

ORARI LEZIONE:

Teoria: lunedì

Laboratorio: mezzogiorno

Le attività di pratica iniziano la terza settimana. Usare =
mo python con libreria numpy.

Punti bonus: è possibile ottenere punti bonus svolgendo settimanalmente gli esercizi del lab e inviandoli. Si creerà una leaderboard sulla base della performance delle soluzioni. I punti bonus assegnati dipendono dalla posizione in classifica. Se non si è tra i primi 3, si ottengono 1 o 2 punti. Se poi si consegna nulla 0 punti. Si può comunque prendere 30 all'esame anche senza aver fatto i lab.

ESAME: 4 domande aperte + qualche esercizio

MATERIALI: 10 pdf di slide su virtuale, tutte le vecchie prove di esame & relative soluzioni sul sito del prof (è linkato su virtuale), registrazioni delle lezioni 2020-2021 su virtuale.

1_ML_Introduzione

Intelligenza: capacità di pensare, comprendere e elaborare modelli astratti della realtà.

“**artificiale:** riproduzione delle azioni legate all'intelligenza svolta da macchine.

RAGIONAMENTO

DEDUTTIVO

le fondamenta si grazie
fosse di **DIM** e teoremi
nella matematica. A partire
da una regola R, posso
scoprire il risultato R₁ su
un caso C₁.



Ho R e C₁, trovo R₁

INDUTTIVO

le più usato nel M.L.
È un ragionamento
probabilistico (= parto
dall'osservazione di C₁ e R₁
e trovo la regola R. Non ho
la certezza.)

La forma più comune di
questo ragionamento è la
GENERALIZZAZIONE.

ABDUTTIVO

Molto recente. Come
quello INDUTTIVO, è **probabilistico**, ma invece di
generalizzare, si ipotizza
che un'implicazione
valga anche al contrario
e quindi si cerca di
ricavare C₁ a partire
da R e R₁. Si usa

Un'altra è il **RACIONAMENTO**
PER ANALOGIA in cui si traggono conclusioni su qualcosa in base alle sue somiglianze con qualcos'altro. Permette di **astrarre** concetti portandoli da un dominio all'altro e di **essere creativi**.
 per questo è alla base di molti sistemi di AI generativa.

per fare ipotesi.

↓
 Ho Re R₁, trovo C₁

Ho C₁ e Re, trovo R

INTELLIGENZA ARTIFICIALE (AI) (in particolare l'AI "simbolica" o classica)

L'AI simbolica ha strumenti formali che usano "simboli", creare rappresentazioni simboli che (=leggibili dall'uomo) di problemi, logico e ricerca. Per approfondire l'argomento, il prof consiglia la "Bibbia" dell'AI simbolica: "Artificial Intelligence - A modern Approach", di S. Russel e P. Norvig.

Le tematiche principali dell'AI simbolica sono:

- PROBLEM SOLVING:
 - **SEARCH ALGORITHMS**: es Dijkstra
 - **ADVERSARIAL SEARCH & GAMES**: come implementare bot in grado di giocare a giochi livelli di strategia, tipo dama o scacchi
 - **CONSTRAINT SATISFACTION PROBLEMS**: problemi con vincoli, es sudoko. Simili alla Ricerca Operativa, ma non hanno una funzione obiettivo da min(o max)imizzare, quindi tutte le soluzioni vanno bene fintanto che i vincoli sono rispettati.

- KNOWLEDGE, REASONING & PLANNING
 - **FIRST ORDER LOGIC, INFERENCE**: es dimostratori di teoremi
 - **PROLOG**: es sistemi esperti

- UNCERTAIN KNOWLEDGE
 - Conoscenza probabilistica, gli eventi non sono V o F, ma hanno delle probabilità
 - **PROBABILISTIC REASONING**
 - **BAUERIAN NETWORKS**: permettono di utilizzare con il teorema di Bayes, e andare verso nella diagnosi re da: malattia \Rightarrow probabilità dei vari sintomi sintomi \Rightarrow probabilità delle malattie

- **MULTI-AGENTS**: reti di AI autonome che collaborano per risolvere problemi, specializzandosi in task comunicando e adattandosi.

- **MACHINE LEARNING** → **CLASSIFICAZIONE**, T: saranno l'80% del corso

parte che studieremo nel
corso

REGRESSIONE,
CLUSTERING

NEURAL NETWORKS e
DEEP LEARNING : Ne faremo un po'. Si studia nel corso di
Deep learning opzionale con Ferrara.

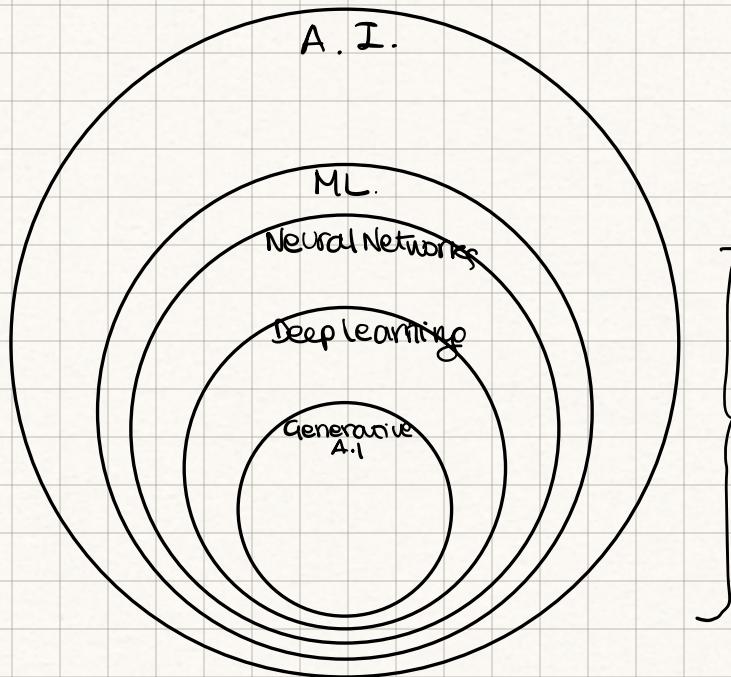
GENERATIVE AI &
LARGE LANGUAGE
MODELS

: Ne parleremo proprio pochissimo. Forse
attiveranno un corso opzionale di NLP

def MODELLO di M.L. : apprende, tramite **training**, a partire da esempi.
Successivamente generalizza e gestisce nuovi dati nello **STESO DOMINIO
APPLICATIVO**



Può gestire dati solo se provengono dallo stesso tipo di sorgente (es:
classificazione di immagini, risponde "cane" e "gatto", se dai una foto
di un pesce va la su gestire correttamente).

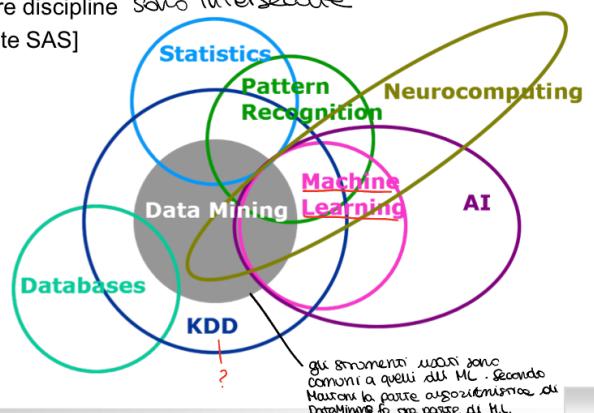


questa è l' **A.I. simbolica**

Per l'A.I., ie M.L. è fondamentale
perché rende possibile esplorare
i dati senza la necessità di
programmazione esplicita.

la **BRUTE-FORCE** (= ricerca esauri-
tiva) permette a volte di risolvere
in modo ottimo i problemi

ML e altre discipline sono intersecate
[Fonte SAS]



ML

semplicemente elenquando e valutando tutte le possibili alternative (scartando quelle non accettabili). Nella maggior parte dei casi, non è computazionalmente gestibile.

Alcune tecniche di ricercae non esauriente prendono il nome di **Weak AI**, hanno la caratteristica di non avere capacità di ragionamento e comprensione.

Una particolare famiglia di algoritmi denominata **EVOLUTIONARY COMPUTING** (che comprende **GENETIC ALGORITHMS**) è oggetto di grande interesse nell'ambito dell'intelligenza artificiale. Un po' come nell'evoluzione Darwiniana, gli algoritmi si "accoppiano" e creano nuovi algoritmi, e solo i migliori si accoppiano successivamente.

Per valutare alcuni modelli di AI si usava il **Test di Turing**, che definiva i modelli "intelligenti" se in grado di essere scambiati per esseri umani. Ormai è obsoleto, e per queste valutazioni si usano benchmark come:

- **BIG-bench (Beyond the Imitation Game)**: (2022) con 200 task su traditional NLP, logic, Math, Coding, competenze scientifiche e umane.
- **GPTA - Google-Proof Q&A** (2023): 448 domande al livello di dottorato.
- **HLE - Humanity's Last Exam** (2025): 2500 problemi difficili su 100 tematiche.

L'AI viene impiegata in giochi (es: scacchi, domino). Le mosse danno un punteggio es: +100 per vittoria, -100 per sconfitta. Si valutano sequenze di mosse possibili scendendo in profondità. I cammini vengono espansi con un algoritmo di **minmax** che consente di scegliere sempre la mossa ottima. Alternativamente vengono usati algoritmi di **pruning**. Esempi: DEEP BLUE (IBM, 1997) aveva una profondità di 6-8 livelli. Computer moderni ponono arrivare fino a 30 livelli. Questo tipo di AI vengono prima trainati su mosaici di campioni e poi giocano contro loro stessi. Altro esempio: WATSON (IBM, 2011) vince contro i campioni del quiz televisivo Jeopardy. In questo caso non aveva di livelli, era stato trainato 400 Terabytes di informazioni, che aveva caricato in RAM. Altro esempio: ALPHAGO (Google, 2016): vinte contro un campione di Go, usava la **ricerca ad albero Monte Carlo (MCTS)** per la selezione delle mosse, guidato da 2 deep neural networks.

Le ML viene anche applicato ai videogames: DEEP MIND (Google, 2013) dimostrò la possibilità di apprendere abilità super-human in numerosi giochi arcade per la console Atari. L'input che ha usato era costituito da 4 frame (immagini) e non da dati, quindi funzionava come un giocatore umano che "guarda" lo schermo. Altro es: nel 2017 OpenAI sviluppò un bot

capace di battere un giocatore professionista al gioco Dots 2 (genere MOSS)

È applicata anche alla robotica, es RoboCup (1997-2050) o la DARPA Challenge

LE STAZIONI dell'AI: chiamate così perché si alternano momenti di grande crescita (=primavera) a momenti di calo (=inverno).

1940-1974 NASCITA E ANNI D'ORO: dopo i primi calcolatori elettrici e i relè, lo sviluppo di computazione di Turing e Teori di Turing, vengono sviluppati i neuroni artificiali (1943). In questo periodo (1956) viene coniato il termine INTELLIGENZA ARTIFICIALE.

- si ottengono i primi risultati nel SYMBOLIC REASONING (=AI non basata sui neuroni artificiali) (es: natural language processing).

1974-1980 PRIMO INVERNO: I risultati ottenuti in precedenza non sono utili a livello pratico. C'è una riduzione dei finanziamenti.

- la scarsa capacità computazionale e l'esposizione combinatoria dei problemi sono ostacoli enormi.

- c'è un ridimensionamento connessionistico (=basato su reti neurali).

1980-1987 NUOVA PRIMAVERA:

- Nascita di nuovi sistemi esperti (creati dall'unione di conoscenza e regole logiche).
- Viene inventato da Rumelhart, Hinton e Williams l'algoritmo di Backpropagation (1986) che riporta l'attenzione sulle reti neurali.
- Vengono finanziate le ricerche sulla II GEN. di calcolatori

1987-1993 SECONDO INVERNO:

- la II gen di calcolatori è un flop, i finanziamenti sono interni, anche perché si diffondono i PC, che sono visibili hardware specializzati.

Le reti neurali ancora non scalano con problemi complessi, e danno risultati concreti solo in campi specifici.

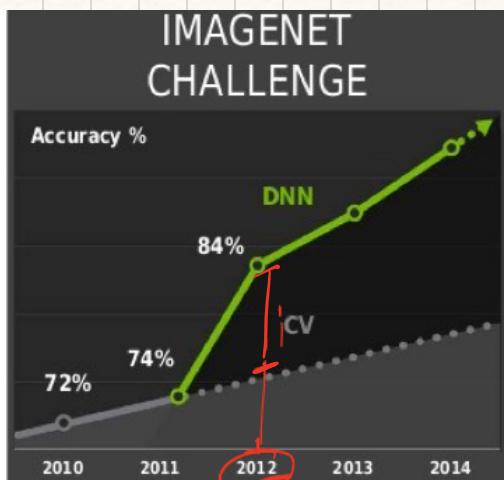
1993-2011 TEMPI MODERNI: introduzione di molti algoritmi

- BAYESIAN NETWORKS, INTELLIGENT AGENTS
- SVM, RANDOM FOREST, ADABOOST
- HMM

Si ha successo in numerose discipline, come Computer Vision, sistemi biometrici, speech recognition, robotica, guida automatica,...

2011-OGLI DEEP LEARNING: (=tecniche basate su reti neurali, sviluppate a seguito dell'impulso dato dal settore della Computer Vision).

- CNN (= Convolutioned Neural Network) introdotte da Yann LeCun (1989) non hanno subito successo, in quanto hanno bisogno di Big Data e elevata potenza di calcolo. Permettono di addestrare reti con molti livelli (DEEP) e milioni di parametri.
- la rivoluzione in Computer Vision (2012) si ha quando la CNN "AlexNet" vince lo ImageNet challenge, una gara di object classification e detection su milioni di immagini e 1000 classi.



Dal grafico poniamo vedere nel 2012 il distacco di AlexNet e gli altri partecipanti

che prima miglioravano anno dopo anno in maniera lineare.

- Grandi progressi anche in Speech Recognition (2016) e la Language Translation negli stessi anni. In particolare la traduzione riceve ampi fenomeni, anche grazie al suo impiego nella Difesa e nello spionaggio.

- Dal 2018 si inizia ad addestrare gli LLM (Large Language Models) modelli generativi basati su reti neurali profonde di tipo Transformer.
 - GPT-3 (OpenAI 2020) → ChatGPT (OpenAI 2022) addestrato su 45 TB di testo. Lavorare il testo e fare embedding richiede reti neurali diverse da quelle usate per lavorare i numeri.
 - APT-4 (2023)

Emergono capacità di reasoning inaspettate quando la scala e l'addestramento aumentano.

Le tecniche di deep learning raggiungono e superano lo stato dell'arte in molti campi:

PERCEPTION

- Object detection e localization
- face recognition
- autonomous car e drones
- speech recognition

LANGUAGE

- LLM
- translation, summarization, Q/A
- generazione di immagini da testo
- copilot
- recommendation

MEDICINE, BIOCHEMISTRY, HUMAN COMPUTER INTERFACES

- medical image analysis
- protein folding, biomolecule combination
- brain implants.

systems.

La **SINGOLARITÀ TECNOLOGICA** è il momento in cui lo sviluppo dell'AI non è più in crescita lineare, ma esponenziale, e non più controllabile non più prevedibile dall'uomo. Ray Kurzweil, inventore e futurista, prevede che avverrà nel 2045.

Verrà raggiunta perché:

- abbiamo hardware sempre più potenti,
- c'è **AGI** (Artificial General Intelligence)

Alcuni personaggi influenti sono preoccupati:

- 2015 Open letter on Arti