



Algoritmi per il Machine Learning

Dott. Antonio Giovanni Lezzi





ORARIO SVOLGIMENTO DEL MODULO

- Svolto in due moduli differenti, suddivise in 80 ore ciascuna
- Si spiegheranno algoritmi e loro implementazione
- Lezioni iniziano il 12 aprile e finiscono 16 giugno 2021 (a meno di imprevisti)
- L'esame sarà sicuramente teorico e progetto da svolgere (16 giugno)
- Eventuali variazioni verranno comunicati tempestivamente



COMPETENZE RICHIESTE

- Linguaggio di programmazione: Python ma è possibile usare qualsiasi altro linguaggio
- Conoscenza di modelli matematici e probabilità e statistica
- Problem solving



ALGORITMI DI MACHINE LEARNING

- Introduzione alla Data Science
 - o Discipline che ne fanno parte
 - o Frequenti cause di errori metodologici nella statistica classica.
- Introduzione al Machine Learning
 - o Esempi di problemi
 - o Classificazione e formalizzazione dei problemi di machine learning: apprendimento supervisionato, classificazione e regressione, apprendimento non supervisionato, clustering, reinforcement.
 - o Algoritmo K-Nearest Neighbors (KNN)



ALGORITMI DI MACHINE LEARNING

- Regressione lineare (supervised - regressione)
 - o Valutazione di un modello di regressione: il root mean square error (RMSE)
 - o Minimizzazione del RMSE per un modello lineare: il metodo dei minimi quadrati a una e ad n dimensioni.
 - o Generalizzazione della regressione lineare per funzioni di base arbitrarie.
 - o Utilizzo delle potenze come funzioni di base: la regressione polinomiale (a una dimensione).



ALGORITMI DI MACHINE LEARNING

- Validazione di un modello
 - o Valutazione di un modello di classificazione: l'accuratezza.
 - o Matrice di confusione per due o più classi.
 - o Definizione di "caso positivo"; gli indici di precisione e sensibilità; la F1 score.
 - o La cross-validation, K-fold cross-validation, leave-one-out: cross-validation stratificata
- Regressione logistica (supervised- classificazione)
 - o Utilizzo di un modello lineare con una funzione di soglia per problemi di classificazione.
 - o La funzione sigmoide.
 - o Ottimizzazione iterativa: la discesa lungo il gradiente



OBIETTIVI

- L'Unità Formativa costituisce un'introduzione teorico-pratica alle tecniche di apprendimento automatico (machine learning) che, a partire da esempi, generano modelli matematici che possono essere generalizzati a nuovi casi per operare previsioni. Esistono decine, se non centinaia, di pacchetti software e di librerie in grado di importare dati, costruire modelli, fare previsioni. Uno sviluppatore può creare una rete neurale, addestrarla e utilizzarla in un suo programma senza nemmeno sapere come funziona, semplicemente utilizzando le API di una libreria e trattando i diversi algoritmi come "scatole nere". Insomma, non è necessario conoscere un algoritmo di machine learning per poterlo utilizzare, come non è necessario conoscere i protocolli di rete in dettaglio per realizzare un'applicazione web, e si possono ordinare gli elementi di un vettore senza conoscere gli algoritmi di ordinamento. Nonostante ciò, le tecniche che studieremo in questo corso vanno applicate con estrema attenzione, perché sono soggette a errori estremamente perniciosi, ma non sempre facili da evitare. Il corso sarà in parte dedicato a studiare le metodologie utili a evitare o a minimizzare l'impatto dei trabocchetti più comuni come il paradosso di Simpson.



ESAME

- **16 Giugno 2021**
- **Modalità: domande a risposta chiusa, aperta e analisi su uno scenario d'uso e progetto da realizzare**

MACHINE LEARNING

Dott. Antonio Giovanni Lezzi





- Introduzione al Machine Learning
- Progettazione di un algoritmo
- Algoritmi di Machine Learning



ALGORITMI DI MACHINE LEARNING

- **Figure professionali**
- Cenni storici
- Algoritmi e complessità computazionale
- Ambiti e applicazioni



FIGURE PROFESSIONALI

- Realizzare un progetto richiede una quantità enorme di conoscenze, difficilmente un solo esperto può gestirle tutte.
- Il **project manager** gestisce il progetto, coordina le persone, individua priorità, verifica l'andamento dei lavori, ecc.
- **L'esperto in modelli matematici** sceglie le formule matematiche, eventualmente crea nuove formule, ottimizza i parametri, senza pensare ai DB e ai linguaggi di programmazione.



FIGURE PROFESSIONALI

- **Lo sviluppatore** prende il modello matematico e realizza il software necessario per realizzarlo, ottimizzando i tempi di calcolo e la memoria.
- **L'esperto di database** si dedica alle strutture per conservare i dati, in modo da fornire ciò che serve in poco tempo e con facili accessi.
- **Altre figure** si incaricano della rete di computer e della **sicurezza informatica**.



ALGORITMI DI MACHINE LEARNING

- Figure professionali
- **Cenni storici**
- Algoritmi e complessità computazionale
- Ambiti e applicazioni





CENNI STORICI

- **Aristotele (384-322 a.C.)** codificò certi stili di ragionamento deduttivo chiamandoli sillogismi, esempio: “Tutte le torte sono dolci, tutti i dolci sono alimenti, ne consegue che tutte le torte sono alimenti”. ($A \rightarrow B, B \rightarrow C \Rightarrow A \rightarrow C$)
- **Leibniz (1646-1716)** intendeva realizzare un'algebra universale con cui tutta la conoscenza, comprese le verità morali e metafisiche, potesse un giorno essere ricondotta entro un singolo sistema deduttivo.
- **Boole (1815-1864)** sviluppò i fondamenti della logica proposizionale, con lo scopo di raccogliere dei probabili indizi riguardanti la natura e la costituzione della mente umana.
- **Frege (1848-1925)** propose un sistema simbolico per il ragionamento meccanico: il calcolo dei predicati.
- I logici del XX secolo, come **Gödel, Kleene, Church e Turing**, hanno specificato in modo chiaro e formale le potenzialità e le limitazioni dei sistemi logici e computazionali.



CENNI STORICI

- Il nome più famoso è quello di **Alan Turing** (1912-1954), per il ruolo svolto nella Seconda guerra mondiale nell'analisi dei codici segreti tedeschi e per le vicissitudini della sua vita.
- Il nome di Turing è legato anche a una macchina ideale che manipola i dati contenuti su un nastro di lunghezza potenzialmente infinita, secondo un insieme prefissato di regole ben definite.
- Il suo documento fondamentale sull'AI si chiama Computing Machinery and Intelligence (1950), in cui viene discussso il test di Turing come criterio per determinare se una macchina sia in grado di pensare.



Dott. Antonio Giovanni Lezzi



CENNI STORICI

- **Samuel (1901-1990)**, pioniere nel campo dei videogiochi per computer, creò il primo programma sul gioco della dama basato sull'apprendimento automatico.
- Gli informatici **Stephen Cook e Richard Karp**, hanno identificato classi di computazioni che potrebbero richiedere enormi quantità di tempo e di memoria e a computazioni che non si possono portare a termine.
- **McCulloch (1898-1969) e Pitts (1923-1969)** hanno studiato nel 1943 i rapporti tra semplici elementi di calcolo e i neuroni biologici.
- **Rosenblatt (1928-1971)** nel 1962 ha studiato l'uso di reti, chiamate "percetroni" (perceptron) e costituite da elementi simili a neuroni, per l'apprendimento e per il riconoscimento di forme. Inizialmente non ebbe il successo sperato con un blocco temporaneo a partire dal 1969.



CENNI STORICI

- Tra gli **anni '70 e '80** vengono studiate varie tipologie di rappresentazioni di problemi, di tecniche di ricerca di soluzioni e di euristica, impiegate in programmi in grado di risolvere semplici puzzle, fare giochi e acquisire informazioni.
- Vengono sviluppati programmi sempre più capaci, contenenti la conoscenza richiesta per raggiungere le prestazioni di esperti umani nello svolgimento di svariati compiti, quali la diagnosi, la progettazione e l'analisi.



CENNI STORICI

- **Facebook** avvia attività di ricerca e sviluppo per migliorare l'interazione sociale dei suoi iscritti.
- **NVIDIA**, produttrice di processori grafici, schede madri e componenti per PC e console, a partire dal 2015 si lancia nell'ambito del deep learning con Tegra X1 per arrivare alla GPU: VOLTA.
- **Nell'AI segue un andamento ondulatorio: nasce un nuovo approccio, si ottengono risultati interessanti, si trovano punti deboli e risultati scoraggianti, non si riesce a scalare verso problemi fattibili, si perde fiducia con drastica riduzione dei fondi, per molti anni si ferma tutto.**



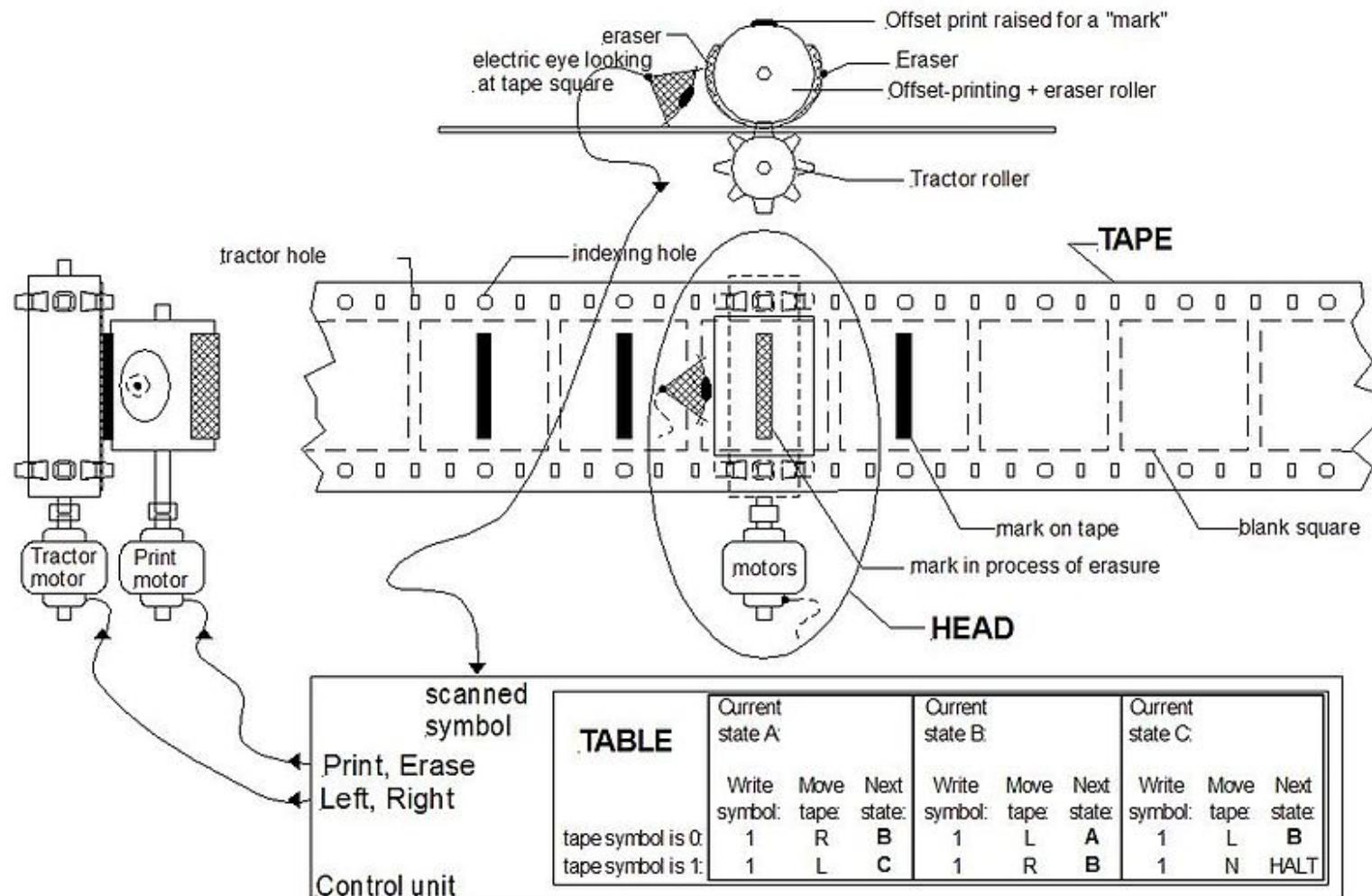
ALGORITMI DI MACHINE LEARNING

- Figure professionali
- Cenni storici
- **Algoritmi e complessità computazionale**
- Ambiti e applicazioni



MACCHINA DITURING

- Si tratta di un **modello astratto** che definisce una macchina in grado di **eseguire algoritmi** e dotata di un **nastro** potenzialmente **infinito su cui può leggere e/o scrivere** dei simboli.
- È un potente strumento teorico usato nella teoria della calcolabilità e nello studio della complessità degli algoritmi, in quanto è di notevole aiuto agli studiosi nel comprendere i limiti del calcolo meccanico.
- Nel 1936 egli aveva già introdotto il concetto di calcolabilità.

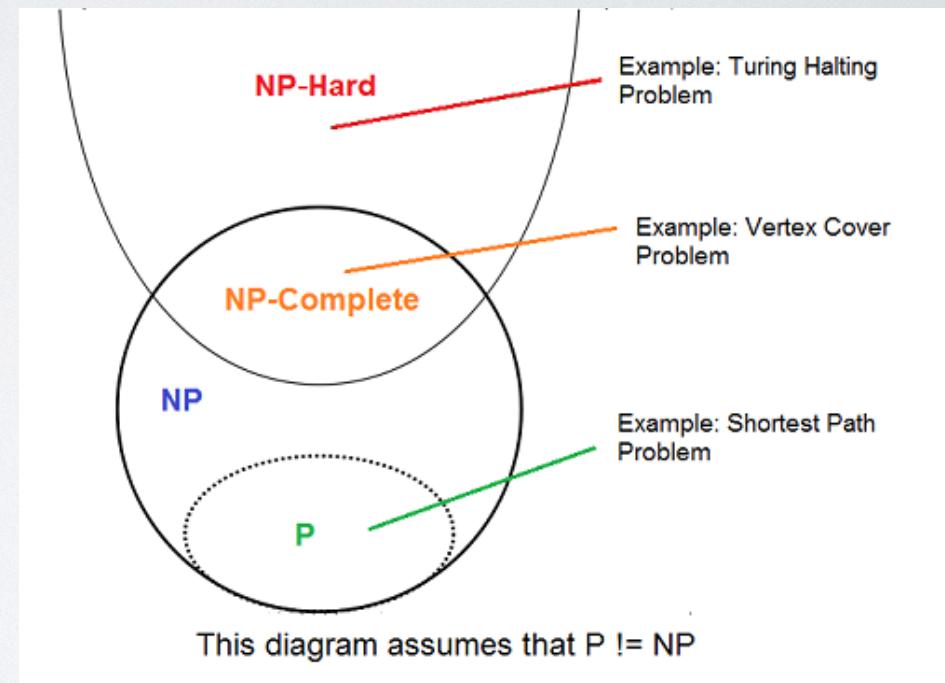


A fanciful mechanical Turing machine's TAPE and HEAD. The TABLE instructions might be on another "read only" tape, or perhaps on punch-cards. Usually a "finite state machine" is the model for the TABLE.



COMPLESSITÀ COMPUTAZIONALE

- **Un algoritmo** è una strategia che serve per risolvere un problema ed è costituito da una sequenza finita di operazioni (istruzioni) in tempo finito.
- La **complessità computazionale** si occupa della valutazione del costo degli algoritmi in termini di risorse di calcolo: tempo di elaborazione equantità di memoria (spazio) utilizzata.
- L'obiettivo è quello di comprendere le prestazioni massime raggiungibili da un algoritmo applicato ad un determinato problema con minime risorse.



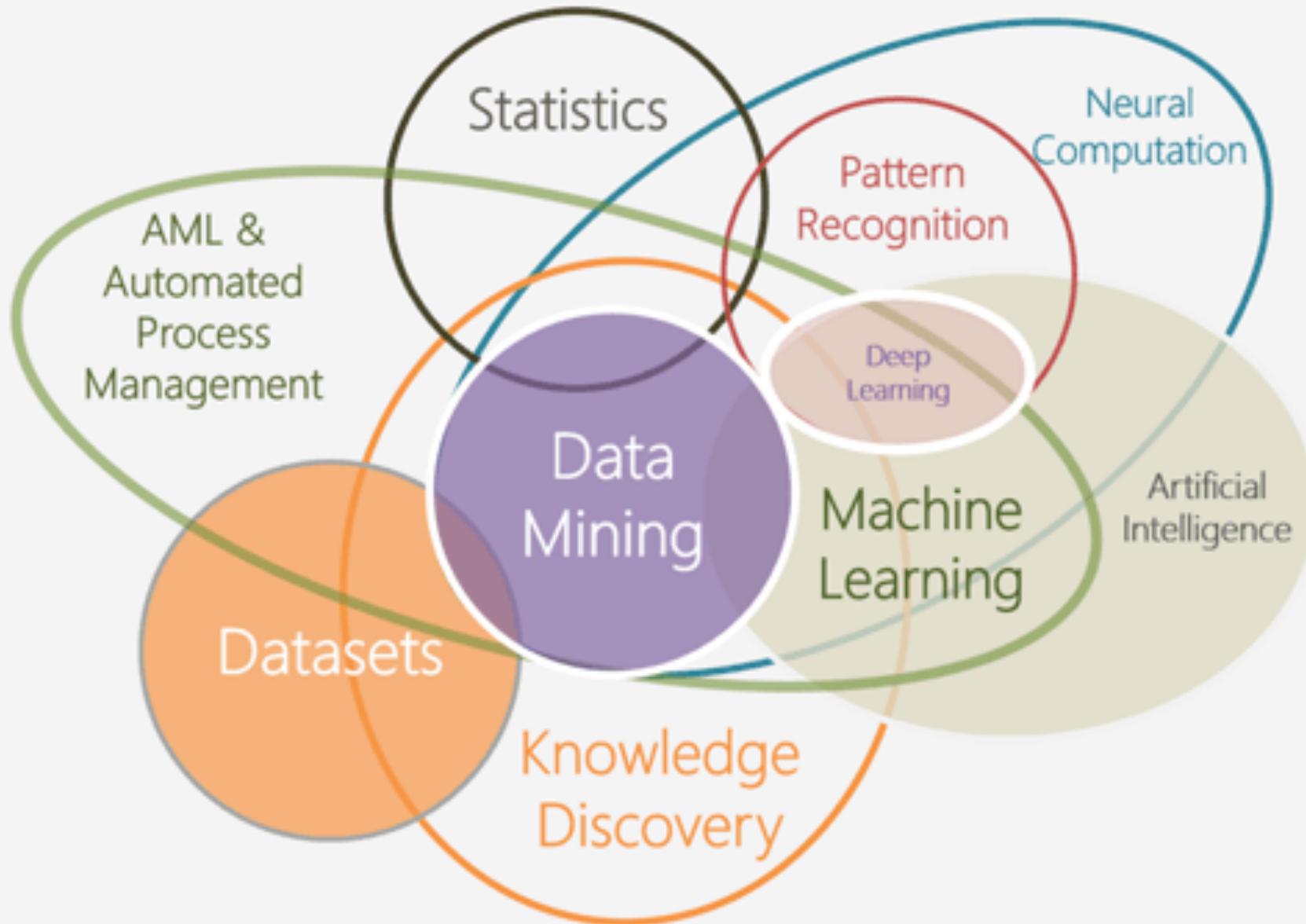


ALGORITMO EURISTICO

- **Algoritmo euristico** è progettato per risolvere un problema più velocemente, nel caso in cui i metodi classici siano troppo lenti nel calcolo o per trovare una soluzione approssimata, nel caso in cui i metodi classici falliscano nel trovare una soluzione esatta.
- Il risultato viene ottenuto cercando di **equilibrare gli obiettivi di maggiori ottimizzazione, completezza, accuratezza e velocità di esecuzione**.
- I metodi euristici costituiscono **spesso una strada obbligata per risolvere problemi molto difficili** (ad esempio quelli di tipo NP-difficile) come il problema del commesso viaggiatore, in quanto per determinate dimensioni delle istanze l'algoritmo euristico riesce a ricavare una soluzione approssimativamente vicina a quella ottima in tempi ragionevoli.

ALGORITMI DI MACHINE LEARNING

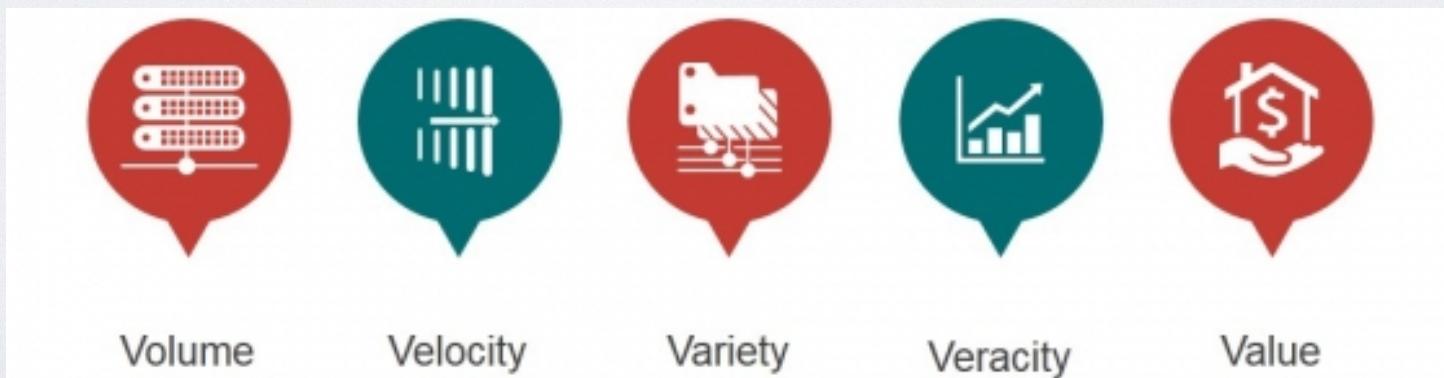
- Figure professionali
- Cenni storici
- Algoritmi e complessità computazionale
- **Ambiti e applicazioni**





BIG DATA

- I big data vengono utilizzati perché per svolgere le varie forme di apprendimento in molti casi sono necessarie grandi quantità di dati, con le caratteristiche denominate come 5V:
- **Volume** in memoria di memorizzazione (oltre ZettaByte), da elaborare in **Velocità** con ampia **Varietà** di strutture e formati dei file; con **Veridicità** delle fonti; creando **Valore**;





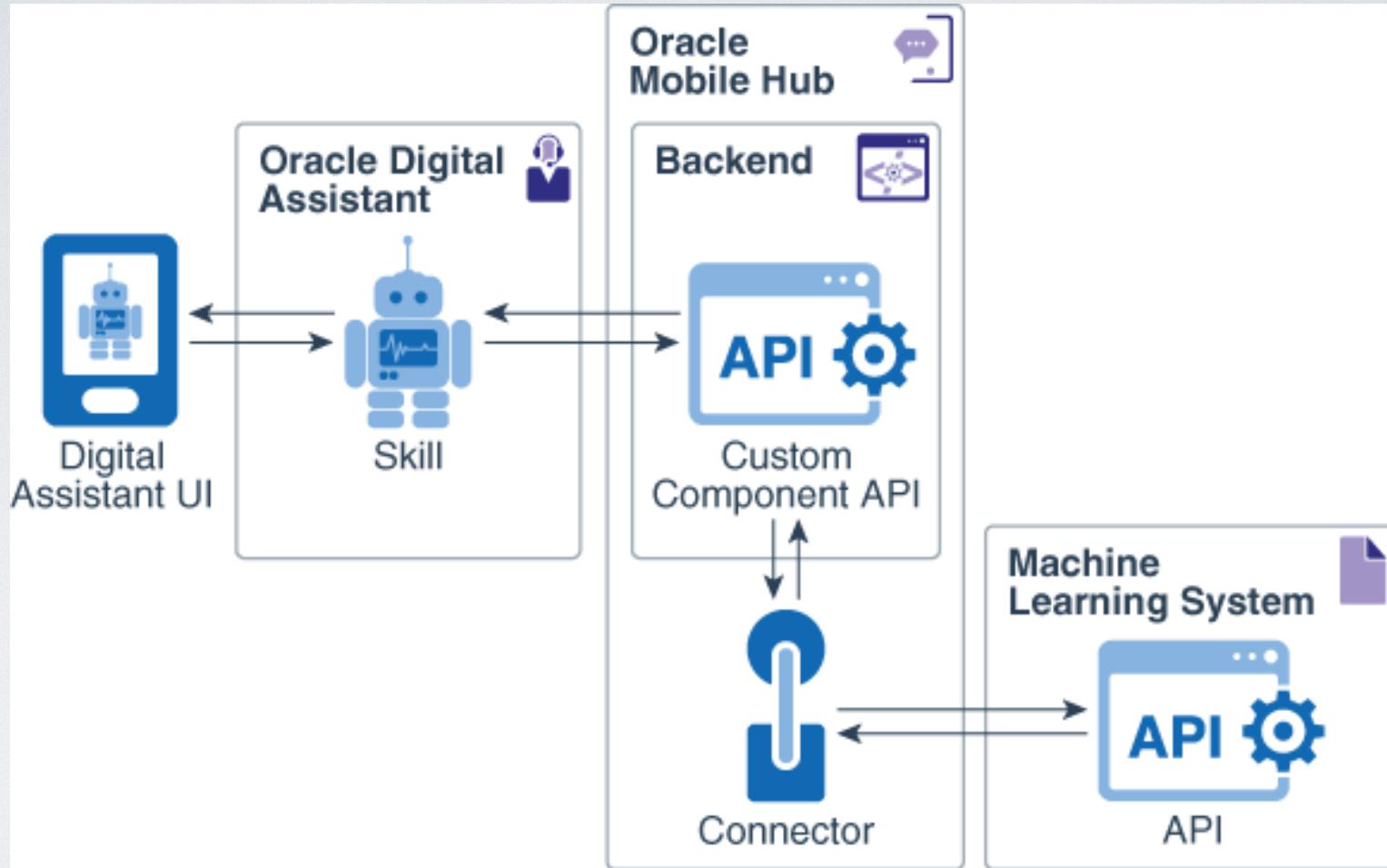
CIBERNETICA

- È la **teoria della retroazione e del controllo**, deriva dal greco kybernetes, “**pilota di nave**”, cioè il timoniere che usa gli strumenti e corregge per dirigere la nave verso l'obiettivo.
- È una disciplina non astratta ma matematica, con cui assegnare alle grandezze altre grandezze valutabili e descrivibili e dare un'interpretazione nuova nell'osservare l'universo e il modo in cui si comportano i suoi componenti.
- Gli esseri umani e le macchine cominciano ad assomigliarsi, il processo di regolazione è un trasferimento di informazione. Cosa li fa funzionare? Cosa li mantiene in moto? Prima la risposta era: l'energia, dopo è: l'informazione.
- Nell'AI vuole simulare l'attività logica del cervello umano riproducendo artificialmente, cioè costruite componenti elettroniche la struttura cerebrale delle reti di neuroni.



CLOUD COMPUTING

- Gli sviluppatori di AI possono usare algoritmi molto sofisticati, hanno ampia disponibilità di memoria e velocità di calcolo.
- Gli svantaggi derivano dal fatto di dover essere sempre collegati online con il cloud, di dover stare attenti ai pagamenti per non perdere la licenza, e dal fatto che i propri dati vengono elaborati dalle aziende e potrebbero esserci problemi di riservatezza su dati delicati che non dovrebbero uscire dal contesto di business.





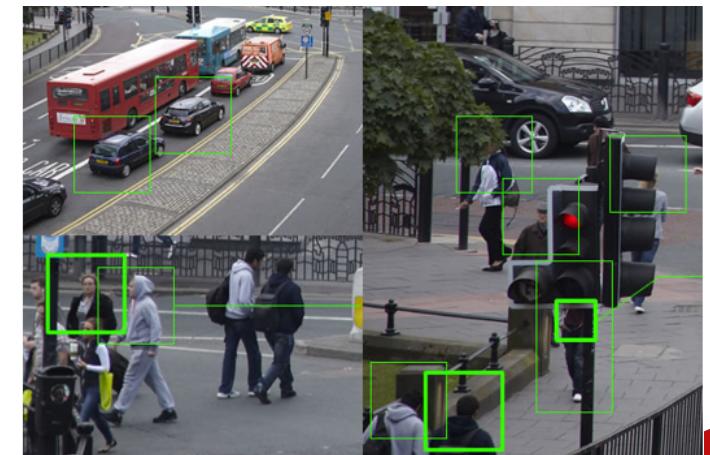
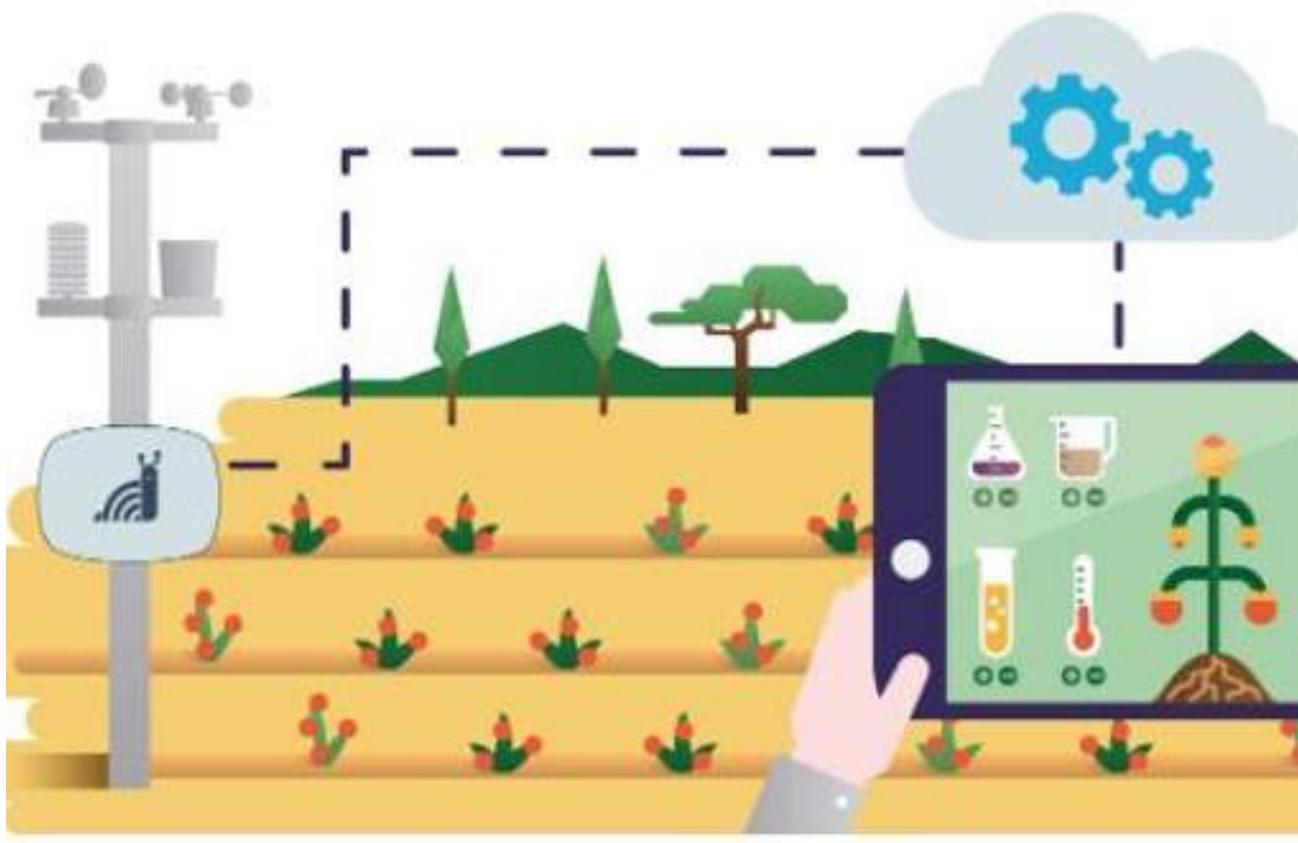
DATA MINING

- **Mining** significa scavare, andare oltre quello che si conosce e la definizione più significativa è sull'analogia con le operazioni dei minatori: **scavano grandi quantità di materiale di poco valore per trovare l'oro.**
- Nel data mining l'oro è **l'informazione**, infatti è una tecnologia da usare per ottenere alcune variabili di sintesi utili come strumento a supporto delle decisioni.
- Se si conoscono in anticipo la forma, il contenuto e la quantità di ciò che si sta cercando, allora si sta facendo una analisi dei dati. Più un'informazione è poco ovvia, strutturata e non se ne suppone l'esistenza, più il suo valore aumenta.
- Anche l'**AI permette di estrarre informazioni con una natura esplorativa**, quindi c'è una certa affinità di obiettivo, tanto che per fare data mining si usano molte tecniche derivanti dagli apprendimenti supervisionato e non supervisionato.



IOT

- Il concetto di internet delle cose (Internet of Things, IoT) per fare riferimento alla rete globale formata dalle **interconnessioni tra dispositivi fisici in grado di trasmettere dati**.
- Data l'enorme quantità di dati generati, tipica dei big data, **gli algoritmi di AI possono esprimere la loro potenza per estrarre informazioni utili, eliminare rumore, ricostruire dati mancanti, prendere decisioni in pochissimo tempo e tanto altro.**
- Un aspetto molto importante deriva dalla condivisione di questi dati tra le macchine, in modo che acquistino autonomia, così la conoscenza entra a far parte del sistema percettivo di internet.
- La convergenza tra IoT, big data, AI è conosciuta come internet cognitivo e ubiquo.





ROBOTICA

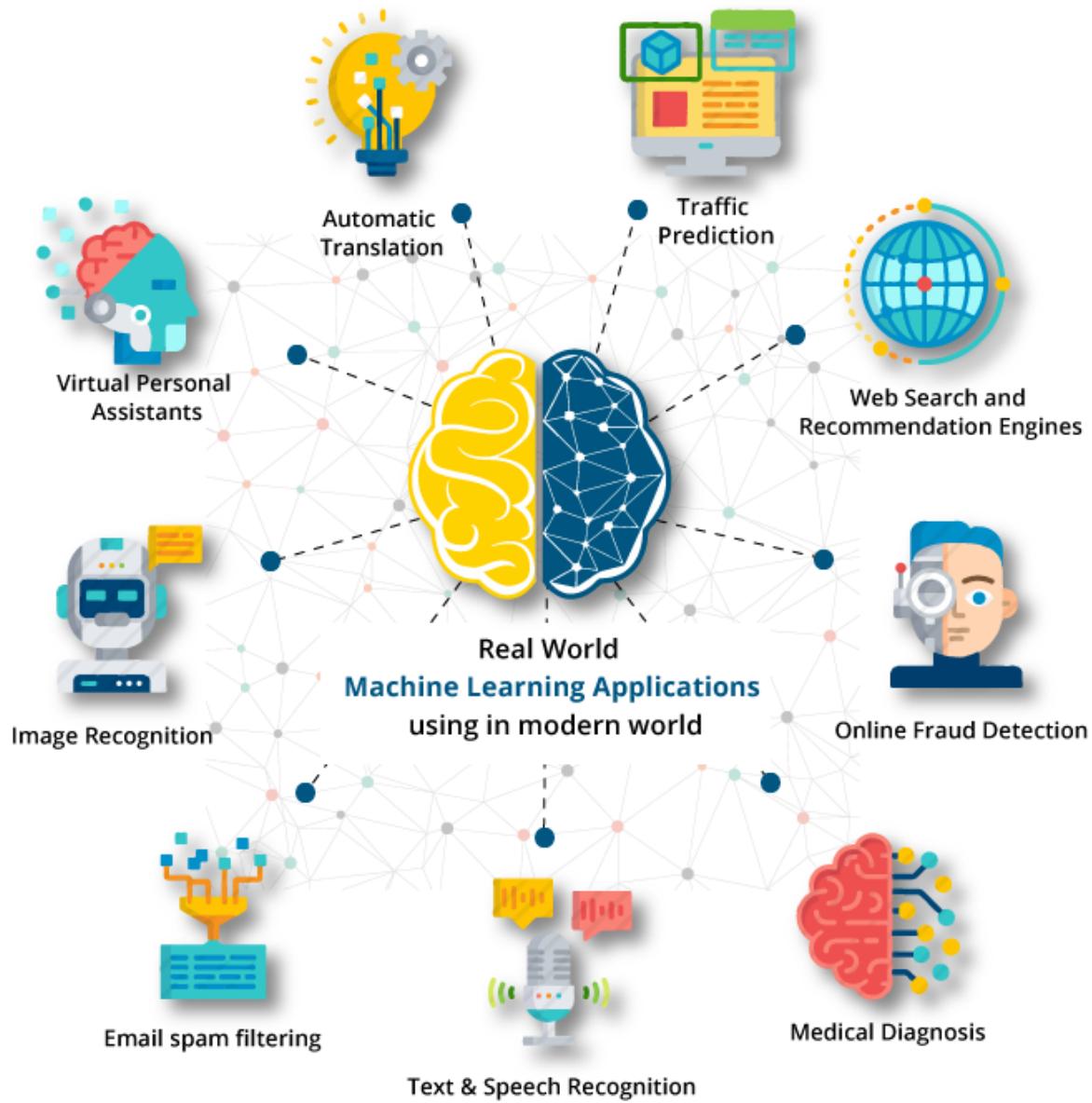
- **La robotica è l'unione di varie scienze**, discipline e tecnologie per studiare e sviluppare metodi che permettano a un robot di **eseguire dei compiti specifici riproducendo il lavoro umano.**
- Il primo progetto documentato di un robot umanoide viene fatto da Leonardo da Vinci, realizzò disegni per realizzare un cavaliere meccanico, in grado di alzarsi in piedi ed agitare le braccia e muovere testa; progetto mai realizzato.
- Il primo robot funzionante fu costruito nel 1738 da Jacques de Vaucanson, che realizzò una serie di macchine che svolgevano azioni animali. Il primo robot industriale venne installato nella fabbrica della General Motors di Trenton (USA) nel 1961 per fare pressofusione.



ROBOTICA

- **L'AI è collegata alla robotica** perché consente di **realizzare** forme di **ragionamento, percezione, pianificazione essenziali per l'interazione** con l'ambiente circostante.
- In particolare, rende possibile al robot la reazione alle emozioni manifestate dall'uomo.
- Inoltre, molte ricerche e applicazioni della robotica hanno permesso di creare nuovi metodi di calcolo.





Dott. Antonio Giovanni Lezzi



APPLICAZIONI

- Le applicazioni in ambito business riguardanti **servizi finanziari, assicurazioni, blockchain**,
- Chi è interessato all'**ambito sociale e umanitario**, soprattutto parlando di paesi in via di sviluppo
- **Intelligent Data Processing**: analizzano dati specifici per estrapolare informazioni e compiere azioni in conseguenza;
- **Virtual Assistant/Chatbot**: in grado di eseguire azioni o erogare servizi per un individuo in base a comandi ricevuti in maniera vocale o testuale;
- **Recommendation**: indirizzano le scelte degli utenti in base a informazioni da essi fornite;



APPLICAZIONI

- **Image Processing:** analizzano immagini statiche o video per il riconoscimento di oggetti, persone o animali;
- **Autonomous Vehicle:** si riferisce ai mezzi di trasporto autoguidati;
- **Intelligent Object:** in grado di compiere azioni senza l'intervento umano e di prendere decisioni in base alle condizioni dell'ambiente circostante;
- **Language Processing:** elaborano il linguaggio con finalità quali la comprensione del testo e la traduzione;
- **Autonomous Robot:** tipologia di robot in grado di muoversi senza l'intervento umano, in base a informazioni raccolte dall'ambiente circostante.