



7-8 Maggio 2013

UNIBO – CSR

Corso di Intelligenza Artificiale

*Progettazione e sviluppo di
applicazioni industriali basate su
Machine Learning*

Matteo Roffilli
roffilli@gmail.com



Agenda - Day 1

- Introduzione
- Applicazioni accademiche vs industriali
- Intelligenza Artificiale
- Machine Learning
- Un progetto concreto
- Tips and tricks
- Riferimenti, tesi, tirocini e stage

Introduzione

- Contestualizziamo ...
- Progettazione?
- Sviluppo?
- Applicazioni?
- Industriali?
- Machine Learning?

Le immagini recuperate da Internet ai fini didattici, eventualmente coperte da Copyright, sono proprietà dei legittimi titolari.

Università vs Industria

Tempo infinito

Risorse limitate

Scoprire e conoscere

Nessun vincolo legale

1 user

Obiettivo: pubblicare

Rischio \$\$ zero

Tempo pianificato

Risorse ampie

Produrre

Molti vincoli legali

1'000'000 users

Obiettivo: vendere

Rischio \$\$ alto

Applicazione industriale



Applicazione industriale ... continua



Applicazione industriale ... continua



Artificial Intelligence vs Machine Learning

"Can machines think?"

AI



"Can machines do what we (as thinking entities) can do?"

ML



Machine Learning: molti obiettivi

**La miglior riproduzione
artificiale delle funzioni
di una foglia**



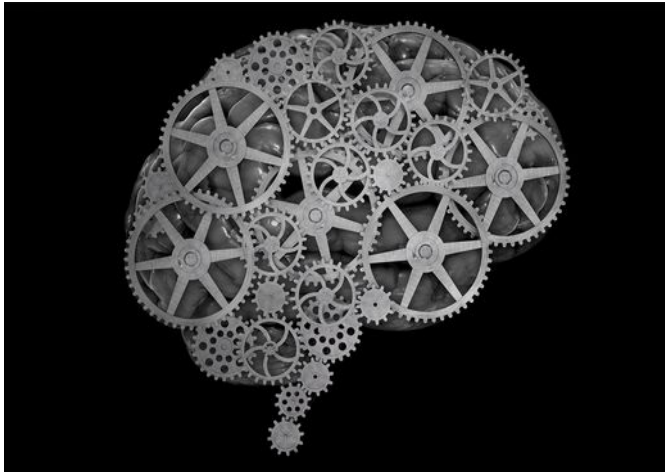
**La miglior riproduzione
artificiale delle funzioni
di un albero**



[wiktionary] *artificiale*: prodotto con mezzi tecnologici a imitazione del naturale

Machine Learning: molti approcci

Computational learning theory



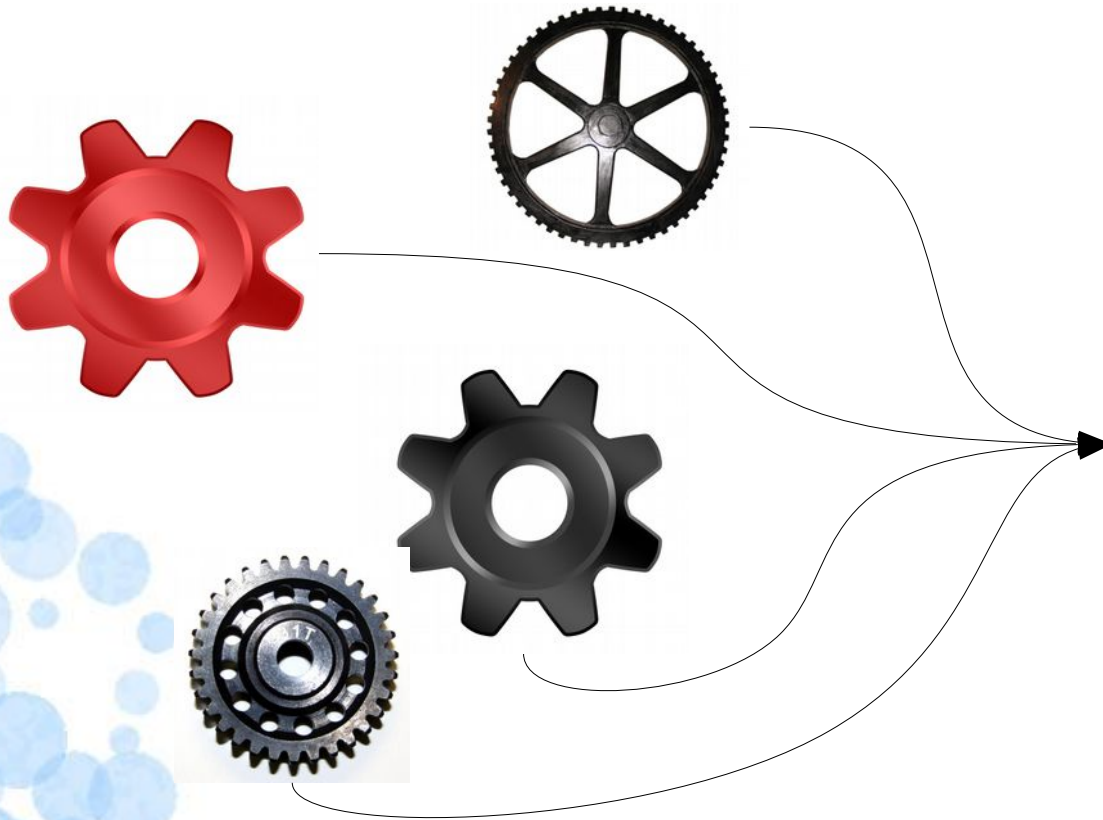
embodiment

Applicazioni ML



Machine Learning industriale

Ricerca universitaria



Framework industriale



Un progetto concreto

...dalla teoria alla pratica....

**Sviluppiamo un progetto
industriale
(*non troppo*) immaginario**

La Machine



Machine: definizione obiettivo

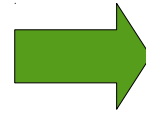
Produrre una *Machine* che si accenda (start) in modo autonomo/automatico se caricata con materiale lavabile



Produrre una *Machine* che NON si accenda (stop) in modo autonomo/automatico se caricata con materiale NON lavabile

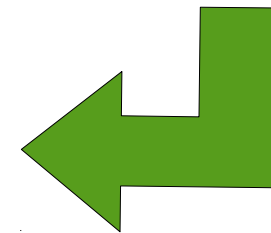
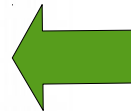


Machine: approccio “classico”



Input: A dataset of graphs D , a distance threshold δ , a minimum support threshold min_sup
 Output: A representative set

- 1: Call CloseGraph (w.r.t. min_sup) to obtain the set of δ -jump patterns JP and the set of closed frequent subgraphs CF ;
- 2: **for** each δ -jump pattern R in JP **do**
- 3: **for** each subgraph P in CF **do**
- 4: **if** P is δ -covered by R **then**
- 5: Put P into E ; /* E is the set of subgraphs covered by JP */
- 6: Output R ;
- 7: $\bar{E} = CF - E$; /* \bar{E} is the set of remaining uncovered subgraphs */
- 8: **for** each subgraph P in \bar{E} **do**
- 9: **for** each candidate representative R in $CF - JP$ **do**
- 10: **if** P is δ -covered by R **then**
- 11: Put P into $Set_\delta(R)$;
- 12: **while** $\bar{E} \neq \phi$ **do**
- 13: $R^* = \operatorname{argmax}_{R \in (CF - JP)} |Set_\delta(R)|$;
- 14: **for** each subgraph P in $Set_\delta(R^*)$ **do**
- 15: Remove P from \bar{E} and other $Set_\delta(R')$ ($R' \in (CF - JP)$);
- 16: Output R^* ;



Machine: approccio "learning"



minimize:

$$W(\alpha) = -\sum_{i=1}^{\ell} \alpha_i + \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y_i y_j \alpha_i \alpha_j \mathbf{x}_i \mathbf{x}_j$$

subject to:

$$\sum_{i=1}^{\ell} y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C$$



k_1	k_2	$w^{[k_1]}$	$w^{[k_2]}$	$w^{[k_1, k_2]}$	$\mathbf{x}^{[k_1, k_2]}$	$\phi(\mathbf{x}^{[k_1, k_2]})$	$w^{[k_1, k_2]} \phi(\mathbf{x}^{[k_1, k_2]})$
[1]	[2]	[3]	[4]	[5]=[3][4]	[6]	[7]	[8]=[5][7]
0	0	1/6	1/4	1/24	(2.00, 0.00)	0.058550	0.002440
0	1	1/6	4/4	4/24	(2.00, 0.75)	0.093562	0.015594
0	2	1/6	2/4	2/24	(2.00, 1.50)	0.085190	0.007099
0	3	1/6	4/4	4/24	(2.00, 2.25)	0.044196	0.007366
0	4	1/6	1/4	1/25	(2.00, 3.00)	0.013064	0.000544
1	0	4/6	1/4	4/24	(2.50, 0.00)	0.016775	0.002796
1	1	4/6	4/4	16/24	(2.50, 0.75)	0.039003	0.026002
1	2	4/6	2/4	8/24	(2.50, 1.50)	0.051670	0.017223
1	3	4/6	4/4	16/24	(2.50, 2.25)	0.039003	0.026002
1	4	4/6	1/4	4/24	(2.50, 3.00)	0.016775	0.002796
2	0	2/6	1/4	2/24	(3.00, 0.00)	0.002915	0.000243
2	1	2/6	4/4	8/24	(3.00, 0.75)	0.009861	0.003287
2	2	2/6	2/4	4/24	(3.00, 1.50)	0.019008	0.003168
2	3	2/6	4/4	8/24	(3.00, 2.25)	0.020877	0.006959
2	4	2/6	1/4	2/24	(3.00, 3.00)	0.013064	0.001089
3	0	4/6	1/4	4/24	(3.50, 0.00)	0.000307	0.000051
3	1	4/6	4/4	16/24	(3.50, 0.75)	0.001512	0.001008
3	2	4/6	2/4	8/24	(3.50, 1.50)	0.004241	0.001414
3	3	4/6	4/4	16/24	(3.50, 2.25)	0.006778	0.004518
3	4	4/6	1/4	4/24	(3.50, 3.00)	0.006171	0.001029
4	0	1/6	1/4	1/24	(4.00, 0.00)	0.000020	0.000001
4	1	1/6	4/4	4/24	(4.00, 0.75)	0.000141	0.000023
4	2	1/6	2/4	2/24	(4.00, 1.50)	0.000574	0.000048
4	3	1/6	4/4	4/24	(4.00, 2.25)	0.001335	0.000222
4	4	1/6	1/4	1/24	(4.00, 3.00)	0.001768	0.000074
							0.130995

“learning” cosa?



minimize:

$$W(\alpha) = -\sum_{i=1}^{\ell} \alpha_i + \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y_i y_j \alpha_i \alpha_j \mathbf{x}_i \mathbf{x}_j$$

subject to:

$$\sum_{i=1}^{\ell} y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C$$



k_1	k_2	$w^{[k_1]}$	$w^{[k_2]}$	$w^{[k_1, k_2]}$	$\mathbf{x}^{[k_1, k_2]}$	$\phi(\mathbf{x}^{[k_1, k_2]})$	$w^{[k_1, k_2]} \phi(\mathbf{x}^{[k_1, k_2]})$
[1]	[2]	[3]	[4]	[5]=[3][4]	[6]	[7]	[8]=[5][7]
0	0	1/6	1/4	1/24	(2.00, 0.00)	0.058550	0.002440
0	1	1/6	4/4	4/24	(2.00, 0.75)	0.093562	0.015594
0	2	1/6	2/4	2/24	(2.00, 1.50)	0.085190	0.007099
0	3	1/6	4/4	4/24	(2.00, 2.25)	0.044196	0.007366
0	4	1/6	1/4	1/25	(2.00, 3.00)	0.013064	0.000544
1	0	4/6	1/4	4/24	(2.50, 0.00)	0.016775	0.002796
1	1	4/6	4/4	16/24	(2.50, 0.75)	0.039003	0.026002
1	2	4/6	2/4	8/24	(2.50, 1.50)	0.051670	0.017223
1	3	4/6	4/4	16/24	(2.50, 2.25)	0.039003	0.026002
1	4	4/6	1/4	4/24	(2.50, 3.00)	0.016775	0.002796
2	0	2/6	1/4	2/24	(3.00, 0.00)	0.002915	0.000243
2	1	2/6	4/4	8/24	(3.00, 0.75)	0.009861	0.003287
2	2	2/6	2/4	4/24	(3.00, 1.50)	0.019008	0.003168
2	3	2/6	4/4	8/24	(3.00, 2.25)	0.020877	0.006959
2	4	2/6	1/4	2/24	(3.00, 3.00)	0.013064	0.001089
3	0	4/6	1/4	4/24	(3.50, 0.00)	0.000307	0.000051
3	1	4/6	4/4	16/24	(3.50, 0.75)	0.001512	0.001008
3	2	4/6	2/4	8/24	(3.50, 1.50)	0.004241	0.001414
3	3	4/6	4/4	16/24	(3.50, 2.25)	0.006778	0.004518
3	4	4/6	1/4	4/24	(3.50, 3.00)	0.006171	0.001029
4	0	1/6	1/4	1/24	(4.00, 0.00)	0.000020	0.000001
4	1	1/6	4/4	4/24	(4.00, 0.75)	0.000141	0.000023
4	2	1/6	2/4	2/24	(4.00, 1.50)	0.000574	0.000048
4	3	1/6	4/4	4/24	(4.00, 2.25)	0.001335	0.000222
4	4	1/6	1/4	1/24	(4.00, 3.00)	0.001768	0.000074
							0.130995

“learning” cosa?

Il “learning” produce un insieme di parametri
che configurano un modello generico
per una specifica applicazione

k_1	k_2	$w^{[k_1]}$	$w^{[k_2]}$	$w^{[k_1, k_2]}$	$x^{[k_1, k_2]}$	$\phi(x^{[k_1, k_2]})$	$w^{[k_1, k_2]} \phi(x^{[k_1, k_2]})$
[1]	[2]	[3]	[4]	[5]=[3][4]	[6]	[7]	[8]=[5][7]
0	0	1/6	1/4	1/24	(2.00, 0.00)	0.058550	0.002440
0	1	1/6	4/4	4/24	(2.00, 0.75)	0.093562	0.015594
0	2	1/6	2/4	2/24	(2.00, 1.50)	0.085190	0.007099
0	3	1/6	4/4	4/24	(2.00, 2.25)	0.044196	0.007366
0	4	1/6	1/4	1/25	(2.00, 3.00)	0.013064	0.000544
1	0	4/6	1/4	4/24	(2.50, 0.00)	0.016775	0.002796
1	1	4/6	4/4	16/24	(2.50, 0.75)	0.039003	0.026002
1	2	4/6	2/4	8/24	(2.50, 1.50)	0.051670	0.017223
1	3	4/6	4/4	16/24	(2.50, 2.25)	0.039003	0.026002
1	4	4/6	1/4	4/24	(2.50, 3.00)	0.016775	0.002796
2	0	2/6	1/4	2/24	(3.00, 0.00)	0.002915	0.000243
2	1	2/6	4/4	8/24	(3.00, 0.75)	0.009861	0.003287
2	2	2/6	2/4	4/24	(3.00, 1.50)	0.019008	0.003168
2	3	2/6	4/4	8/24	(3.00, 2.25)	0.020877	0.006959
2	4	2/6	1/4	2/24	(3.00, 3.00)	0.013064	0.001089
3	0	4/6	1/4	4/24	(3.50, 0.00)	0.000307	0.000051
3	1	4/6	4/4	16/24	(3.50, 0.75)	0.001512	0.001008
3	2	4/6	2/4	8/24	(3.50, 1.50)	0.004241	0.001414
3	3	4/6	4/4	16/24	(3.50, 2.25)	0.006778	0.004518
3	4	4/6	1/4	4/24	(3.50, 3.00)	0.006171	0.001029
4	0	1/6	1/4	1/24	(4.00, 0.00)	0.000020	0.000001
4	1	1/6	4/4	4/24	(4.00, 0.75)	0.000141	0.000023
4	2	1/6	2/4	2/24	(4.00, 1.50)	0.000574	0.000046
4	3	1/6	4/4	4/24	(4.00, 2.25)	0.001335	0.000222
4	4	1/6	1/4	1/24	(4.00, 3.00)	0.001768	0.000074
							0.130995

“learning” come?

Il “learning” analizza come lo stesso nostro obiettivo è stato perseguito da altri e cerca di riprodurne/replicarne/imitarne il funzionamento

Gli altri sono entità in grado di perseguire in modo corretto il nostro obiettivo

Ovvero....

Altri: Machine che perseguono l'obiettivo



start



start



start



stop



stop



start



stop

...

La Machine... all'opera



Machine ... with e w/o learn(ing)



esegue



start

Machine programmata con learning

Machine programmata da esperti



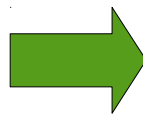
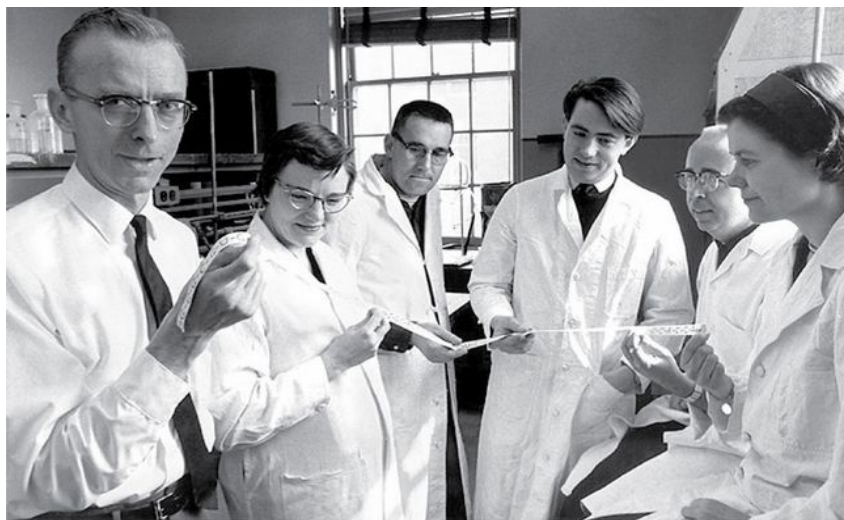
esegue



stand-by

Machine w/out learning

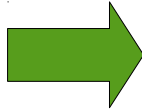
ML: quali vantaggi?



La creazione dell'algoritmo non è guidata (*driven*) dagli esperti/programmatori ma dai dati (*data driven*) processati da risorse HPC.

Si sostituisce energia chimica (cervello) in energia elettrica (CPU).

ML: quali svantaggi?



Il dataset (*data driven*) di supporto deve essere molto molto più grande (*large scale*) da cui deriviamo molte questioni computazionali da gestire

Tips and tricks

Molto rumore per nulla ?

Tips and tricks

Machines can never take us by surprise

Lady Lovelace's Objection

cfr Turing, Alan (October 1950), "Computing Machinery and Intelligence", *Mind* LIX (236): 433–460

Machine Learn(ing) : generalizzazione

stop



Machine Learn(ing) : generalizzazione

stop



Machine Learn(ing): “tutti” possono sbagliare

start



Alcune questioni in sospeso

Come formalizzare l'obiettivo: classificazione, regressione, novelty detection, clustering

Quale modello generico usare: SVM, RVM, Deep Learning, ANN, etc

Come rappresentare l'input degli altri (feature): sensori, immagini, etc

Come rappresentare l'output degli altri: supervisionato, semi-supervisionato, non supervisionato, trasduttivo, con universo, etc

Analizziamo insieme un progetto step-by-step

Raggruppatevi in gruppi di 3-5 studenti

Ideate una possibile applicazione industriale che vorreste vedere sul mercato il prossimo anno

**Inviare via email l'idea (titolo e descrizione)
e i componenti del gruppo**

(nome,cognome,matricola,email) a roffilli@gmail.com

**Domani sceglieremo una proposta e la analizzeremo
insieme mostrando gli strumenti software di
prototipazione utilizzabili**

(OpenCV, Octave, MLDemos, libsvm, etc)

Tesi, tirocini e stage in ML

Sono disponibili:

**Tesi, Tirocini con credito
Stage post laurea retribuiti**

anche in collaborazione con la società:

Bioretics srl

con sedi in Cesena e Cesenatico



Agenda - Day 2

- Scelta dell'idea
- Proof-of-concept
- Prototipo accademico
- Prototipo industriale
- Cenni sulla commercializzazione
- Riferimenti, tesi, tirocini e stage

The Big picture



SVM



RVM



ANN



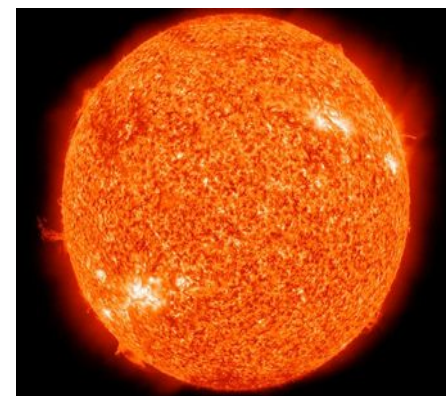
**Deep
Net**



cluster



HPC



dati



CPU



GPU

Tool & Link interessanti

- OpenCV: <http://opencv.org>
- GNU Octave: www.gnu.org/software/octave
- MLDemos: <http://mldemos.epfl.ch>
- LibSVM: www.csie.ntu.edu.tw/~cjlin/libsvm
- UniverSVM: <http://mloss.org/software/view/19>
- Torch: <http://www.torch.ch>
- http://videolectures.net/Top/Computer_Science/Machine_Learning/
- <http://deeplearning.net/>

Tesi, tirocini e stage in ML

Sono disponibili:

**Tesi, Tirocini con credito
Stage post laurea retribuiti**

anche in collaborazione con la società:

Bioretics srl

con sedi in Cesena e Cesenatico

