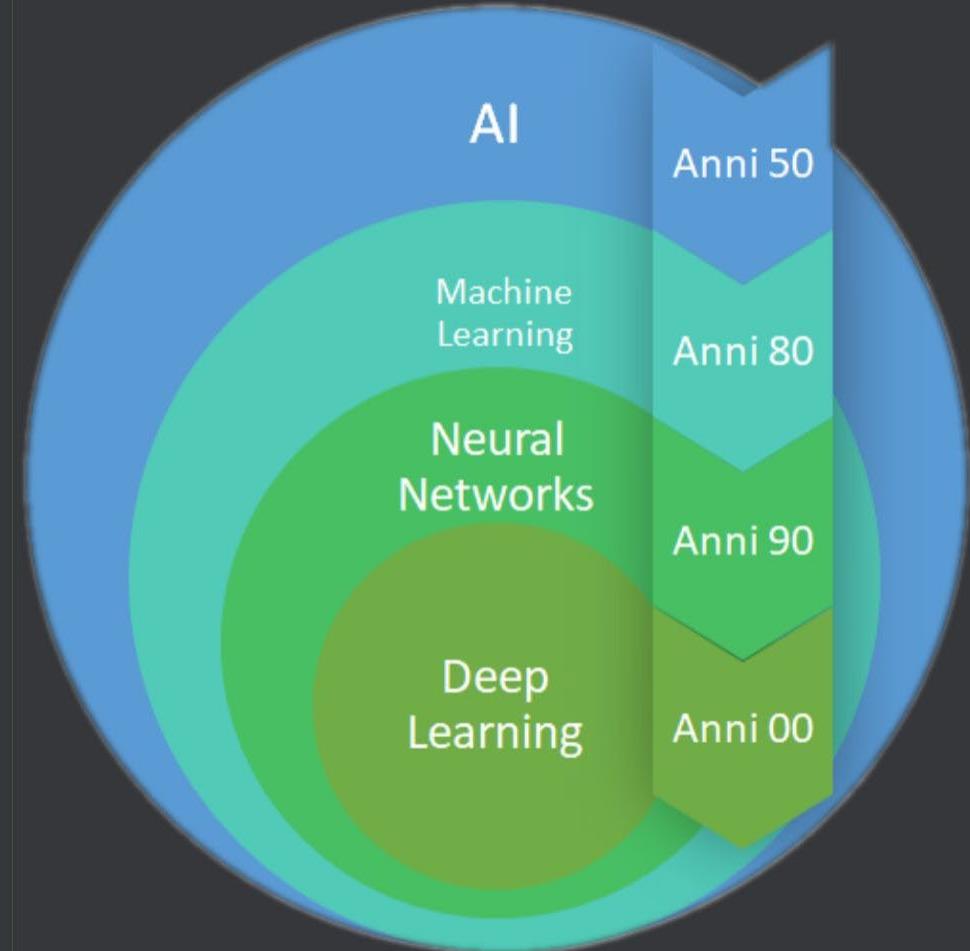


Disciplina Informatica di progettazione di hardware e software che simulano il comportamento umano

Algoritmi di apprendimento automatico per simulare il comportamento umano

Algoritmi di apprendimento automatico che simulano le connessioni neurali del cervello umano

Neural Network profondi con molti strati di apprendimento.

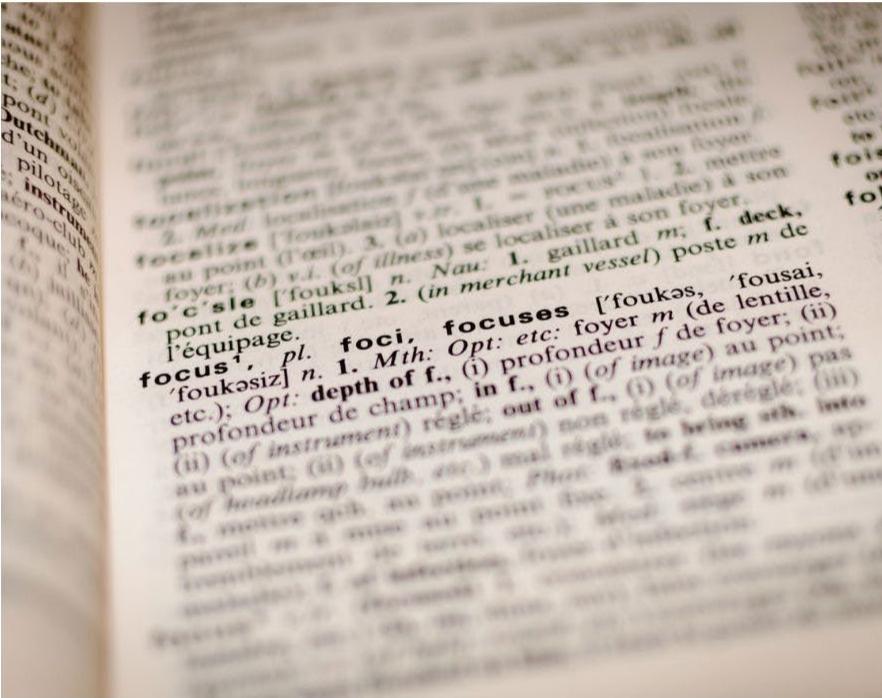


RETI NEURALI

Una semplificazione



1. Definizione

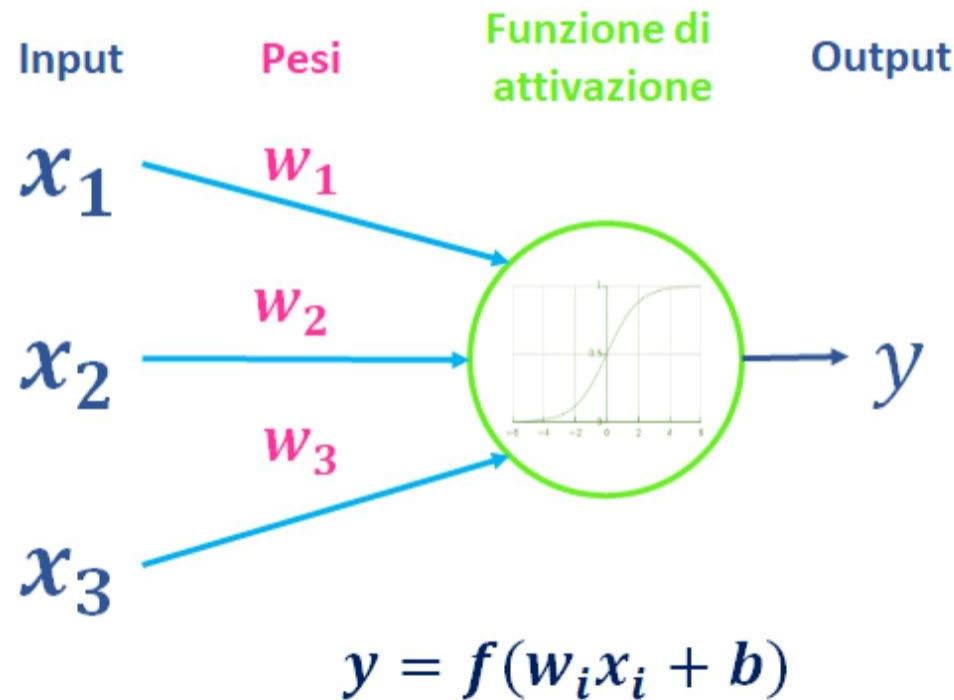


Una rete neurale è un
algoritmo
computazionale per la
trasformazione di INPUT
in OUTPUT

IL NEURONE o NODO

Una rete neurale gira intorno al concetto di **neurone o nodo**.

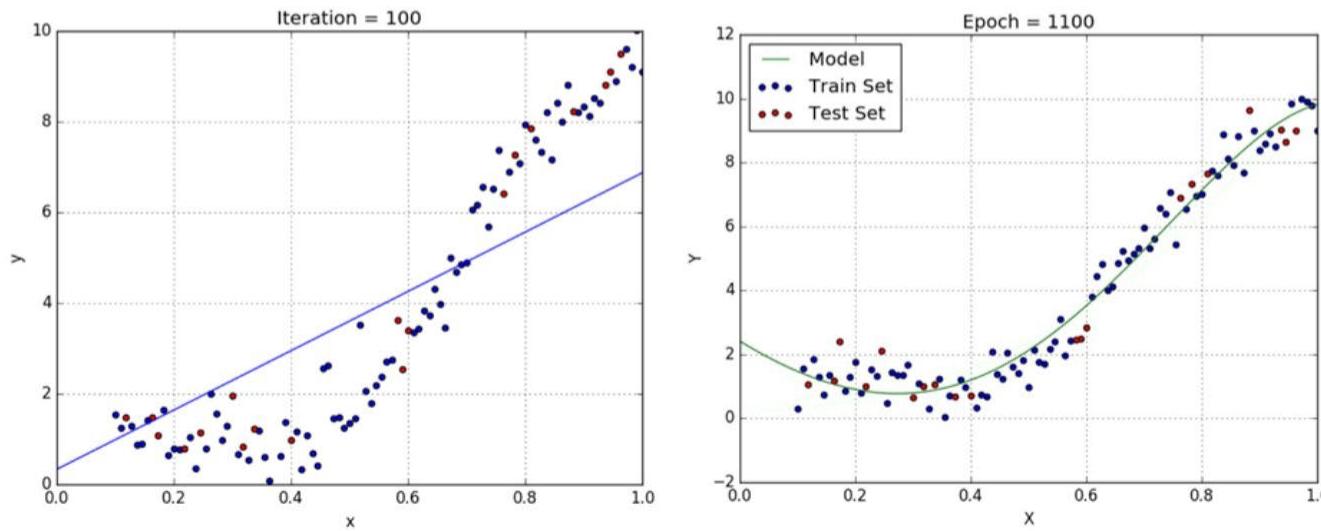
Un **nodo** è l'unità computazionale di una rete ed è una combinazione lineare pesata che trasforma un input in un output per il tramite di una **FUNZIONE DI ATTIVAZIONE** non lineare



Una rete neurale senza funzione di attivazione equivale semplicemente a un modello di regressione, ovvero cerca di approssimare la distribuzione dei dati con una retta

Lo scopo delle reti neurali è di rappresentare funzioni di approssimazione universali, cioè essere in grado di approssimare qualsiasi funzione. Per fare questo è necessario introdurre un fattore di non linearità.

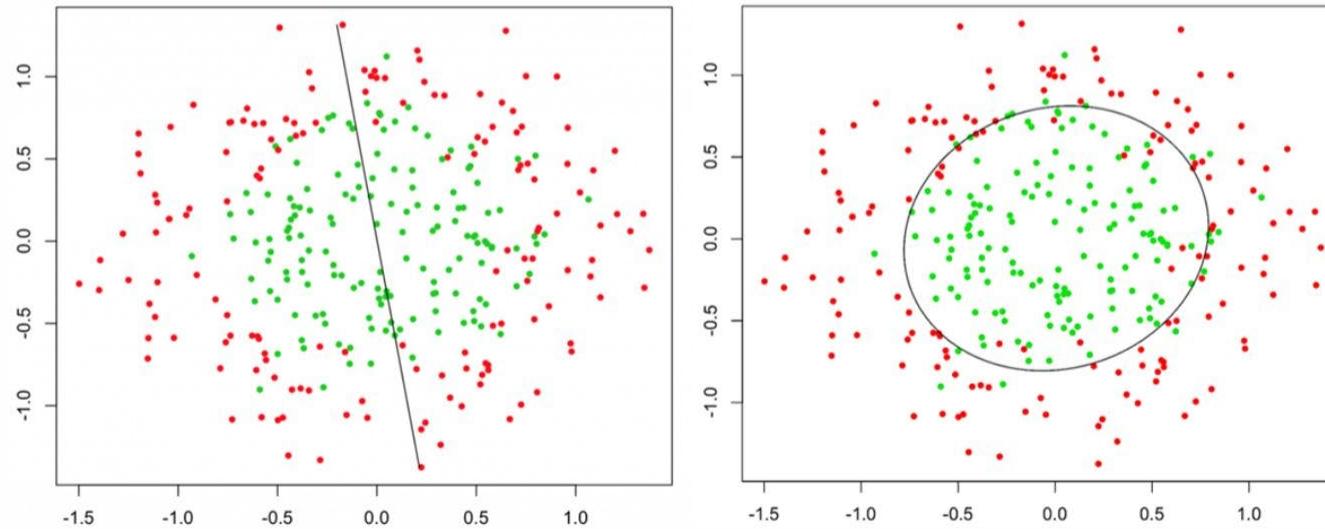
Questo fattore è rappresentato dalle funzioni di attivazione.



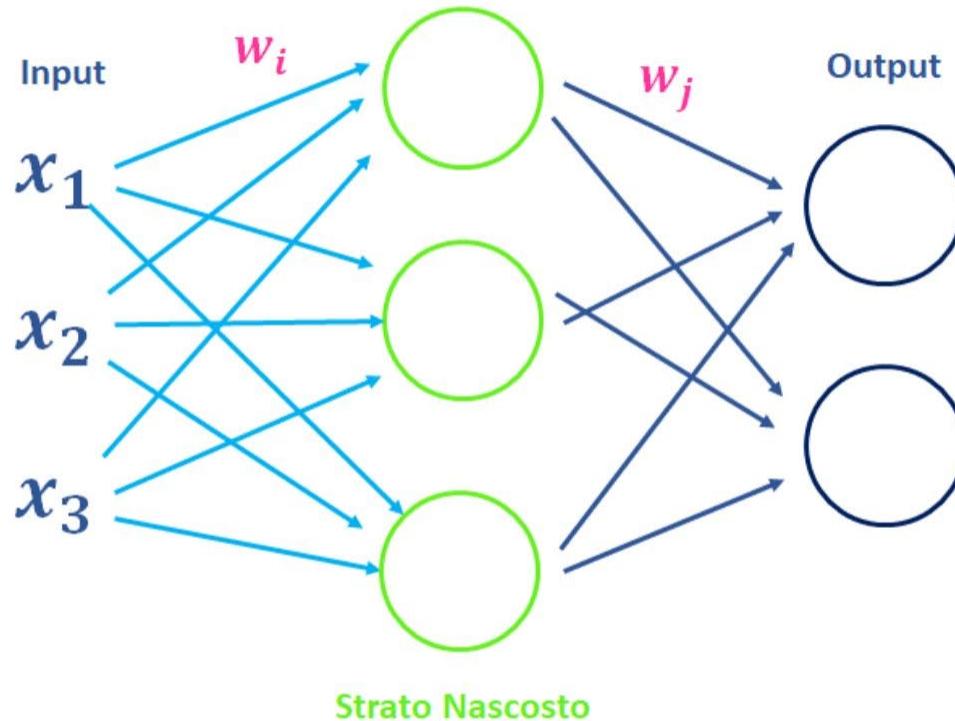
Una rete neurale senza funzione di attivazione equivale semplicemente a un modello di regressione, ovvero cerca di approssimare la distribuzione dei dati con una retta

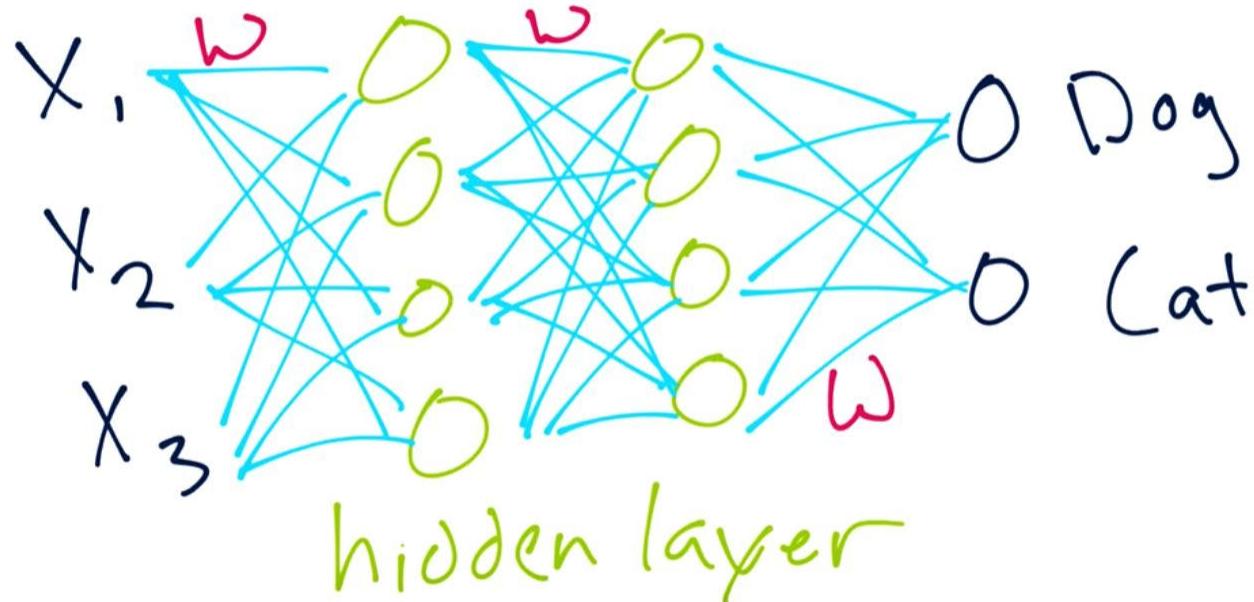
Lo scopo delle reti neurali è di rappresentare funzioni di approssimazione universali, cioè essere in grado di approssimare qualsiasi funzione. Per fare questo è necessario introdurre un fattore di non linearità.

Questo fattore è rappresentato dalle funzioni di attivazione.



Gli Strati Nascosti





A photograph of a field of yellow flowers, likely canola or mustard, with a single red poppy flower in sharp focus in the foreground. The background is blurred.

2. Gli elementi che caratterizzano le RETI NEURALI

- 1 Funzione di attivazione
- 2 Pesi
- 3 Scelta dell'architettura



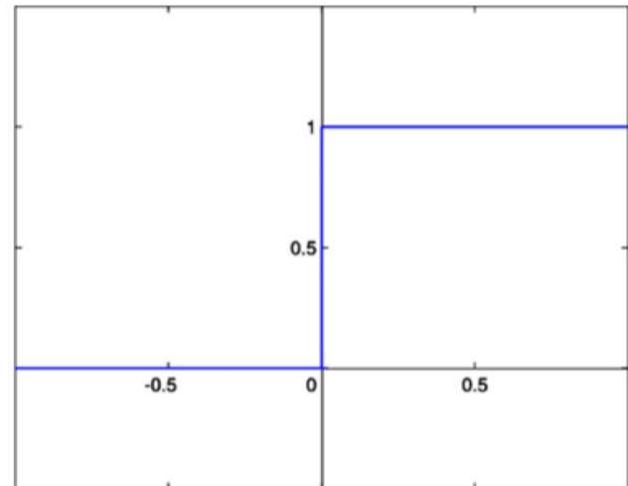
A photograph of a field of yellow flowers, likely canola or mustard, with a single red poppy flower in sharp focus in the foreground. The background is blurred.

2. Gli elementi che caratterizzano le RETI NEURALI

- 1 Funzione di attivazione
- 2 Pesi
- 3 Scelta dell'architettura

Funzione a Gradino

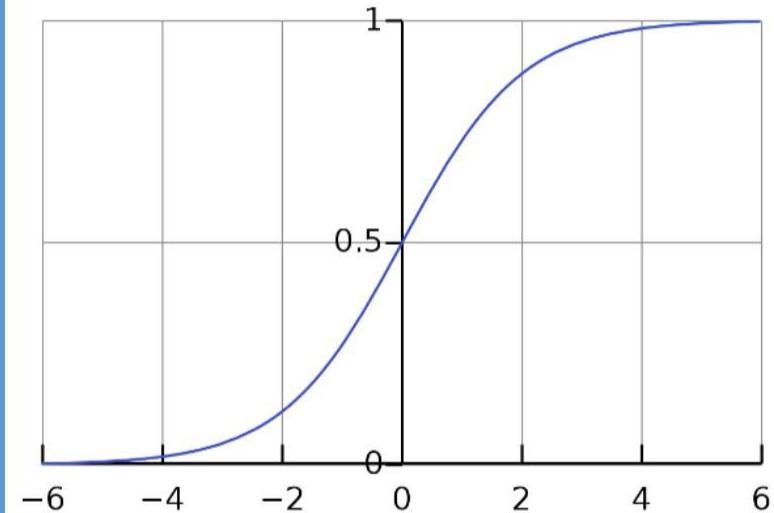
La funzione più intuitiva è quella a gradino, in un certo senso più simile al funzionamento biologico. Per tutti i valori negativi risposta rimane 0, mentre salta a +1 appena il valore raggiunge o supera lo zero anche di poco. Il vantaggio è che è semplice da calcolare, e “normalizza” i valori di output, comprimendoli tutti in un range compreso tra 0 e +1.

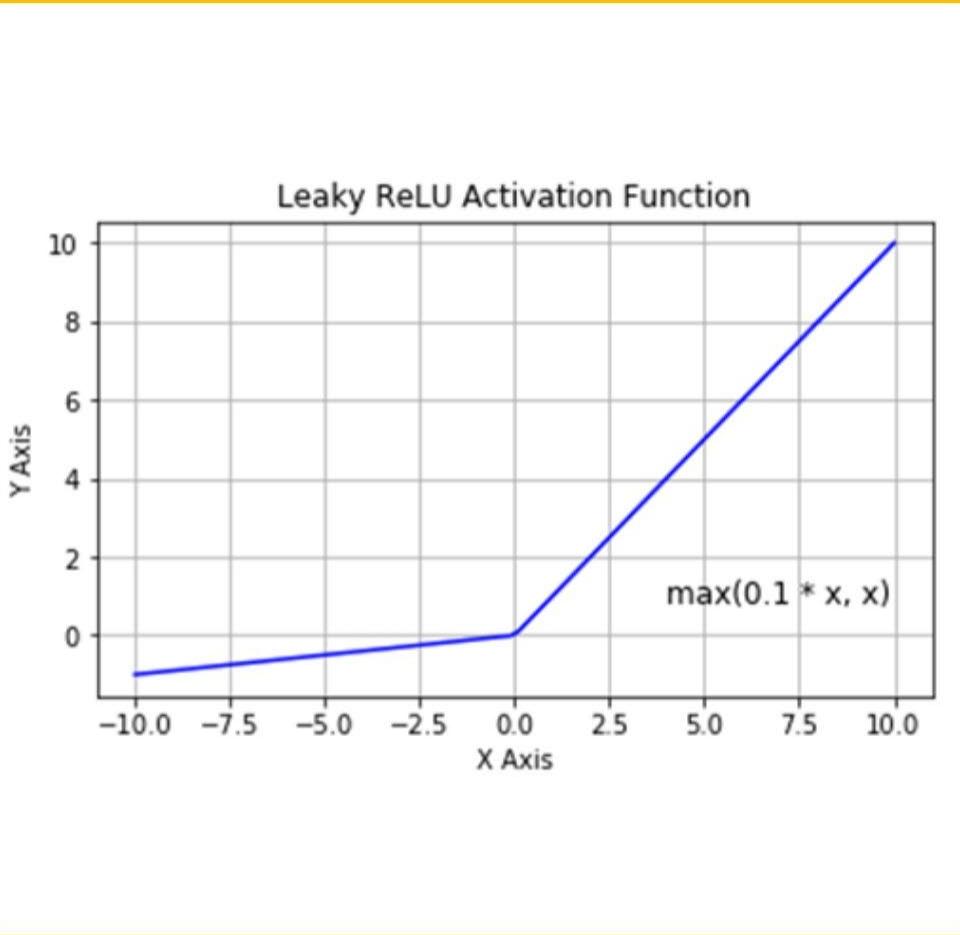


Funzione a gradino

Funzione Sigmoide

Ha delle similitudini con quella a gradino, ma il passaggio da 0 a +1 è più graduale, con un andamento a forma di "s", appunto.





Funzione ReLU

La funzione ReLU (Rectifier Linear Unit) è una funzione ultimamente divenuta molto utilizzata, soprattutto nei layer intermedi. Si tratta di una funzione molto semplice da calcolare: appiattisce a zero la risposta a tutti i valori negativi, mentre lascia tutto invariato per valori uguali o superiori a zero.



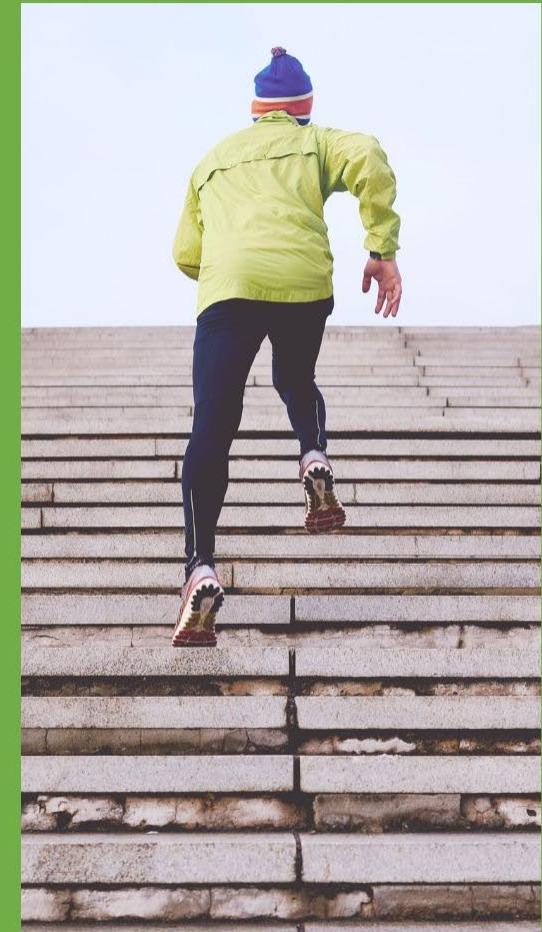
2. Gli elementi che caratterizzano le RETI NEURALI

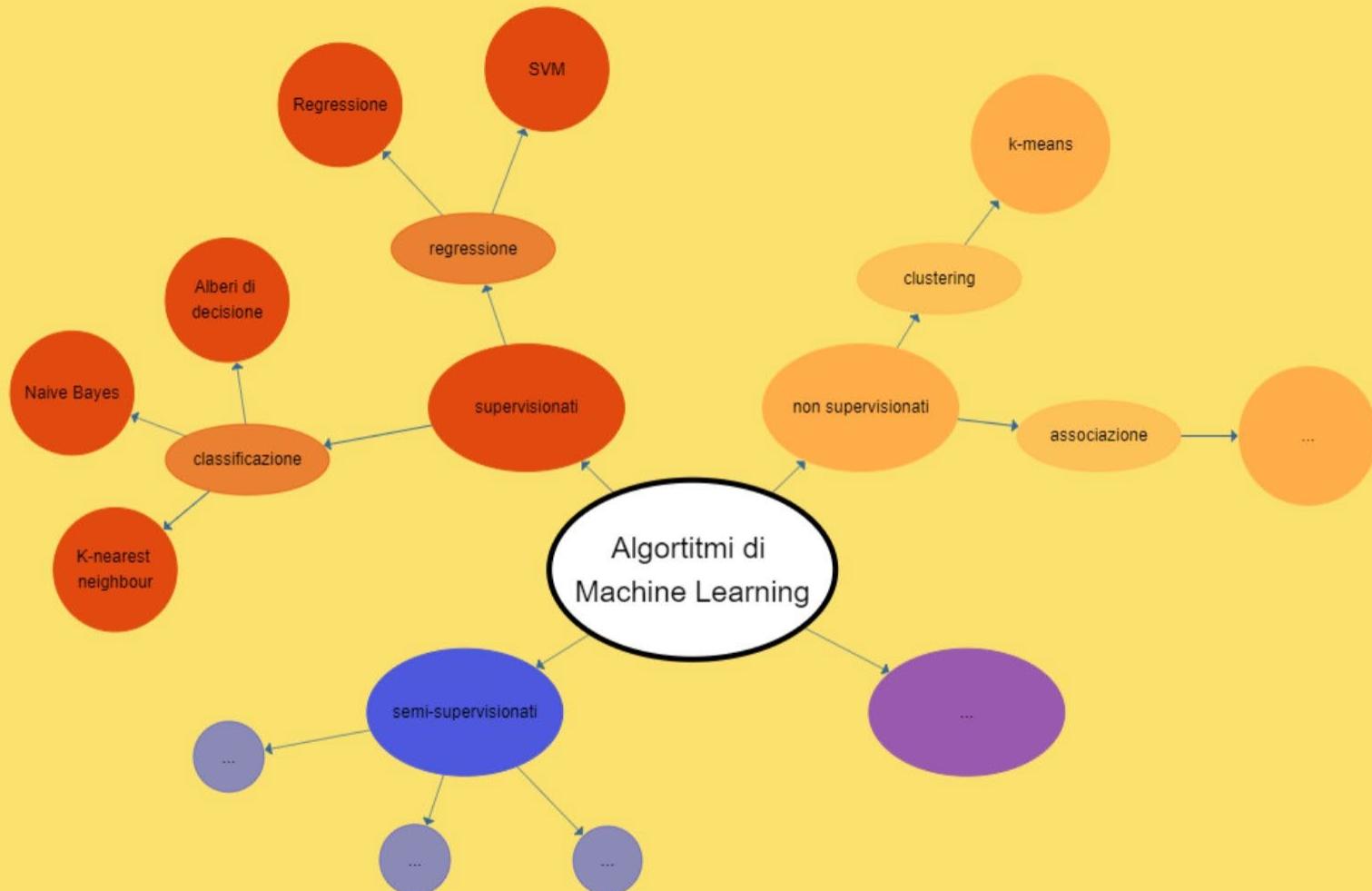
- 1 Funzione di attivazione
- 2 Pesi: allenamento del modello
- 3 Scelta dell'architettura

Determinazione dei pesi

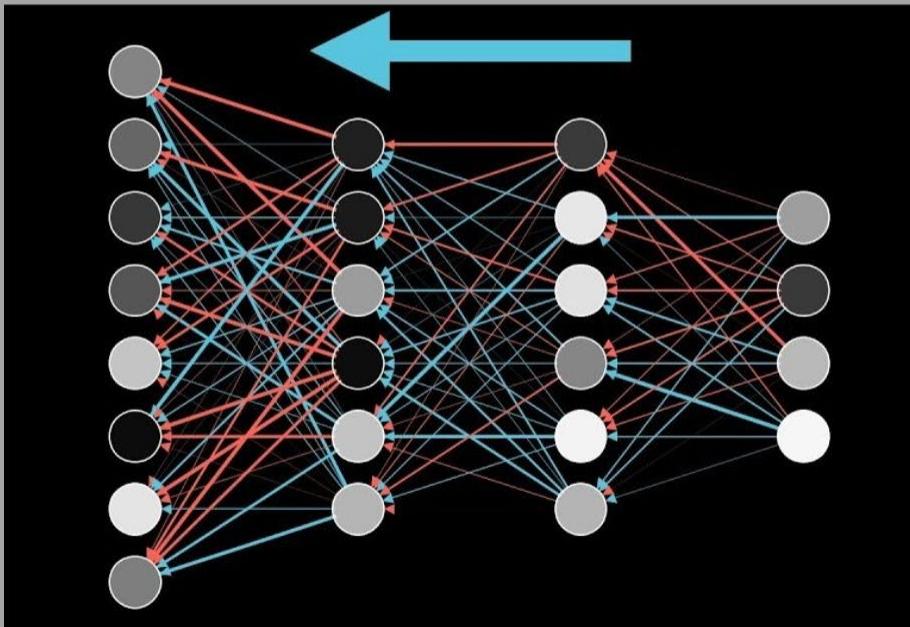
ALLENARE IL MODELLO

- 1 Algoritmo di apprendimento
- 2 Funzione di costo
- 3 Regolarizzazione
- 4 Ottimizzazione
- 5 Architettura





L'algoritmo di apprendimento del deep learning: la Backpropagation



- algoritmo di apprendimento supervisionato
- algoritmo iterativo
- Parte da valori random dei pesi
- Calcola l'output
- lavora a ritroso e ricalcola ad ogni iterazione i pesi in modo da ridurre al minimo l'errore determinato nell'output ->

Riduzione dell'errore significa:

Minimizzazione di una funzione di costo/loss

- Errore Quadratico medio (MSE) (regressione)
- Errore Assoluto medio (AME) (regressione)
- Cross Entropy Loss (classificazione)
- ...

Ottimizzazione e regolarizzazione

- Gradient Descent
- ADAM = Adaptive Moment estimation
- ...





2. Gli elementi che caratterizzano le RETI NEURALI

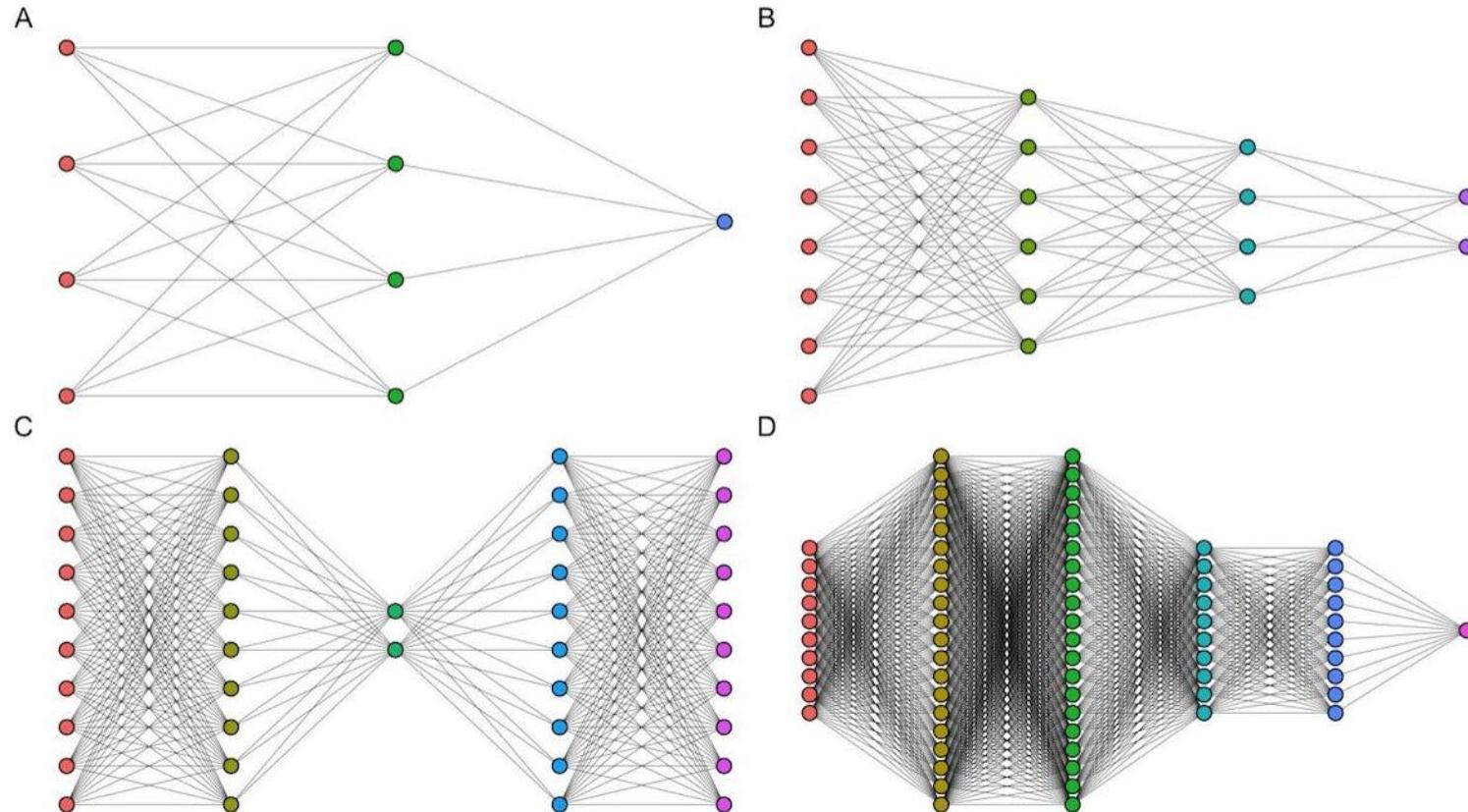
- 1 Funzione di attivazione
- 2 Pesi: allenamento del modello
- 3 Scelta dell'architettura

Numero di layers: non esiste una regola

- 1 Si può *cominciare* con pochi strati e aumentarne progressivamente il numero
- 2 Per problemi più semplici di solito bastano un paio di Strati
- 3 Per problemi più complessi si può arrivare ad utilizzare dozzine di layers.

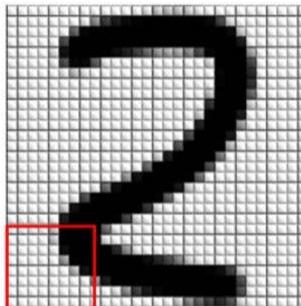


Diverse architetture delle reti neurali (->)



Un esempio concreto: i dati MNIST

immagine in scala di grigi

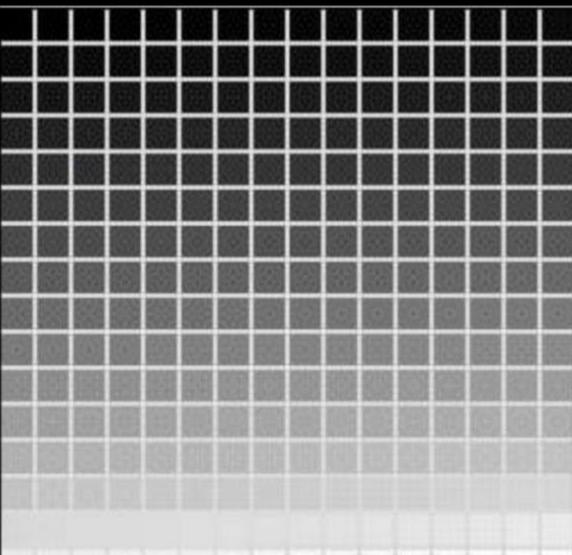


| | | | | | | | | |
|---|---|---|---|---|----|-----|-----|-----|
| 0 | 0 | 0 | 0 | 0 | 85 | 210 | 255 | 255 |
| 0 | 0 | 0 | 0 | 0 | 92 | 245 | 255 | 255 |
| 0 | 0 | 0 | 0 | 0 | 0 | 245 | 255 | 255 |
| 0 | 0 | 0 | 0 | 0 | 0 | 175 | 255 | 255 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 182 | 255 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 4 | 1 | 9 | 2 | 1 | 3 | 1 | 4 | 3 |
| 5 | 3 | 6 | 1 | 7 | 2 | 8 | 6 | 9 | 4 |
| 0 | 9 | 1 | 1 | 2 | 4 | 3 | 2 | 7 | 3 |
| 8 | 6 | 9 | 0 | 5 | 6 | 0 | 7 | 6 | 1 |
| 8 | 7 | 9 | 3 | 9 | 8 | 5 | 9 | 3 | 3 |
| 0 | 7 | 4 | 9 | 8 | 0 | 9 | 4 | 1 | 4 |
| 4 | 6 | 0 | 4 | 5 | 6 | 1 | 0 | 0 | 1 |
| 7 | 1 | 6 | 3 | 0 | 2 | 1 | 1 | 7 | 9 |
| 0 | 2 | 6 | 7 | 8 | 3 | 9 | 0 | 4 | 6 |
| 7 | 4 | 6 | 8 | 0 | 7 | 8 | 3 | 1 | 5 |

$$8 \text{ bit} \rightarrow 2^8 = 256$$

I valori 0-255 che indicano il
livello di grigio vengono
normalizzati tra 0 e 1
dividendo per 255

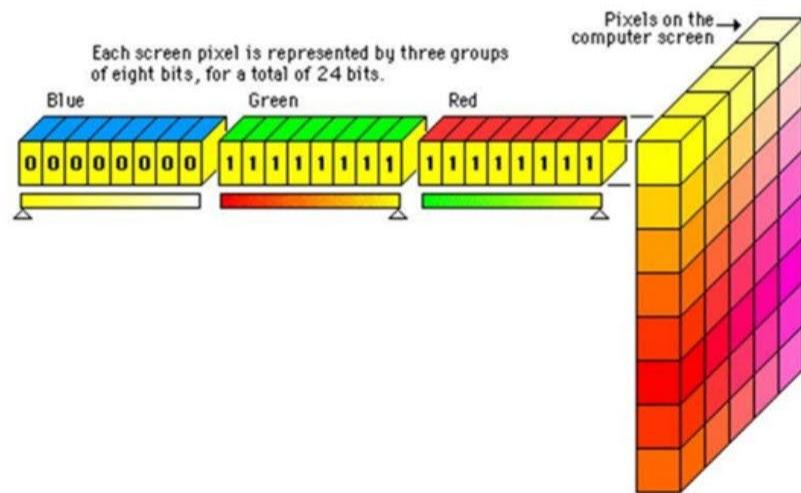


Immagina i colori RGB

RGB 24-bit

canale R: $2^8=256$
canale G: $2^8=256$
canale B: $2^8=256$

} 16.777.216 colori
→ 8 bit x 3 canali = 24-bit





64

64

| → Red | → Green | → Blue | |
|-------|---------|--------|-----|
| 255 | 134 | 93 | 22 |
| 255 | 231 | 42 | 22 |
| 123 | 94 | 83 | 2 |
| 34 | 44 | 187 | 92 |
| 34 | 76 | 232 | 124 |
| 67 | 83 | 194 | 202 |

These tables are only 5 by 5!

Image 64 by 64

y =

- 1: Cat
- or
- 0: Non cat

x = Array con dimensione (64, 64, 3)

x =

| |
|-----|
| 255 |
| 231 |
| 42 |
| .. |
| 255 |
| 134 |
| .. |
| .. |
| 255 |
| 134 |
| .. |
| .. |
| 123 |
| 94 |
| .. |
| .. |
| 142 |

Tensor with all data from the 3 matrices.

Length:
64x64x3
= 12.288 = n_x

64x64 pixels by 3 RGB channels



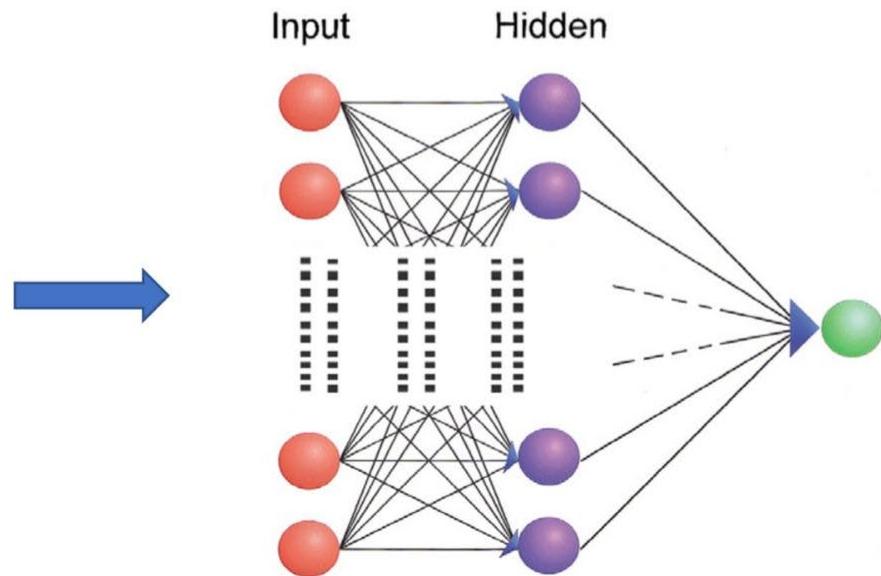
$$64 \times 64 \times 3 = 12,288$$
$$1000 \times 1000 \times 3 = 3,000,000$$

Input data for each image

With a fully connected network, if we have 1000 hidden units, then the total number of weights is 3 billion!



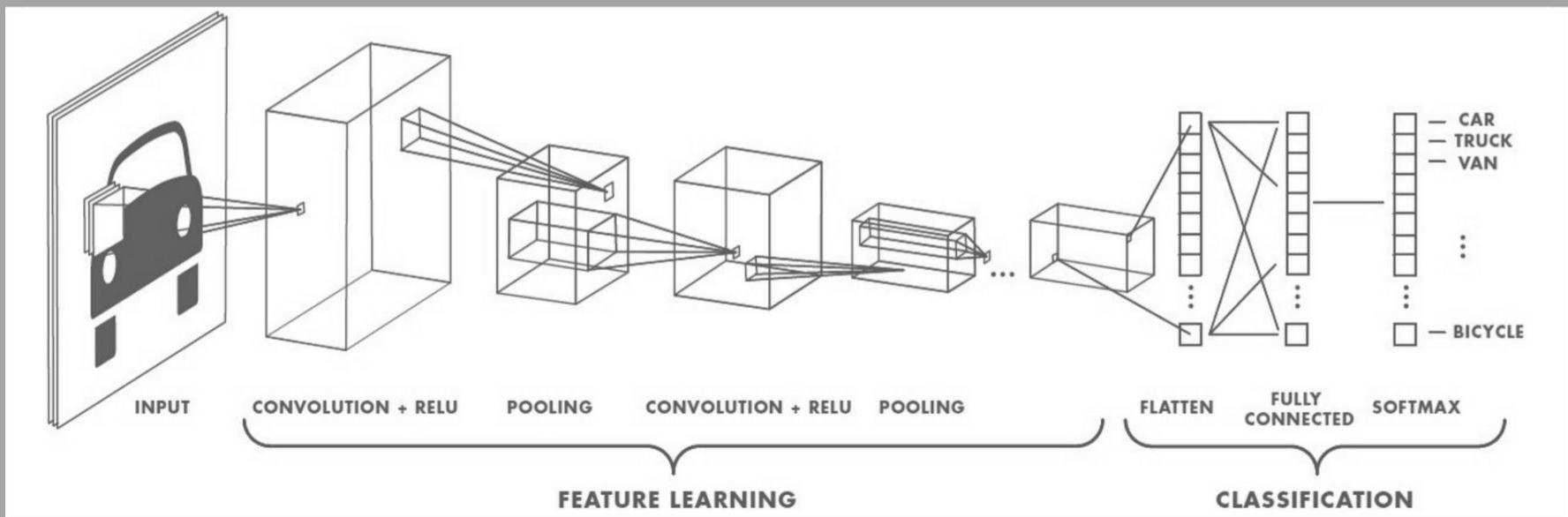
1000x1000 pixels by 3 RGB channels



LE RETI NEURALI PIU' "FAMOSE"

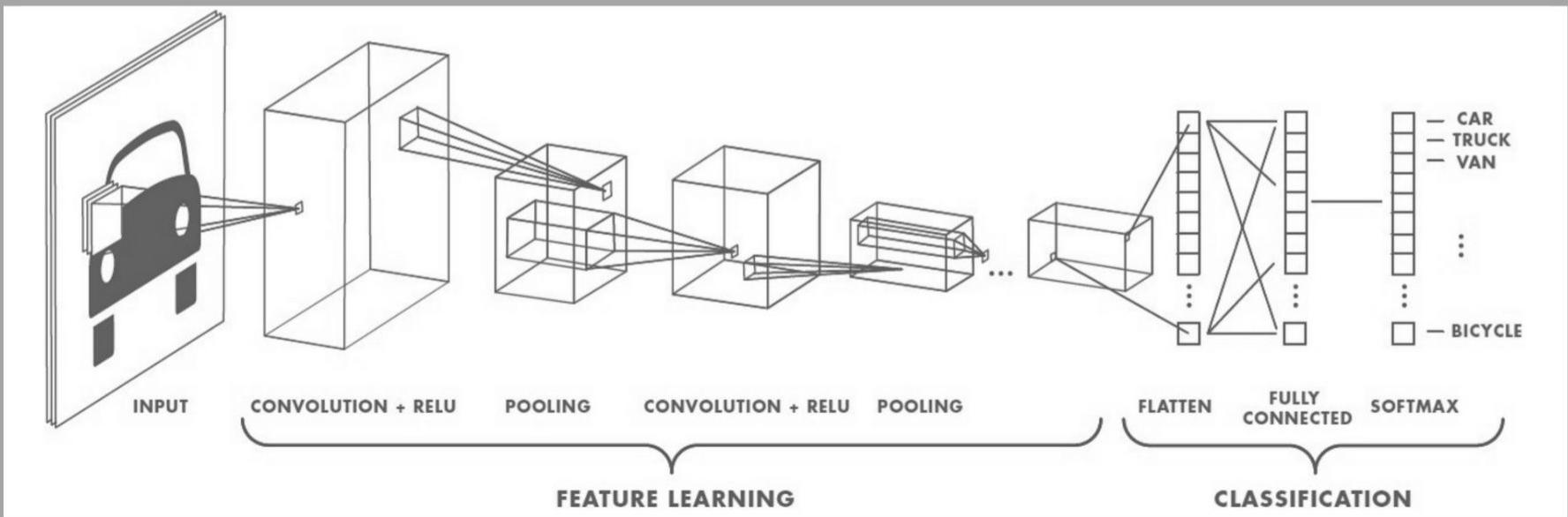
CONVOLUTIONAL NEURAL NETWORKS

Usate nella classificazione delle immagini. Mappare tutti i pixel di un'immagine è computazionalmente molto costoso. Le CNN semplificano il calcolo in larga misura senza perdere l'essenza dei dati. La convoluzione è fondamentalmente la moltiplicazione di matrici e la somma di tali risultati.



Ciascuno strato rappresenta una mappa di alcune caratteristiche dell'immagine.

Gli strati passano attraverso strati di attivazione ed aggregazione prima di essere incanalate in strati di classificazione.



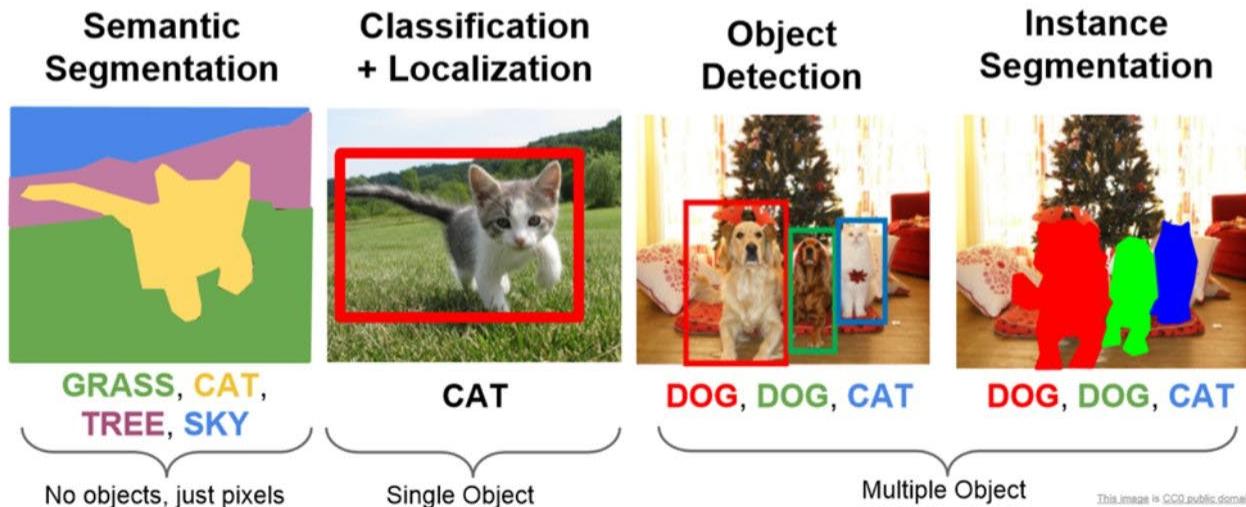
ESEMPI

Applicazioni del deep Learning



Computer Vision (CNN)

Lo scopo principale della visione artificiale è quello di riprodurre la vista umana. Il risultato dell'elaborazione è il riconoscimento di determinate caratteristiche dell'immagine per varie finalità di controllo, classificazione, selezione, ecc.



Amazon Rekognition

Servizio di riconoscimento delle immagini basato sull'apprendimento profondo che rileva oggetti, scene e volti, estrae il testo, riconosce celebrità e individua i contenuti inappropriati nelle immagini. Consente anche di ricercare e confrontare i volti.

Il servizio restituisce un punteggio di affidabilità per tutti gli oggetti che è in grado di identificare, consentendo di prendere decisioni informate su come utilizzare i risultati. Inoltre, tutti i volti rilevati vengono evidenziati con una cornice rettangolare, in modo da risalire alla posizione del volto all'interno delle immagini.

Amazon Rekognition

Analisi intelligente di immagini e di video per applicazioni

RILEVAMENTO DI OGGETTI E SCENE

PERSON
99.3%

CREST
83.0%

OUTDOORS
83.1%

MOUNTAIN BIKE
99.1%

ROCK
82.8%



RICONOSCIMENTO FACCIALE

PERSON 2

PERSON 1

PERSON 4

PERSON 3

RICONOSCIMENTO DI SOGGETTI IN
MOVIMENTO, ANCHE SENZA VEDERNE IL VOLTO



Female
100%

Eyes are open
100%

Happy
97.4%

Smiling
100%

Auto a guida autonoma: Waymo (google)

Video

Auto a guida autonoma - la tecnologia

Video

Robotica: Boston Dynamics ...

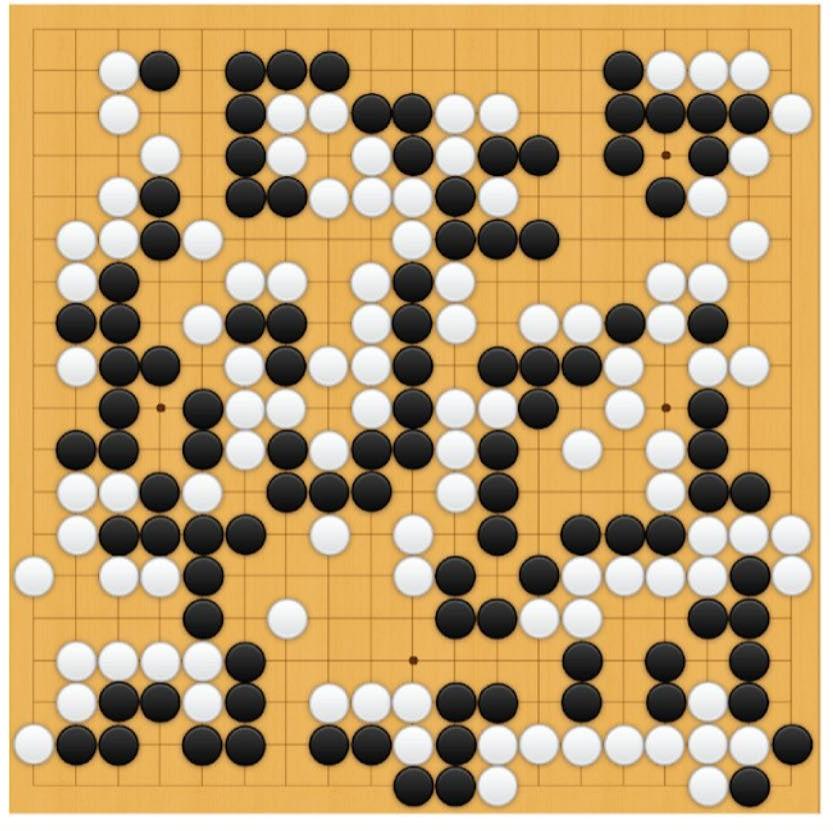
Video

Robotica: PEPPER

Utilizzato come receptionist in diversi uffici nel Regno Unito
in grado di identificare i visitatori con l'uso del riconoscimento facciale, inviare avvisi per gli organizzatori
di riunioni e organizzare la preparazione di bevande.

Il prezzo ai privati è di circa 1.650 \$

Video



THE ULTIMATE GO CHALLENGE

GAME 3 OF 3

27 MAY 2017



AlphaGo

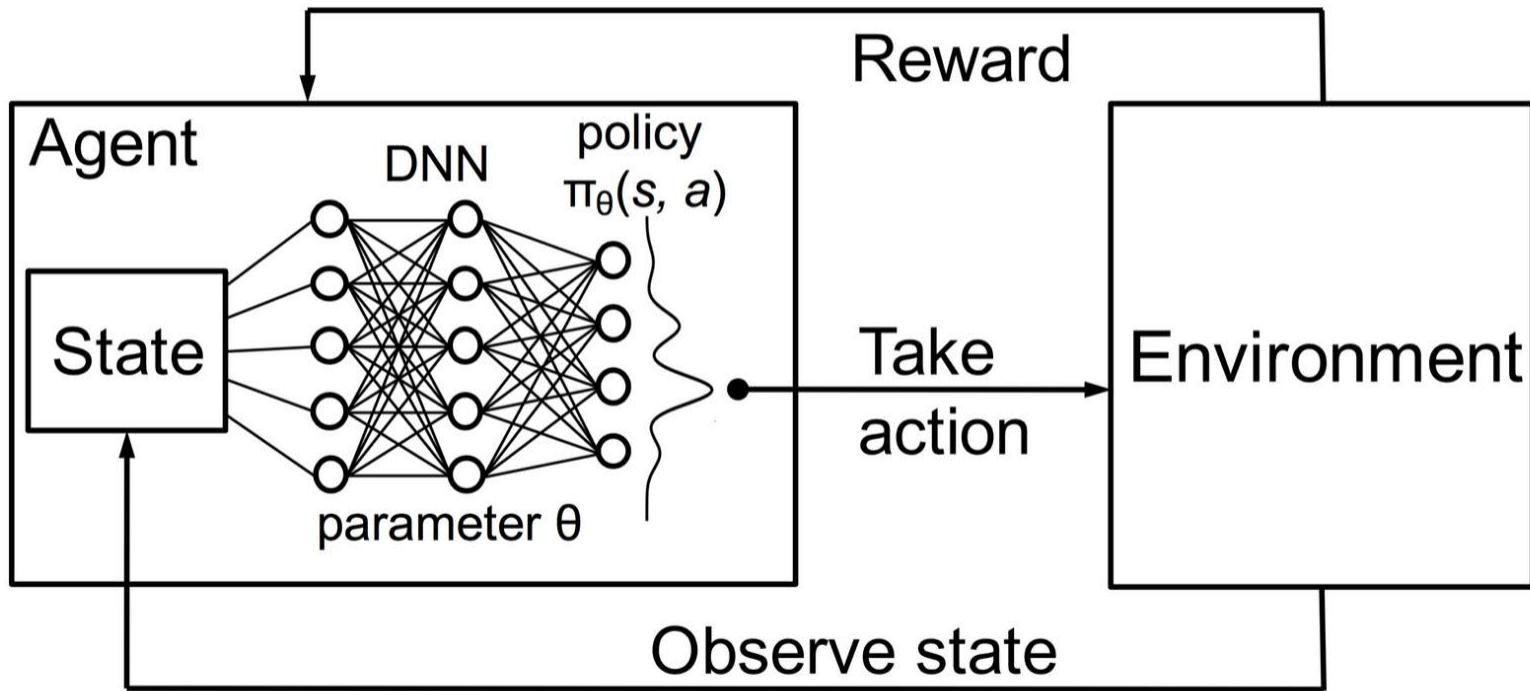
Winner of Match 3

Ke Jie

RESULT B + Res

AlphaGo

L'ALGORITMO DI REINFORCEMENT LEARNING DI ALPHAGO

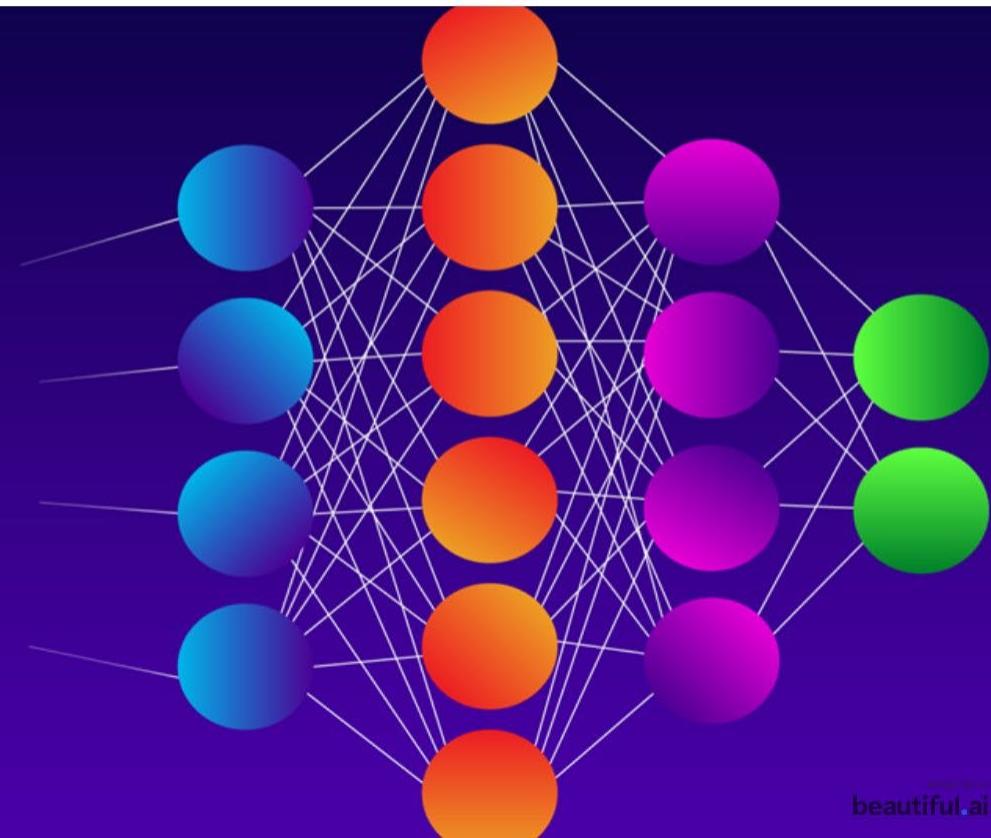


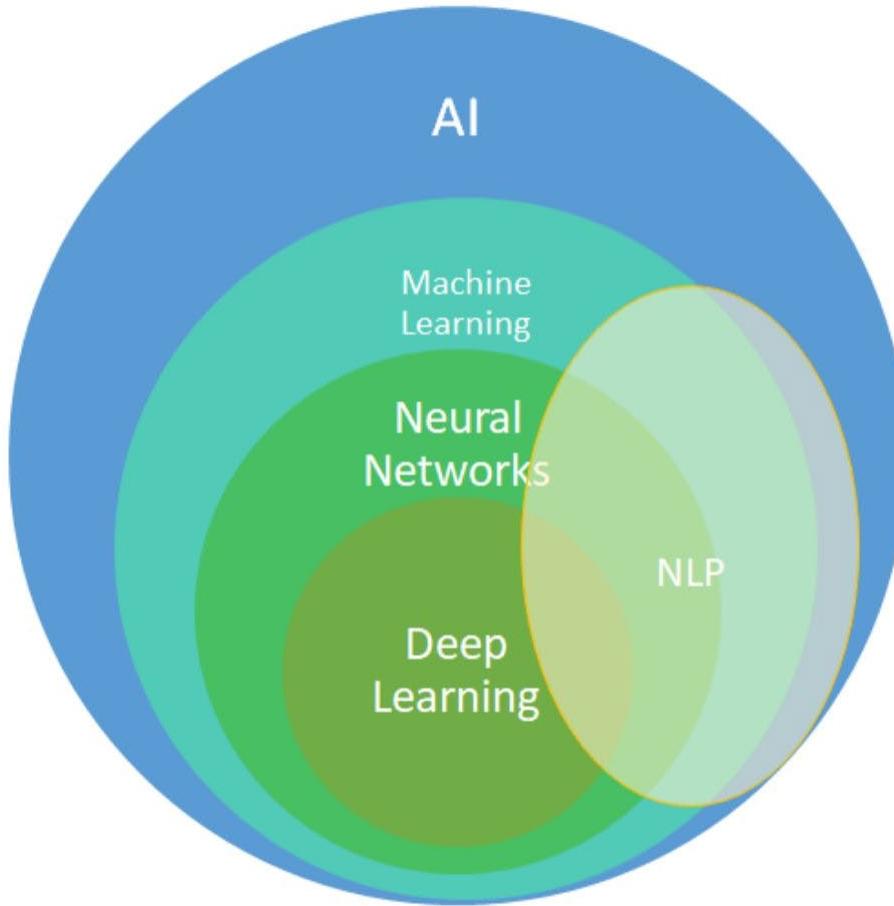
Natural Language Processing (NLP)

ChatBots

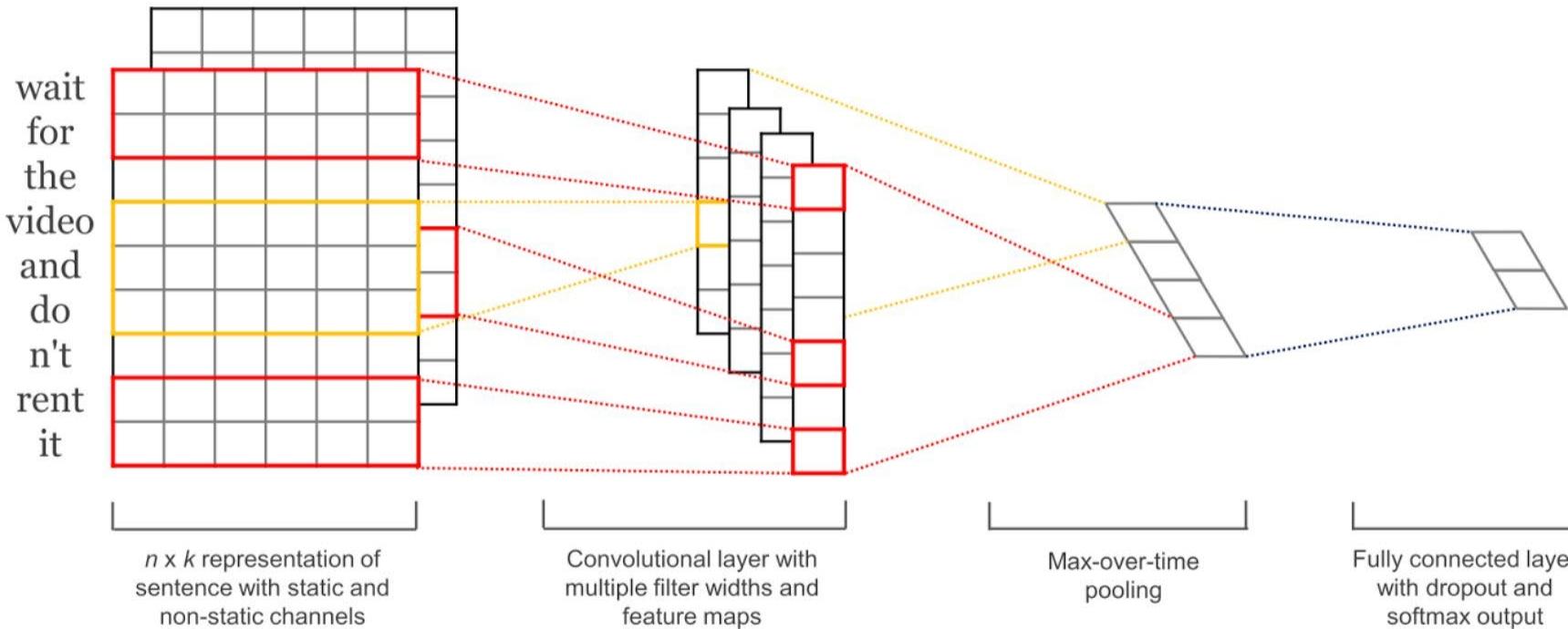
Alexa

...





Es: Sentiment Analysis



Applicazioni controverse



Riconoscimento Facciale per monitorare l'attenzione
dei bambini nelle scuole cinesi



Sistema di riconoscimento di volti o azioni sospette
per le strade di Mosca

Divertiamoci un po'
il Playground di tensorflow



APPLICAZIONE

Costruiamo la nostra prima rete neurale in python: i dati MNIST



python™

Must know



Libreria open source per la costruzione di modelli di machine learning su larga scala.

È di gran lunga la biblioteca più popolare per la costruzione di modelli di deep learning. Ha anche la più forte ed enorme comunità di sviluppatori, ricercatori e collaboratori.



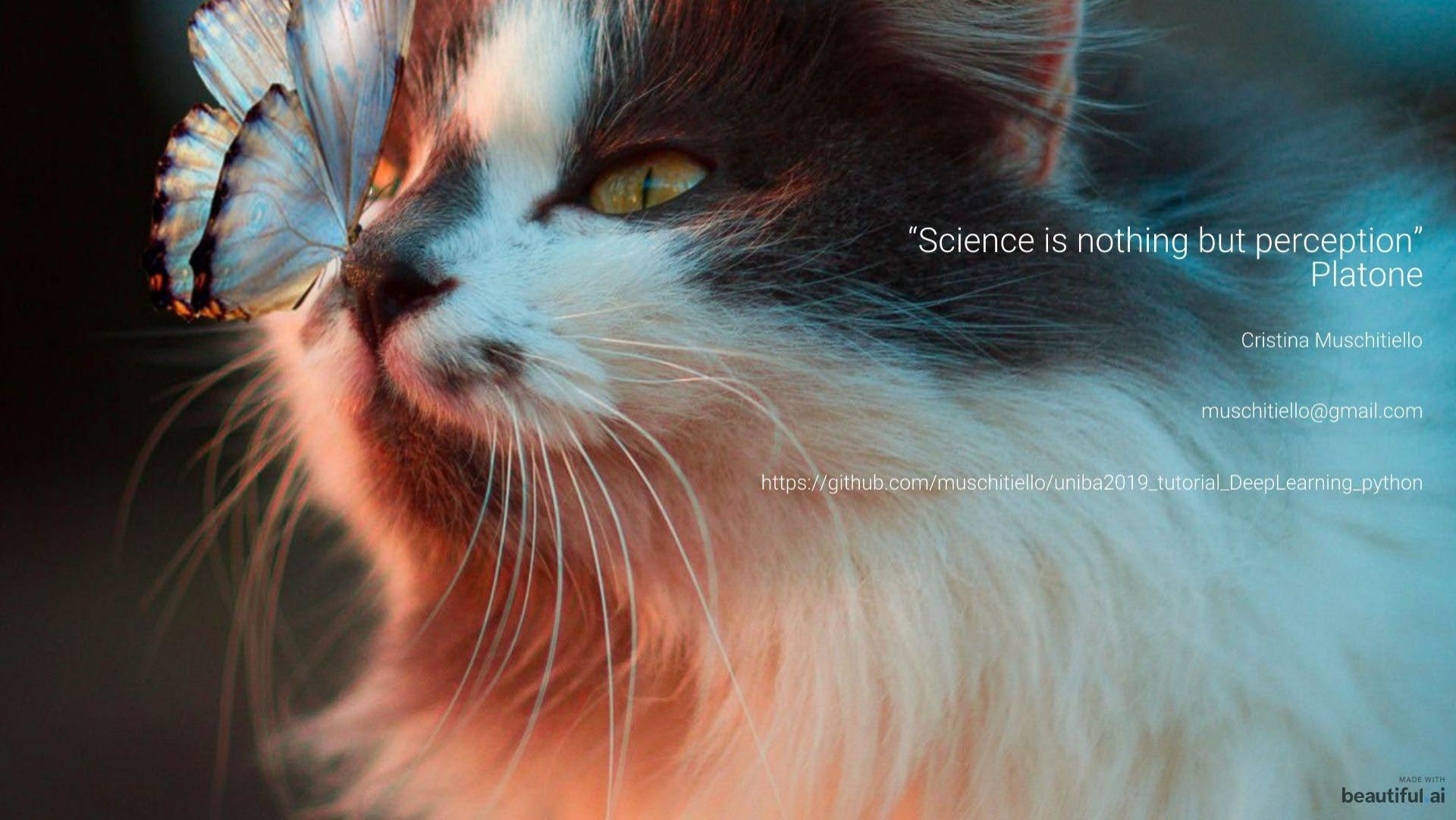
API (application program interface) di reti neurali di alto livello, scritte in Python e in grado di funzionare su Tensorflow, Theano o CNTK.

Supporta la sperimentazione rapida, la prototipazione e API intuitive.



Esempio di riconoscimento di immagini con l'algoritmo inception_v3

colab



"Science is nothing but perception"
Platone

Cristina Muschitiello

muschitiello@gmail.com

https://github.com/muschitiello/uniba2019_tutorial_DeepLearning_python