

Apprendimento Automatico (Intro)

Docente: Prof. Fabio Aioli

www.math.unipd.it/~aioli

Collaboratore: Dott. Mirko Polato

www.math.unipd.it/~mpolato

Sito web del corso

www.math.unipd.it/~aioli/corsi/1920/aa/aa.html

Orario delle lezioni

✓ 40 ore di lezione in aula (5cfu)

✓ 8 ore di laboratorio (1cfu)

✓ AULA: Lunedì dalle 12:30 alle 14:30 e
Mercoledì dalle 14:30 alle 16:30, Aula
~~1BC50~~ (Torre Archimede)

✓ LABORATORI: Da definire (verso la fine del
corso)

Testi e Esame

- T. Mitchell, "Machine Learning", McGraw Hill, 1998 (disponibile in biblioteca)
- E. Alpaydin, "Introduction to Machine Learning", Cambridge University Press, 2010
- C.M. Bishop, "Pattern Recognition e Machine Learning" Springer, 2006
- Ian Goodfellow, Yoshua Bengio, Aaron Courville, "Deep Learning" MIT Press, 2016

- Slides e altro materiale presentato nel corso

- Esame 1/4 progetto (attività, PE) e 3/4 scritto (+ orale opzionale)

Attività del corso (punti esperienza)

Non necessariamente un reale progetto!

Consisterà nell'applicazione di tecniche di machine learning su uno o più dataset presi da kaggle <https://www.kaggle.com> o raccolti su un repository nostro. La valutazione sarà incrementale (a cadenza mensile).

E' premiata ogni [attività collaborativa](#) (tramite moodle)

- descrizione di nuovi task o datasets
- relazione su attività svolte e risultati ottenuti
- revisione di attività svolte da colleghi

P.e. Kannada MNIST

<https://www.kaggle.com/c/Kannada-MNIST>

Per info iscriversi a.s.a.p. al MOODLE

In alternativa è sempre possibile fare normale progetto dopo la fine del corso.

Da dove cominciare.. per non rimanere indietro

Algebra Lineare (vettori e matrici):

- Youtube ('linear algebra for machine learning')
- Capitolo 2 (Goodfellow et al.)

Probabilità:

- Appendice A (Alpaydin)
- Capitolo 3 (Goodfellow et al.)

Python e librerie scientifiche:

- <http://www.python.it/doc/>
- Aiolli. *Appunti di programmazione (scientifica) in Python*, Esculapio
- Guardare numpy, scipy, matplotlib, sklearn: tutti moduli già disponibili in Google CoLab
- Avremo un laboratorio dedicato prossimamente

Ragionamento Deduttivo

- **Deduttivo (Aristotele 384 aC – 322 aC)**
 - REGOLA($C \rightarrow R$): Tutti gli uomini sono mortali
 - CASO(C_1): Socrate è un uomo
 - Quindi..
 - RISULTATO(R_1): Socrate è mortale
- Il ragionamento deduttivo è il fondamento di gran parte delle dimostrazioni e teoremi della matematica. Si parte da assunzioni vere e si inferiscono nuovi teoremi mediante dimostrazione. Molto usata anche in IA classica.

Ragionamento Induttivo

- Induttivo (F. Bacon, filosofo 1561–1626)
 - CASO(C_1): Socrate è un uomo
 - RISULTATO(R_1): Socrate morì
 - Quindi..
 - REGOLA($C \rightarrow R$): Tutti gli uomini sono mortali
- Nel ragionamento induttivo le premesse (casi particolari) forniscono evidenza a sostegno della conclusione (generalizzazione) ma non ne garantiscono la verità (possiamo avere premesse vere, conclusioni false). Quello che usiamo in ML e quello usato dagli scienziati in generale.

Ragionamento Abduittivo

- **Abduittivo (C.S. Peirce, 1839–1914)**
 - REGOLA($C \rightarrow R$): Tutti gli uomini sono mortali
 - RISULTATO(R_1): Socrate morì
 - Quindi..
 - CASO(C_1): Socrate è un uomo
- Nel ragionamento abduittivo invece di generalizzare ci muoviamo «lateralmente», ipotizzando che un implicazione valga anche al contrario. Quello usato da Sherlock Holmes e Dr. Gregory House per intenderci.



Dr House

- DEDUZIONE

- REGOLA: Tutti i malati di lupus muoiono in 5 gg
- CASO: Il paziente ha il lupus
- RISULTATO: Il paziente morirà entro 5 gg

- INDUZIONE

- CASO: Il paziente ha il lupus
- RISULTATO: Il paziente è morto in 5 gg
- REGOLA: Tutti i malati di lupus muoiono in 5 gg

- ABDUZIONE

- REGOLA: Tutti i malati di lupus muoiono in 5 gg
- RISULTATO: Il paziente è morto in 5 gg
- CASO: Il paziente aveva il lupus

Cos'è il Machine Learning?

« Il Machine Learning è il campo di studio che dà ai computer l'abilità di apprendere (a realizzare un compito) senza essere esplicitamente programmati a farlo »

Arthur Samuel, 1959

«Un programma impara da un'esperienza E rispetto a dei compiti T ottenendo una performance P , se quest'ultima migliora con l'esperienza E »

Tom Mitchell

Domande chiave..
(a cui tenteremo di dare risposte)

Quando e perché è utile un approccio
basato su Machine Learning?

Come si può apprendere?

Ma la più importante di tutte (a cui daremo
risposta in seguito):

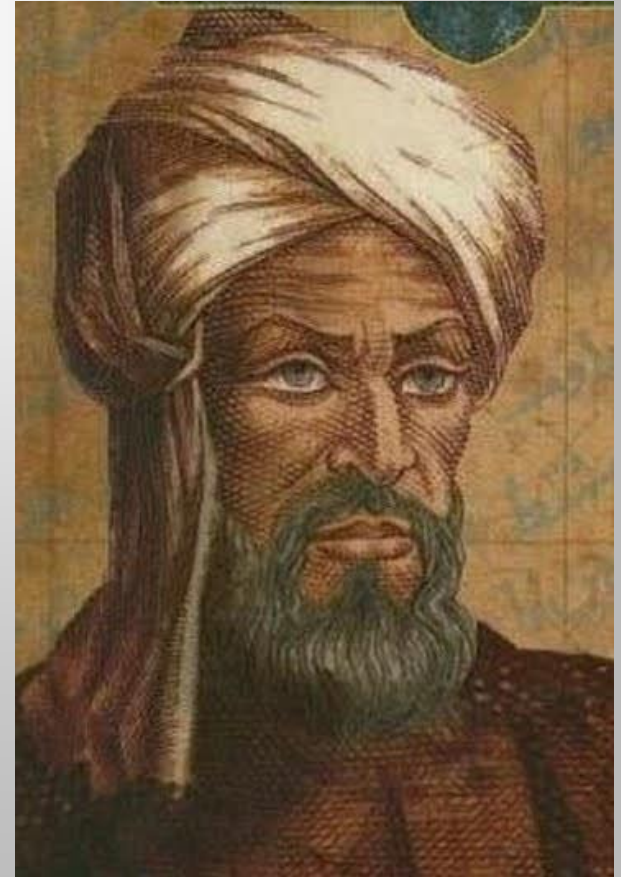
Possiamo veramente apprendere?

Algoritmo

WIKIPEDIA: Un **algoritmo** è un procedimento che risolve un determinato problema attraverso un numero finito di passi elementari. Il termine deriva dalla trascrizione latina del nome del matematico persiano al-Khwarizmi, che è considerato uno dei primi autori ad aver fatto riferimento a questo concetto. L'algoritmo è un concetto fondamentale dell'informatica, anzitutto perché è alla base della nozione teorica di calcolabilità: un problema è calcolabile quando è risolvibile mediante un algoritmo. Inoltre, l'algoritmo è un concetto cardine anche della fase di programmazione dello sviluppo di un software: preso un problema da automatizzare, la programmazione costituisce essenzialmente la traduzione o codifica di un algoritmo per tale problema in programma, scritto in un certo linguaggio, che può essere quindi effettivamente eseguito da un calcolatore.

Più formalmente...

*«una sequenza ordinata e finita di passi (operazioni o istruzioni) elementari che conduce a un **ben determinato** risultato in un tempo finito»*



Cos'è un algoritmo?

« Procedura che risolve un determinato problema attraverso un numero finito di passi elementari »

Wikipedia



INGREDIENTI:

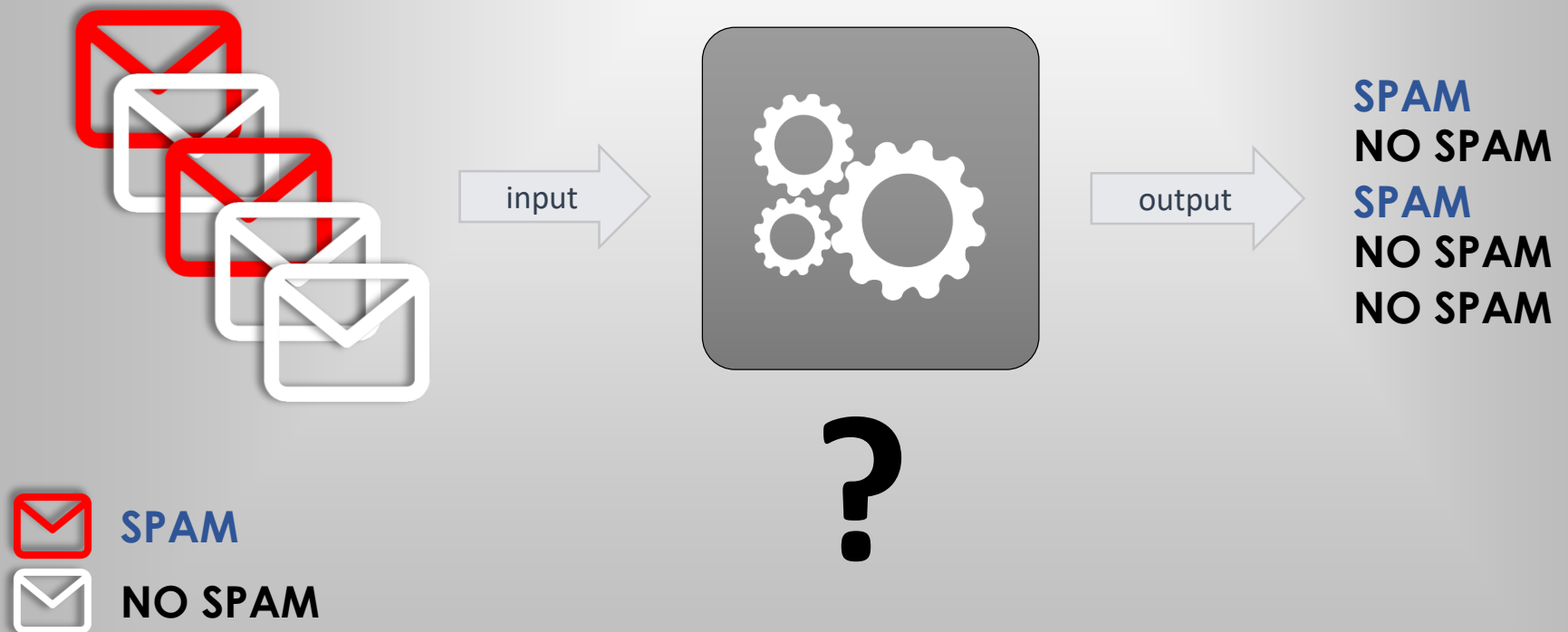
- 300 g. di farina di riso consentita
- 150 g. di burro a temperatura ambiente
- 100 g. di zucchero
- 2 tuorli
- un cucchiaino di marsala
- pizzico di sale

PROCEDIMENTO:

Impastare tutti gli ingredienti e mettere in frigorifero per 1 ora.
Poi stendere la pasta e mettere in una tortiera.
Stendere la marmellata.
Mettere in forno già caldo a 180° per circa 20 minuti.

PASTA FROLLA

Quale algoritmo?



Perché no un approccio algoritmico?

Per uno o più dei seguenti motivi:

- ✓ Impossibilità di formalizzare esattamente il problema
- ✓ Rumore e/o incertezza in input/output
- ✓ Alta complessità della soluzione
- ✓ Inefficienza della soluzione
- ✓ Mancanza di conoscenza 'compilata' sul problema da risolvere

Quando è importante l'apprendimento?

Quando il sistema deve:

- ✓ **Adattarsi** all'ambiente in cui opera (anche personalizzazione automatica)
- ✓ **Migliorare** le proprie prestazioni rispetto ad un particolare compito
- ✓ **Scoprire** regolarità e nuova informazione (conoscenza) a partire da dati empirici

Dati Vs. Conoscenza

In ML, si studiano metodi per trasformare l'informazione empirica (presente nei **dati**) in nuova **conoscenza**

Grazie all'evoluzione dei computer e le reti, i dati oramai sono presenti ogni dove, e abbondanti!

- ✓ Scontrini di una catena di supermercati,
- ✓ Contenuto di pagine web,
- ✓ E-commerce
- ✓ Transazioni bancarie, ecc.

Assunzione fondamentale

Esiste un processo (stocastico) che spiega i dati che osserviamo. Magari non ne conosciamo i dettagli, ma esiste!

P.e. comportamento sociale non è puramente casuale

Lo scopo dell'apprendimento è costruire buone (o meglio, utili) approssimazioni di questo processo.

Il main goal del ML

L'obiettivo finale del ML è definire dei criteri di performance e ottimizzarli usando i dati o esperienza pregressa

I modelli saranno definiti su parametri che vogliamo **apprendere** ottimizzando un dato **criterio**. Usando dati e, eventualmente, conoscenza a priori sul dominio

- ✓ **Modelli Predittivi** (predizioni sul futuro)
- ✓ **Modelli Descrittivi** (ottenere nuova conoscenza)

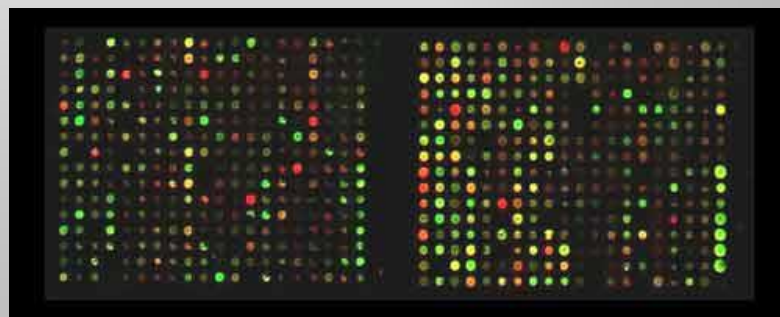
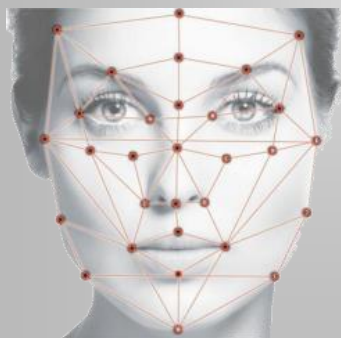
Esempi di Applicazioni (1)



In 1917, Einstein applied the general theory of relativity to model the large-scale structure of the universe. He was visiting the United States when Adolf Hitler came to power in 1933 and did not go back to Germany, where he had been a professor at the Berlin Academy of Sciences. He settled in the U.S., becoming an American citizen in 1940. On the eve of World War II, he endorsed a letter to President Franklin D. Roosevelt alerting him to the potential development of "extremely powerful bombs of a new type" and recommending that the U.S. begin similar research. This eventually led to what would become the Manhattan Project. Einstein supported defending the Allied forces, but largely denounced using the new discovery of nuclear fission as a weapon. Later, with the British philosopher Bertrand Russell, Einstein signed the Russell-Einstein Manifesto, which highlighted the danger of nuclear weapons. Einstein was affiliated with the Institute for Advanced Study in Princeton, New Jersey, until his death in 1955.

Tag colours:

LOCATION TIME PERSON ORGANIZATION MONEY PERCENT DATE



Esempi di Applicazioni (2)

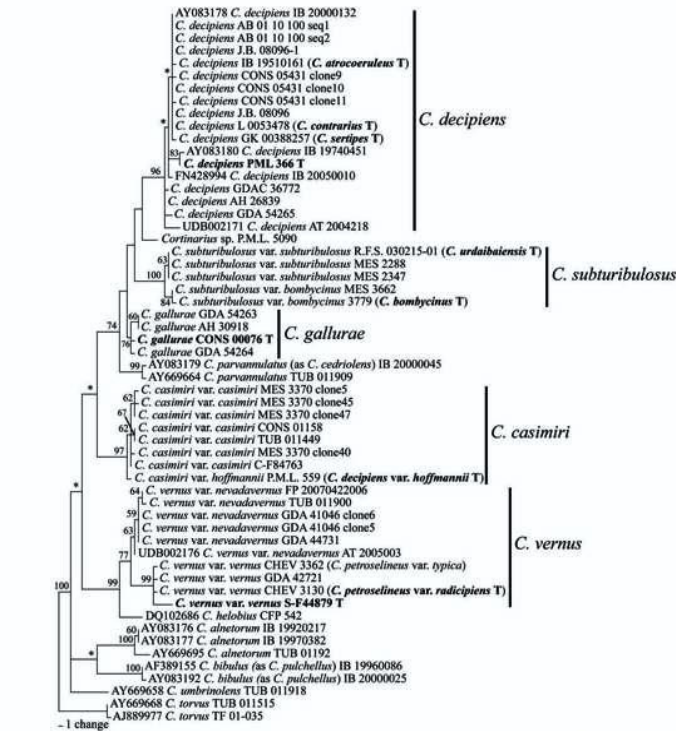
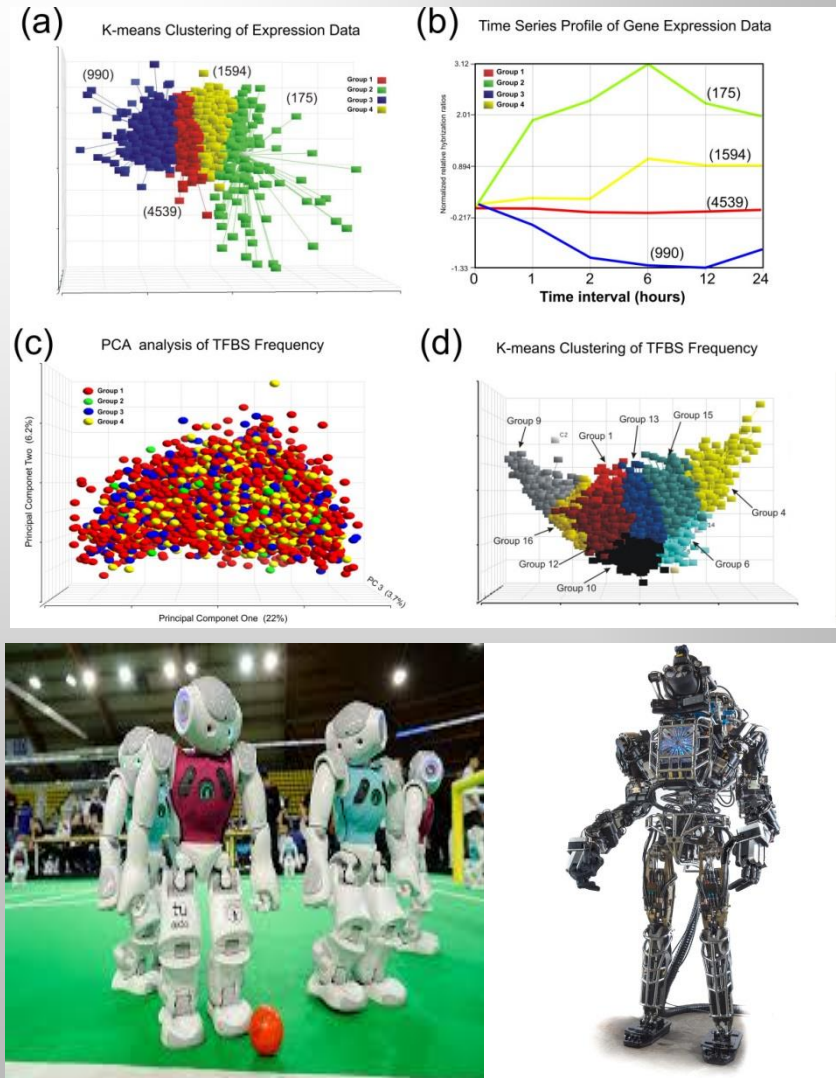
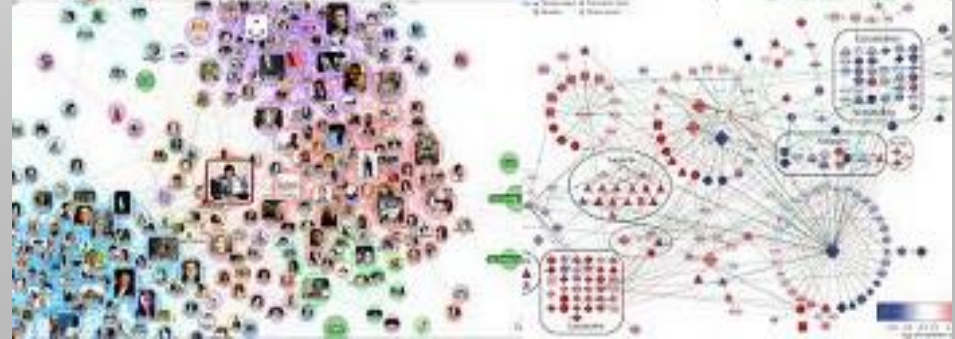
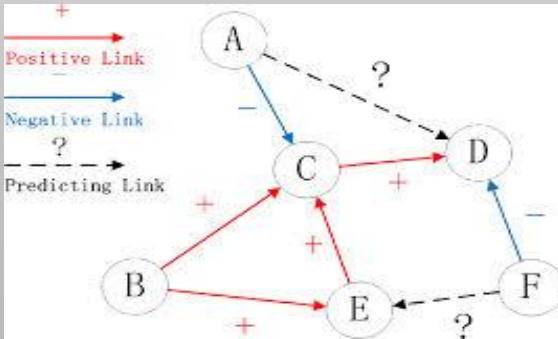
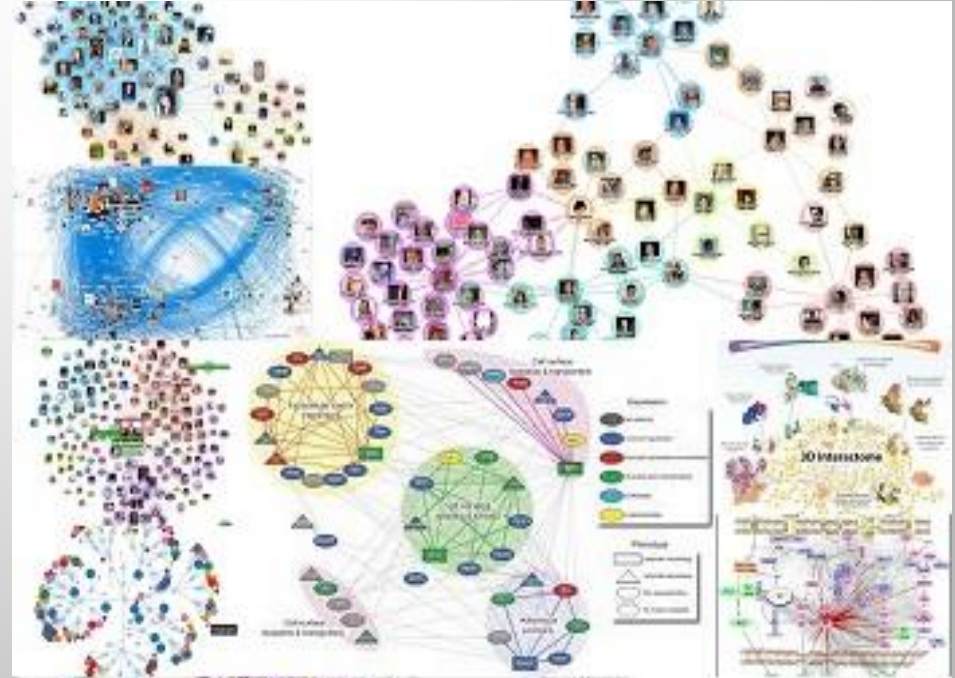


Fig 1 – Phylogenetic relationships of species of Cortinariaceae sect. *Hydrocybe* (subg. *Telamonia*) sensu Brandrud et al. (1992) and Bidaud et al. (1994). Phylogram of one of the most parsimonious trees (length: 173; CI: 0.775; RI: 0.937) obtained from the parsimony analysis. Bootstrap values $\geq 50\%$ are shown above branches. Branches collapsing in the strict consensus tree are marked with an asterisk. The herbarium references (for all sequences) and accession number (for the sequences taken from GenBank database and which specimens were not included in the morphological analysis) are shown after and before each taxon name, respectively. The bolded names followed by "T" represent sequences obtained from type specimens. Species studied are indicated at right.



Esempi di Applicazioni (3)



Esempi di Applicazioni

- Riconoscimento di Facce

- Controllo degli accessi da registrazioni video o fotografiche.
Quali sono le caratteristiche veramente rilevanti di una faccia?

- Named Entity Recognition

- Il problema di identificare entità in una frase: luoghi, titoli, nomi, azioni, ecc. Partendo da un insieme di documenti già marcati/taggati

- Classificazione di documenti

- Decidere se una email è spam o meno, dare una classificazione ad un documento tra un insieme di topics (sport, politica, hobby, arti, ecc.) magari gerarchicamente organizzati

Esempi di Applicazioni

- Giochi e Profilazione Avversario

- Per alcuni giochi ad informazione incompleta (giochi di carte, geister, risiko, ...) vogliamo predire l'informazione mancante basandosi sulle strategie che l'avversario ha usato nel passato (minacce, reazioni, ecc.).

- Bioinformatica

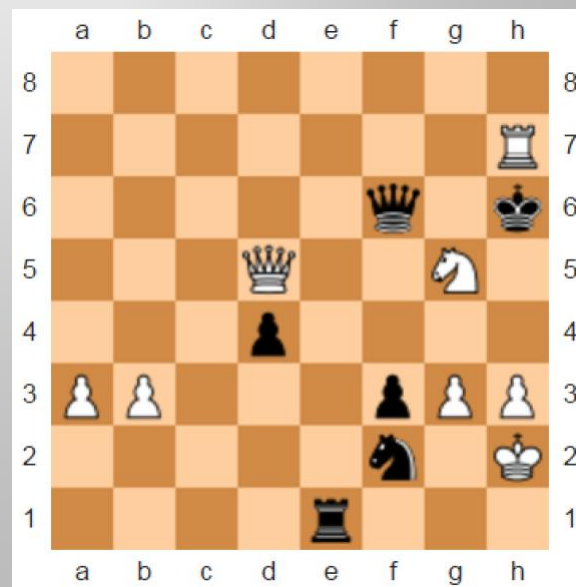
- I microarray sono dispositivi che rilevano l'espressione genica da un tessuto biologico. E' possibile a partire da questi determinare la probabilità che un paziente reagisca in modo positivo ad una certa terapia? ...

- Speech Recognition, Handwritten Recognition, Social Network Analysis, e molto altro ancora.

Deep Blue vs Garry Kasparov

- 1997 Deep Blue (IBM) vince a scacchi contro il campione del mondo Garry Kasparov
- Hardware: capace di valutare 200M di disposizioni sulla scacchiera al sec. La potenza di calcolo (11.38 GFLOPs) rilevante all'epoca è inferiore a quella di un moderno smartphone
- Ricerca in profondità (6-8 livelli in media)
- Valutazione dello score complessa e con molti parametri. I valori ottimali per questi parametri appresi dal sistema stesso (learning), analizzando migliaia di partite di campioni
- La lista di aperture fu fornita da campioni di scacchi

1996 – Prima partita persa: Kasparov (nero) si ritira



Google DeepMind vs Lee Sedol

- 2016 AlphaGo (Google) batte il campione Lee Sedol a GO
- GO antico gioco cinese, regole semplici ma molte più mosse rispetto agli scacchi. Ancora più difficile applicare forza bruta.
- AlphaGo non usa metodi euristici ma è basato pesantemente su ML
- DeepNN supervisionate per imitare le mosse di professionisti a partire da partite memorizzate su Go Servers (30M mosse)
- + Self training giocando milioni di partite contro se stesso (Reinforcement Learning)

2016 – AlphaGO vs Lee Sedol (4-1)



Generative Adversarial Learning



MNIST handwritten digit dataset, (b) Toronto Face Database,
(c-d) CIFAR-10 small object photograph dataset
[Goodfellow et al. 2014]

Altro su applicazioni di ML su TED talk...

The wonderful and terryfying implications of computers that can learn

[Jeremy Howard]



How we're teaching computers to understand pictures

[Fei Fei Li]



Contenuti del corso

- 1) Introduzione al SL
- 2) Basi teoriche: PAC e SRM
- 3) Alberi di decisione
- 4) Reti Neurali
- 5) GLM e SVM
- 6) Preprocessing e feature selection
- 7) Model Selection e valutazione
- 8) Rappresentazione (kernels, embeddings, CNNs)
- 9) Metodi Bayesiani
- 10) Ensemble
- 11) Clustering
- 12) Sistemi di raccomandazione