

SUPSI

Machine Learning per il Multimedia Processing

Fondamenti di Multimedia Processing

Tiziano Leidi

18.01.2019

Machine learning

Il machine learning è un campo dell'informatica con l'obiettivo di sviluppare funzionalità per rendere i computers capaci di imparare senza che li si programmi in maniera esplicita.

È l'evoluzione degli studi sul riconoscimento di pattern e sulle teorie di computational learning dell'intelligenza artificiale.

Il machine learning esplora lo studio e la realizzazione di algoritmi che imparano dai dati in modo da poter fare successivamente previsioni su di essi.

Machine learning

In varie situazioni gli algoritmi di machine learning si rivelano più efficaci di programmi scritti a mano, perché producono autonomamente un modello a partire da dati campione.

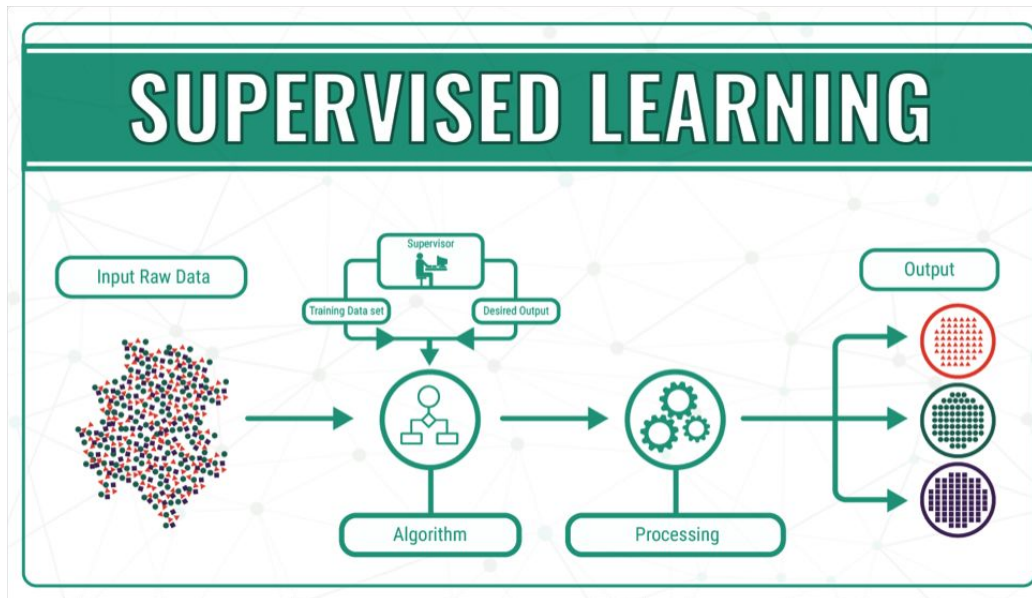
Viene impiegato soprattutto in domini applicativi in cui il design e la programmazione esplicita di algoritmi con buone performances è difficile o addirittura impossibile.

Machine learning

La realizzazione di buoni algoritmi di machine learning è però un compito complesso perché riconoscere i patterns può essere difficile e spesso i dati disponibili non sono sufficienti. Capita che molti programmi di machine learning non siano efficaci.

Supervised learning

Nel supervised learning all'algoritmo vengono forniti sia esempi di input, che i rispettivi dati di output, da un "educatore" dell'algoritmo. L'algoritmo cerca di scoprire la regola generale che associa gli inputs agli outputs.

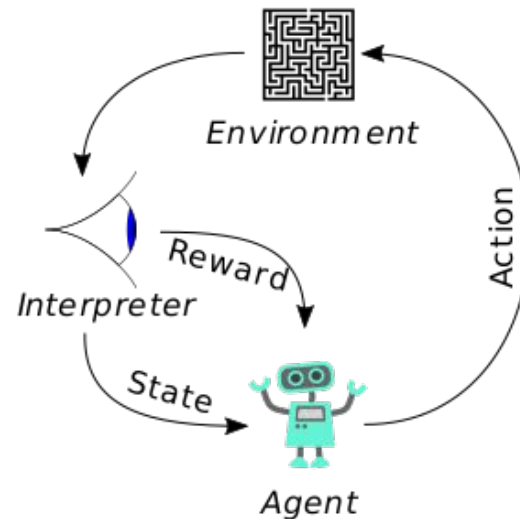


Supervised learning

L'apprendimento può anche continuare in fase di esecuzione dell'algoritmo, sulla base di feedback fornito dall'utente, nel qual caso si parla di active learning.

Reinforcement learning

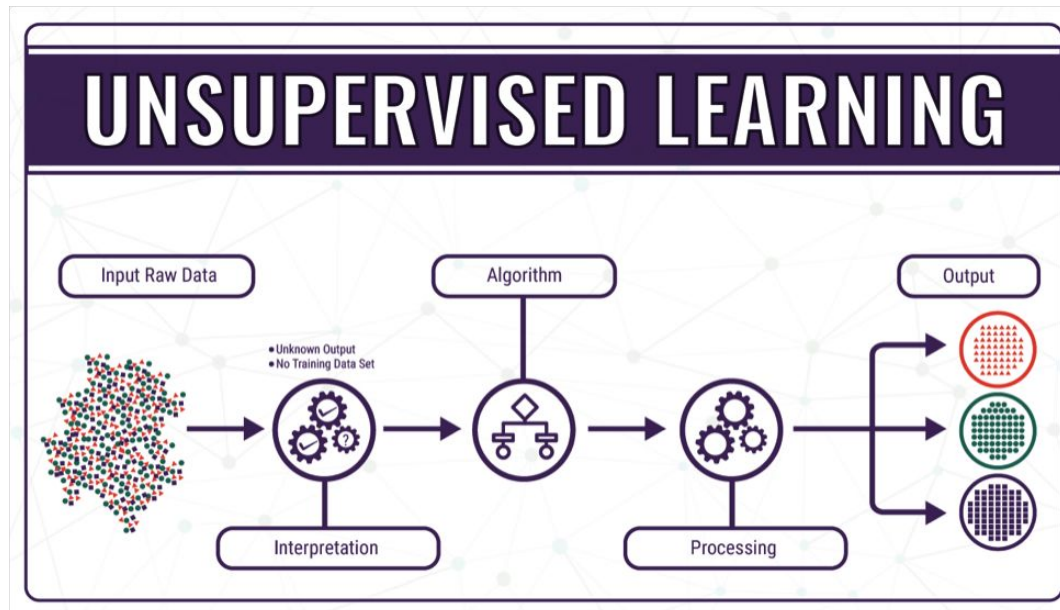
Nelle forme più estreme, come ad esempio nel reinforcement learning, come unica forma di feedback l'output viene approvato o respinto. Il reinforcement learning viene ad esempio sfruttato nei sistemi di guida automatica.



Unsupervised learning

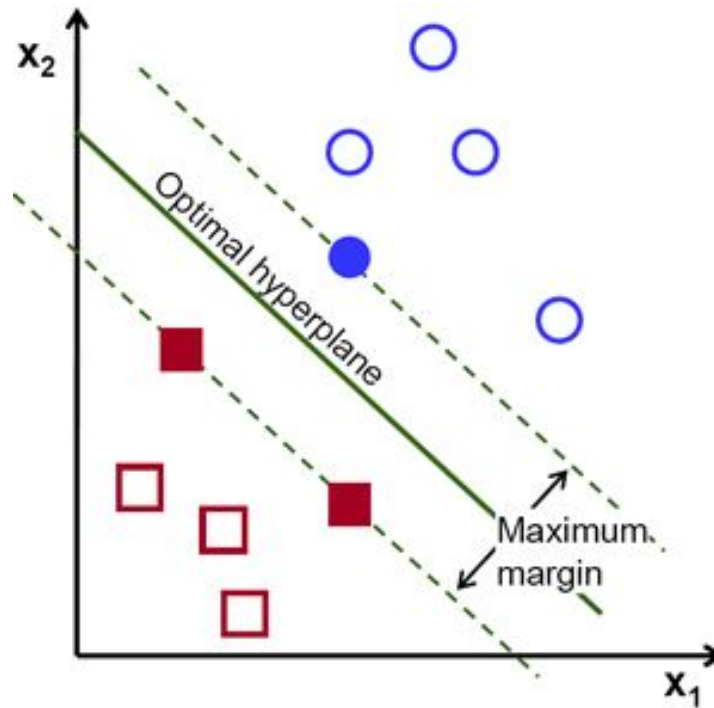
Nel unsupervised learning, nessuna informazione viene fornita all'algoritmo, che è costretto a scoprire in maniera autonoma la struttura dei dati in input.

È la forma più complessa di machine learning.



Esempio: support vector machine

Una support vector machine è un classificatore che divide lo spazio di input in due regioni separate da un margine lineare.



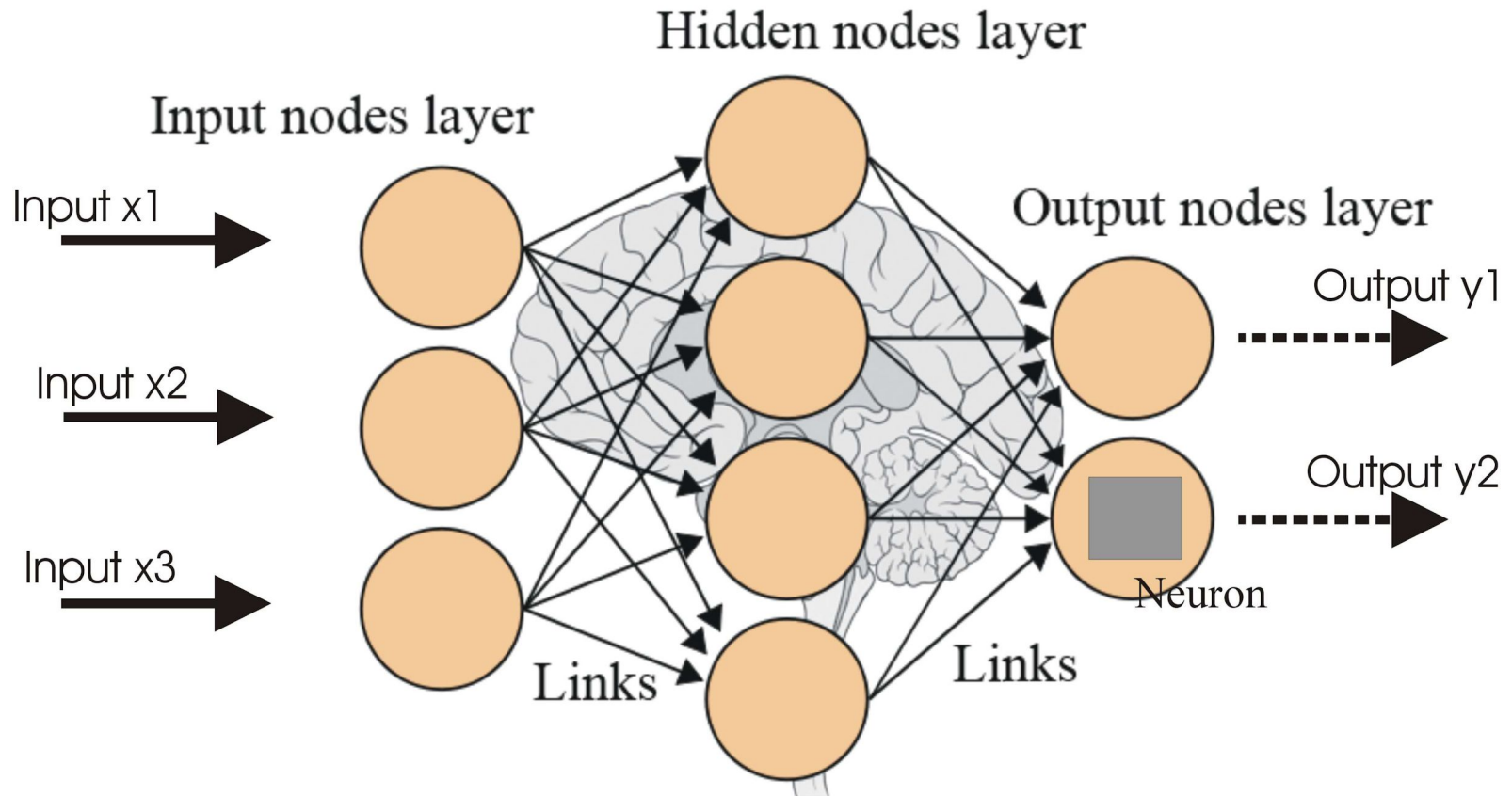
Reti neurali artificiali

Una delle famiglie più utilizzate di algoritmi di machine learning è quella delle reti neurali.

Una rete neurale è un algoritmo che prende ispirazione dalla struttura e dall'aspetto funzionale di reti neurali biologiche. I calcoli vengono strutturati in termini di gruppi di neuroni artificiali interconnessi, che processano collettivamente l'informazione.

Le neural networks utilizzate oggi sono strumenti non lineari per la modellazione di tipo statistico di dati con una certa complessità.

Reti neurali artificiali



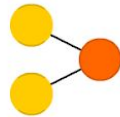
A mostly complete chart of

Neural Networks

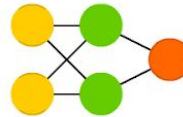
©2016 Fjodor van Veen - asimovinstitute.org

-  Backfed Input Cell
-  Input Cell
-  Noisy Input Cell
-  Hidden Cell
-  Probabilistic Hidden Cell
-  Spiking Hidden Cell
-  Output Cell
-  Match Input Output Cell
-  Recurrent Cell
-  Memory Cell
-  Different Memory Cell
-  Kernel
-  Convolution or Pool

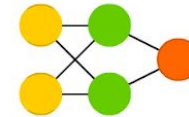
Perceptron (P)



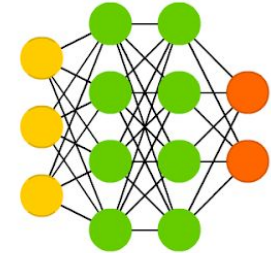
Feed Forward (FF)



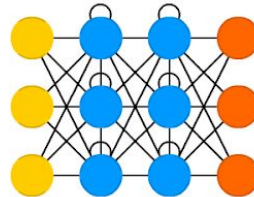
Radial Basis Network (RBF)



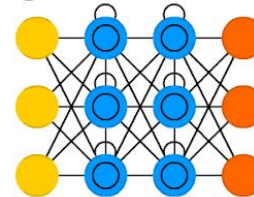
Deep Feed Forward (DFF)



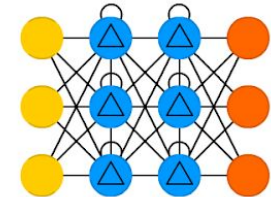
Recurrent Neural Network (RNN)



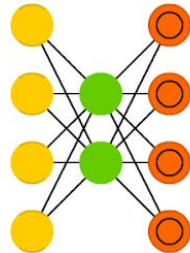
Long / Short Term Memory (LSTM)



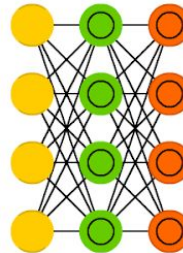
Gated Recurrent Unit (GRU)



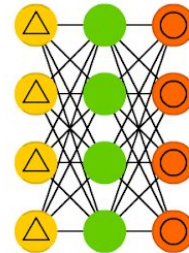
Auto Encoder (AE)



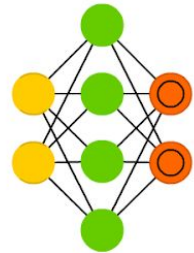
Variational AE (VAE)

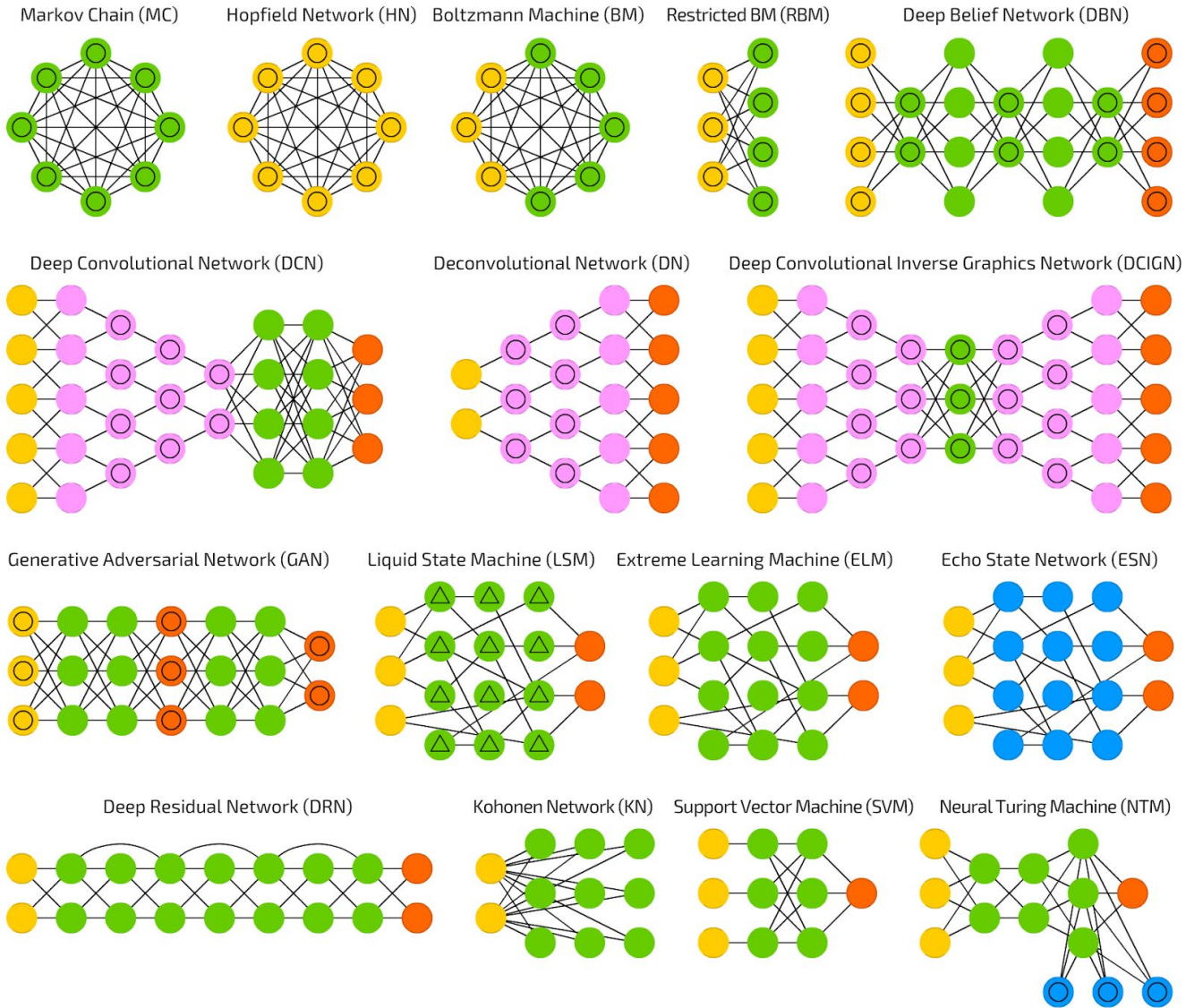


Denoising AE (DAE)



Sparse AE (SAE)





Reti neurali artificiali - vantaggi

Il vantaggio principale delle reti neurali artificiali è la loro capacità di essere generalmente più efficaci degli altri algoritmi di machine learning.

Sono anche generalmente più semplici da allenare rispetto ad altri algoritmi di machine learning ed offrono più flessibilità di parametrizzazione.

Reti neurali artificiali - svantaggi

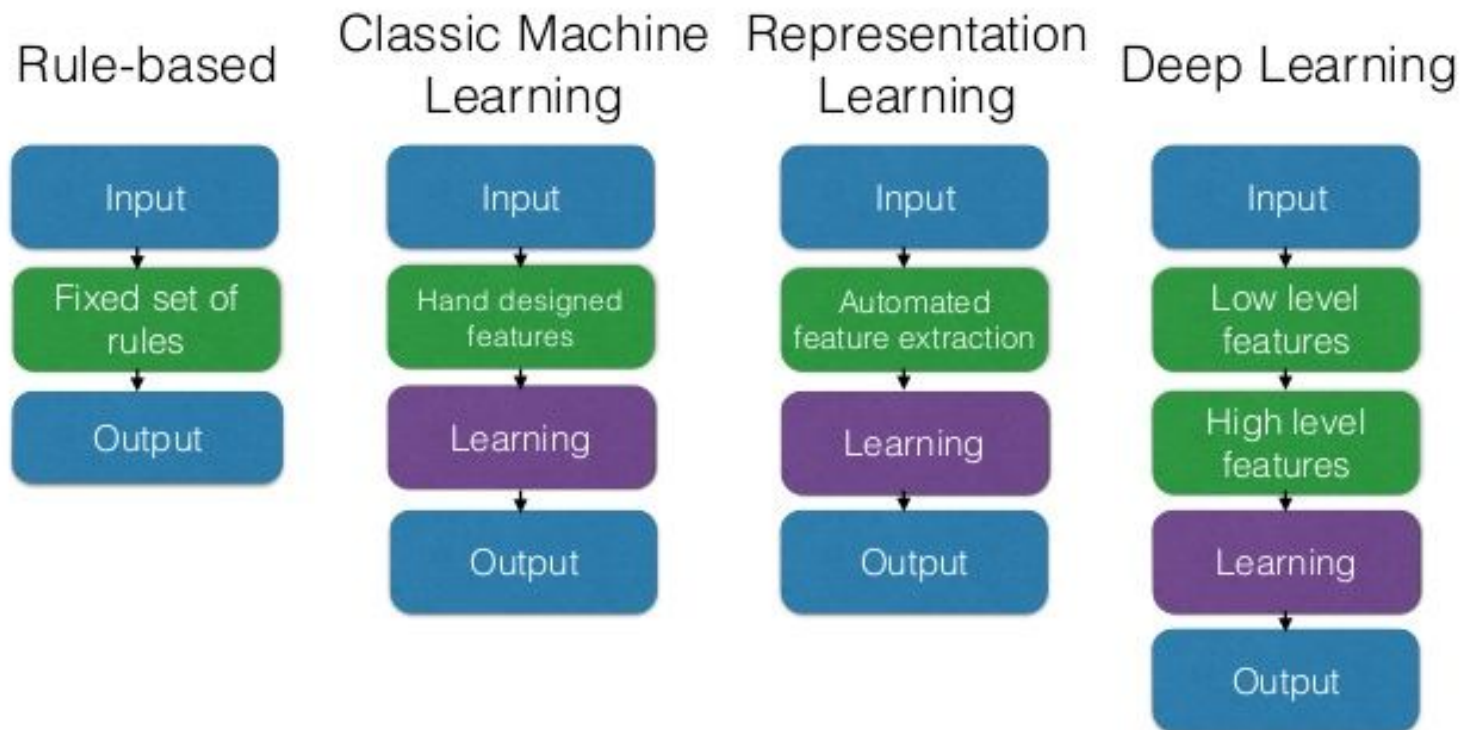
D'altro canto, la loro natura di tipo black-box può risultare svantaggiosa in casi in cui è necessario il controllo del comportamento dell'algoritmo (predictability).

Allo stesso tempo le reti neurali artificiali sono facilmente soggette ad overfitting.

Con pochi dati non identificano la caratteristica generica desiderata ma il caso specifico.

Per renderle capaci di generalizzare è necessario allenarle con data sets molto ampi e ben costruiti.

Evoluzione delle tecniche di ML



Deep learning

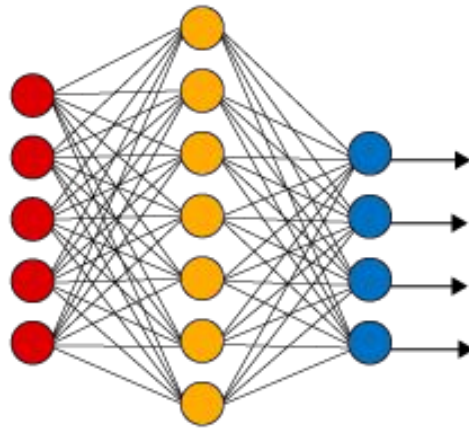
La crescita delle capacità di calcolo, la progressiva riduzione del costo dell'hardware e l'evoluzione delle GPUs degli ultimi anni, hanno contribuito allo sviluppo del concetto di deep learning.

Nel deep learning si cerca di emulare il processing che il cervello esegue ad esempio sulla luce e sul suono nelle capacità di visione e di udito.

Deep learning

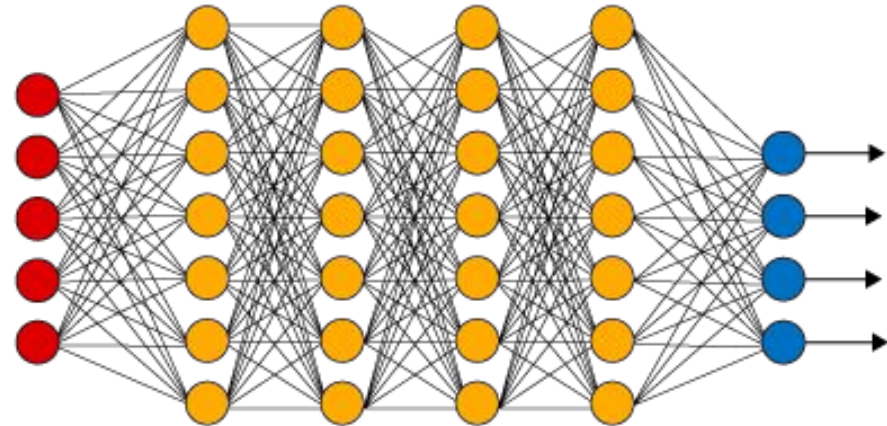
Nel deep learning vengono utilizzate reti neurali composte da layers multipli.

Simple Neural Network



● Input Layer

Deep Learning Neural Network



● Hidden Layer

● Output Layer

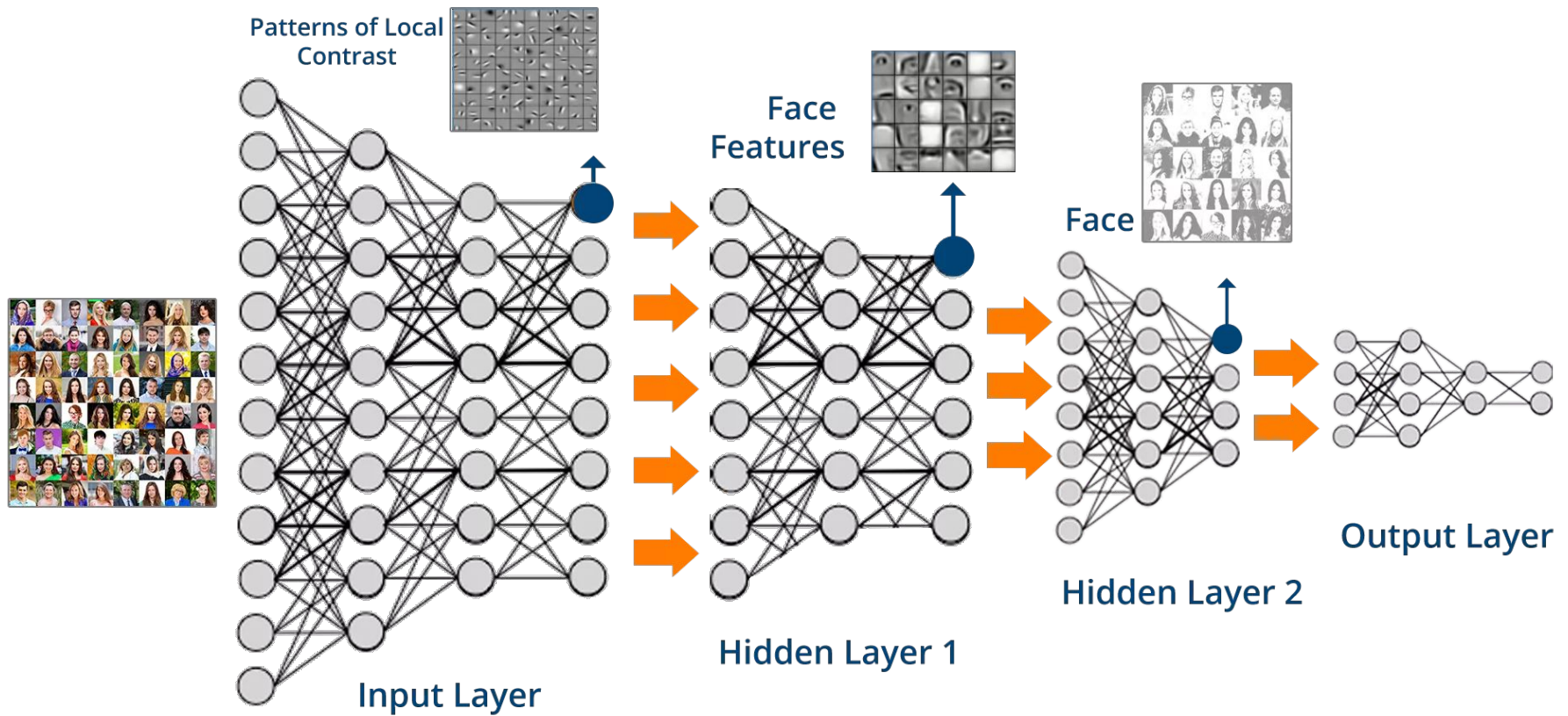
Deep learning

Nel deep learning, ogni layer impara a trasformare i dati di input in una rappresentazione ogni volta leggermente più astratta e composta.

Il termine deep si riferisce al numero di layers attraverso i quali i dati vengono trasformati.

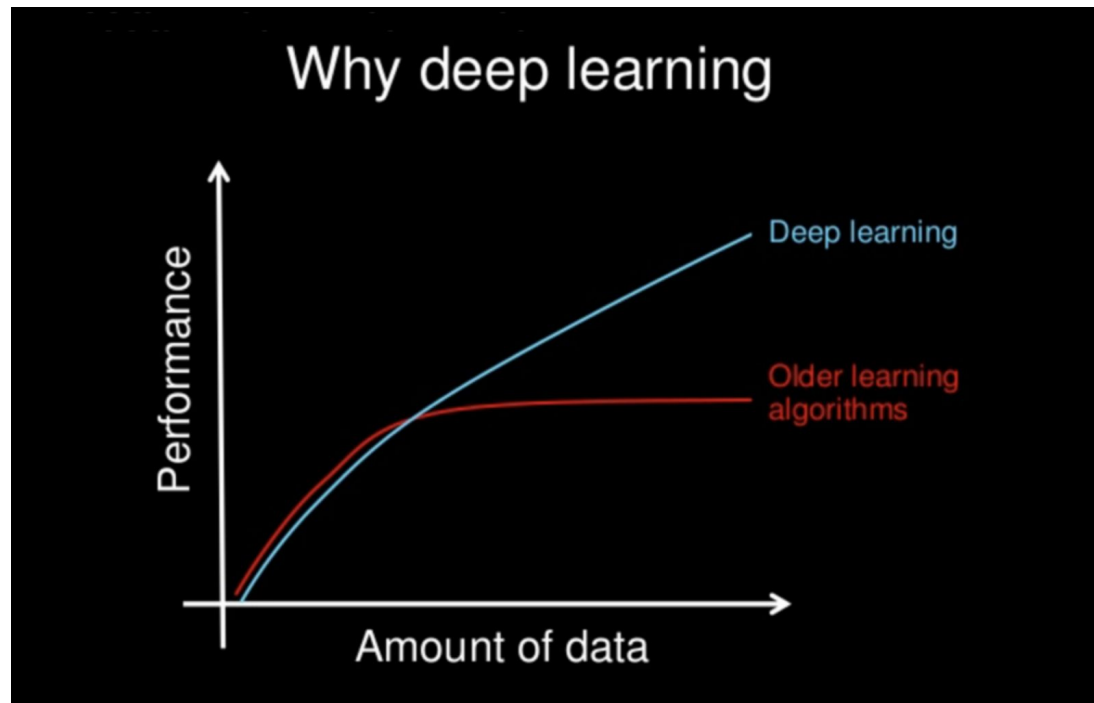
Nelle soluzioni end-to-end l'intero processo , dai sensori agli attuatori, viene eseguito da un'unica layered o recurrent neural network senza modularizzazione.

Deep learning



Deep learning

Il vantaggio degli algoritmi di deep learning è la capacità di continuare a migliorare le performances all'aumentare dei dati di apprendimento disponibili.

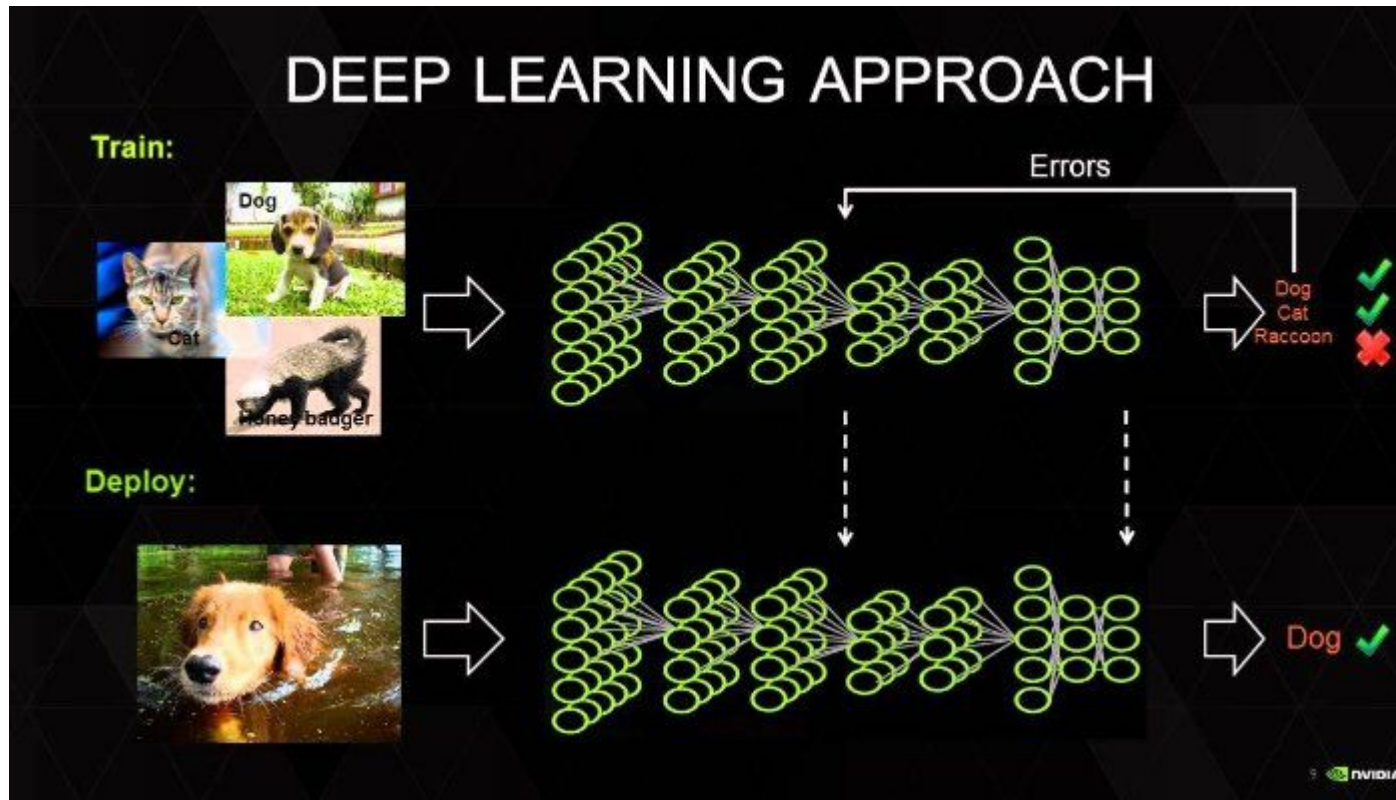


Deep learning

Per quanto concerne l'ambito del multimedia, oggi esistono applicazioni di successo soprattutto nella computer vision e nel riconoscimento vocale.

Deep learning

Ad esempio, il deep learning viene utilizzato con successo nella classificazione di immagini.

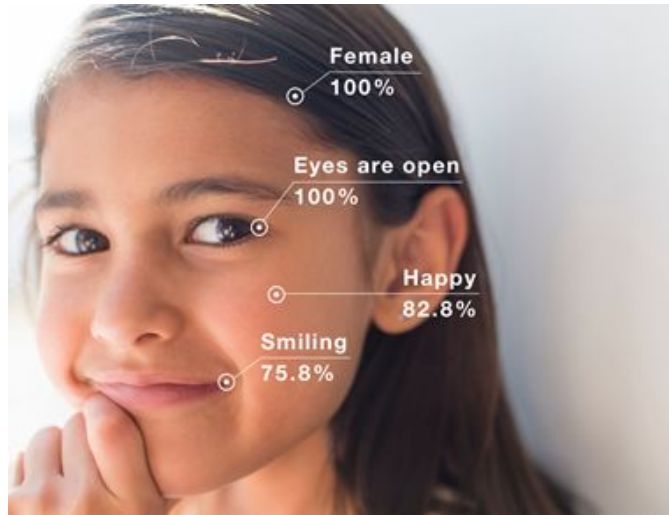


Esempio: Amazon Rekognition

Amazon Rekognition è un servizio della piattaforma AWS che permette di aggiungere funzionalità di analisi d'immagine alle applicazioni. Con Rekognition è possibile identificare oggetti, scene e facce. È anche possibile confrontare le facce.

L'API di Rekognition's permette l'accesso ad algoritmi di deep learning per la classificazione d'immagine. Utilizza modelli di tipo deep neural network.

Esempio: Amazon Rekognition



Esempio: Amazon Transcribe

Amazon Transcribe è un servizio della piattaforma AWS per il riconoscimento vocale automatico.

Facilita l'aggiunta di funzionalità di speech-to-text alle applicazioni.

Con Amazon Transcribe è possibile analizzare files audio, ottenendo come risultato files di testo contenenti la trascrizione.

È possibile utilizzare Transcribe anche in modalità streaming.

Esempio: Caffe

Caffe è un framework di deep learning sviluppato con l'obiettivo di essere veloce, modulare e facile da utilizzare. è realizzato dal Berkeley AI Research (BAIR) e da una comunità di contributors. Viene rilasciato sotto licenza di tipo BSD 2-Clause.