Machine Learning for Networks: Introduction -

Andrea Araldo

January 7, 2022

Outline

- Artificial intelligence and Machine Learning
- Applications to Communication Networks
- Supervised and Unsupervised Learning
- Data exploration

Section 1

Presentation

The allure of Artificial Intelligence (AI)

AI allows machines to sense, comprehend, act and learn.

...

*AI promises a new era of disruption and productivity, where human **ingenuity** is enhanced by speed and precision*

Accenture website.

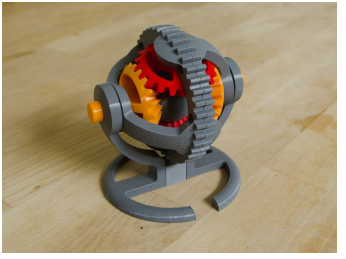
A consultant (or a researcher) today.

The AI allure

Common vision of AI

The AI allure

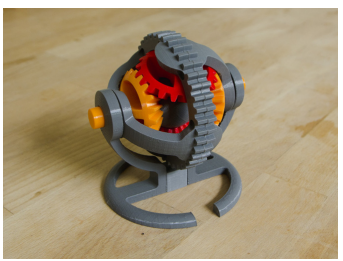
Common vision of AI



Your vision after this course

The AI allure

Common vision of AI



Your vision after this course

- No mysterious “intelligence” or “understanding” in machines
- Machines are still **stupid**.
  - They can only minimize functions.
- Why is AI powerful
  - Smart combination of statistics and optimization
  - Evolution of computing architectures

Figures from Pixabay and Savage Rodent.

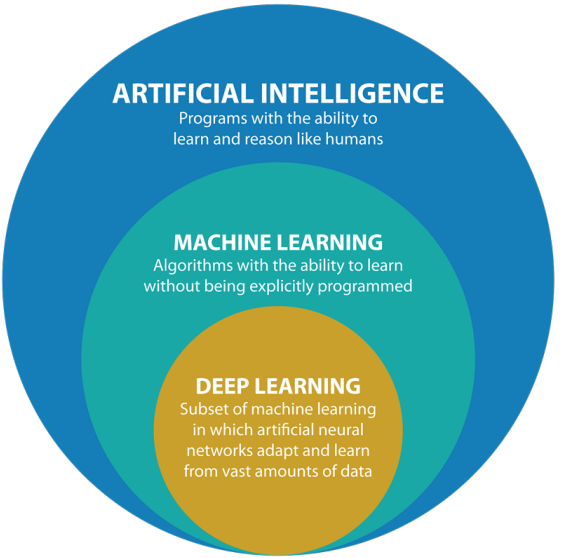
The danger of AI

Blindly trusting Machine "intelligence":

- Financial loss [Fun20]
  - e.g., Knight Capital loss 440 mln $ in 45 min due to an algorithm error [SM18].
- Social inequity [Cou18, CDCs19]
  - e.g., racially discriminatory policing [Cou18].
- Deaths [AVC19]
  - Uber car decided not to stop and killed a woman [Li19]

To avoid this: **UNDERSTANDING**.

Taxonomy of AI



Picture from Victoria Holt's blog

Sometimes is better to avoid deep learning
(see Akamai: predicting users' behavior)

Definitions

The term “Artificial Intelligence” was coined in the Dartmouth workshop.

Claude Shannon (front right), John McCarthy (back right), Marvin Minsky (center), Ray Solomonoff (front left), and Nathaniel Rochester (back left).

Photo credit Margaret Minsky via www.achievement.org. Description from R. Guinness.

AI: algorithms able to take complex decisions or give complex answers in order to maximize a utility function [Bri19, (1)]:

$$f^* = \arg\max_f \mathbb{E}U(f, e)$$

where:

- f: All possible decisions or answers
- e: Observations of the environment

Machine Learning (**ML**)[Bri19, §4.1]: “Algorithms . . . that improve their performance with repeated experience on the task”

Definitions

The term “Artificial Intelligence” was coined in the Dartmouth workshop.

Claude Shannon (front right), John McCarthy (back right), Marvin Minsky (center), Ray Solomonoff (front left), and Nathaniel Rochester (back left).

Photo credit Margaret Minsky via www.achievement.org. Description from R. Guinness.

AI: algorithms able to take complex decisions or give complex answers in order to maximize a utility function [Bri19, (1)]:

$$f^* = \arg\max_f \mathbb{E}U(f, e)$$

where:

- f: All possible decisions or answers
- e: Observations of the environment

Machine Learning (**ML**)[Bri19, §4.1]: “Algorithms . . . that improve their performance with repeated experience on the task”

ML takes advantage of **big data** to achieve AI.

Goals

Theory	**Practice**
Understand ML methods	
	Use ML Python libraries
	Carry on data science project
Carry on data science project	
Robust interpretation of the results	
	Apply ML to practical network problems

ML in Communication Networks

Challenges:

- Networks are extremely complex (many users, protocols, applications, events, . . .)
- 5G calls for flexible and adpative networks (with small human intervention)
- Decisions within milliseconds
- Traffic is encrypted

$\Longrightarrow$ The classic "Model and control" approach is infeasible
$\Longrightarrow$ Need for **data-driven**, approximated and fast decisions

Opportunities:

- **Big data**: a lot of measures from network equipment
- Nobody dies if the network is inefficient

Use cases in Communication Networks

For your project

- Regression:
  - Poqemon dataset: predict user quality of experience in videos
  - Encrypted Traffic: predict flow duration
- Classification:
  - Encrypted Traffic: predict application
  - KDD: classify attacks
  - Industry 4.0: Defect prediction.
    - You can participate to a competition of Collège de France.
    - Register as student of the course "Apprentissage automatique pour les réseaux"
    - Insert a password (ask me)
- Anomaly detection:
  - KDD: find attacks.
  - Industry 4.0

Other in this course:

- Recommendation system for a video service
- Applications to high speed virtual network functions
- Predictive Maintenance
- Find fraudulent credit card transactions
- Find anomalous traffic from traffic matrices

Organization

- One topic per class.
- Groups of 4 students
  - 1 dataset $\iff$ 1 group.
- Cours integré: theory / practice / your presentation
- Final Project:
  - Each group must show regression, classification, anomaly detection.
  - Collect what you did during the module
- Exam
  - All that has been discussed in class can be asked to the exam, including proofs of theorems.
- Mark:
  - 20% Presence and Participation (questions, presentations, forum)
  - 40% Exam (Theory)
  - 40% Project.

Material

- Google is your friend!
- Books
  - Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras and Tensorflow. O'Reilly Media.
    (Accessible from Drawsonera - Mediathèque)
  - James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An introduction to Statistical Learning (Vol. 7). (online version)
- My website
  - Slides
  - Notebooks on colab

What you need

- Account on Google Colab
  - Video tutorial
- Account on Google Drive

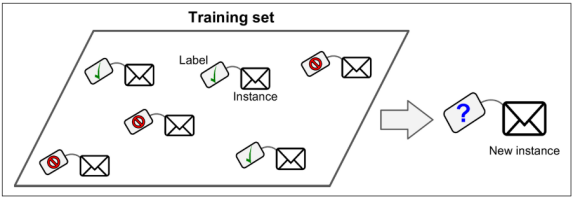
Section 2

Introduction to Machine Learning

Taxonomy of Learning Methods

Supervised L (label)
Unsupervised L (no label, insights from features)
Semi-supervised L
Reinforcement L
Recommender Sys

Pictures from [Gér17].



Classification, e.g. traffic class

Regression, e.g. QoE

Taxonomy of Learning Methods

Supervised L (label)

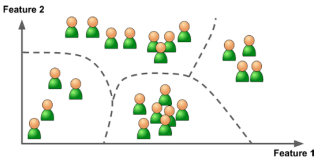
Unsupervised L
(no label, insights from features)

Semi-supervised L

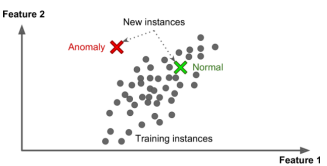
Reinforcement L

Recommender Sys

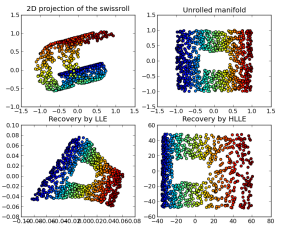
Pictures from [Gér17].



Clustering (e.g. users for targeted ads)



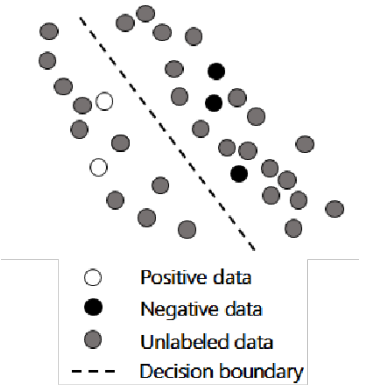
Anomaly Detection (e.g. attack or errors)



Dimensionality reduction
(reduce num of features)

Taxonomy of Learning Methods

Supervised L (label)
Unsupervised L
(no label, insights from features)
Semi-supervised L
Reinforcement L
Recommender Sys

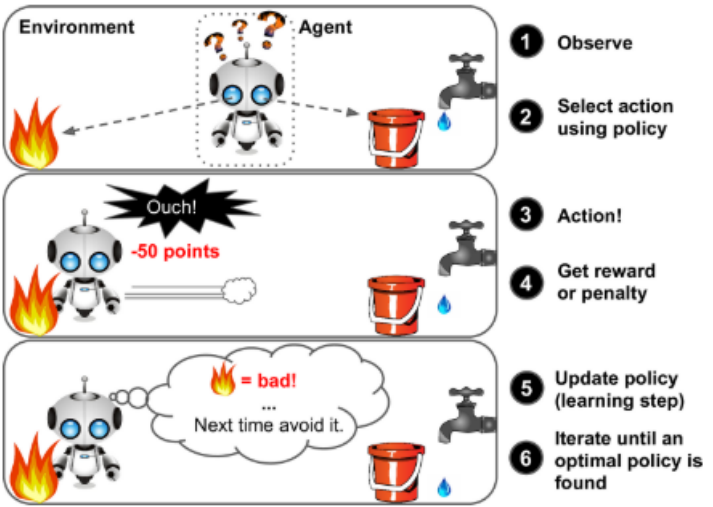


(e.g., only few labeled data: cluster data (unsupervised) and assign to cluster the prevalent label - src: Wikipedia)

Taxonomy of Learning Methods

Supervised L (label)
Unsupervised L
(no label, insights from features)
Semi-supervised L
Reinforcement L
Recommender Sys

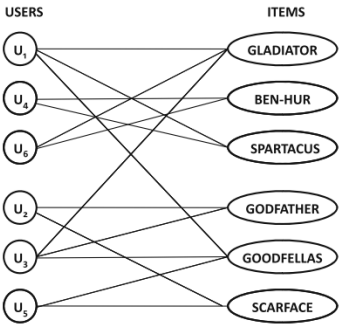
Pictures from [Gér17].



Taxonomy of Learning Methods

Supervised L (label)
Unsupervised L
(no label, insights from features)
Semi-supervised L
Reinforcement L
Recommender Sys

	GLADIATOR	GODFATHER	BEN-HUR	GOODFELLAS	SCARFACE	SPARTACUS
U_1	1			5		2
U_2		5			4	
U_3	5	3		1		
U_4			3			4
U_5				3	5	
U_6	5		4			



Source: [Agg16].

Combination of Unsupervised L (Dimensionality reduction) and Semi-supervised L.

The methods you will learn

	ML task	Linear Regression	Logistic Regression	Tree-based learning	Neural Networks	k-Neirest Neighbors
Supervised	Regression	x			x	
	Classification		x	x	x	
Unsupervised	Clustering				x	x
	Dimensionality reduction				x	
	Anomaly detection			x	x	x
	Recommender Systems				x	

Table: Note: this table is not exhaustive and just summarizes our module.

Common concepts

BufferHealth	BufferProgress	BufferValid	label	label_num
10.241165	0.015357	true	q360p	360
4.446780	0.007103	true	q144p	144
3.989780	0.006509	true	q144p	144
3.700462	0.005897	true	q360p	360
4.512780	0.007156	true	q360p	360
9.454706	0.016805	true	q360p	360
4.606780	0.008046	true	q144p	144
5.301853	0.007990	true	q720p	720
3.638107	0.005493	true	q240p	240
5.314732	0.00[illegible]	true	q240p	[illegible]
8.554780	0.01[illegible]	true	q480p	[illegible]
4.189780	0.007516	true	q360p	360
3.633641	0.005897	true	q480p	480
1.495841	0.002473	true	q720p	720
8.802211	0.014076	true	q1080p	1080
4.611142	0.009263	true	q144p	144
5.590378	0.009113	true	q480p	480
4.940168	0.008851	true	q1080p	1080
4.940168	0.008851	true	q1080p	1080
9.239532	0.016335	true	q720p	720

Correct **“protocol”**:

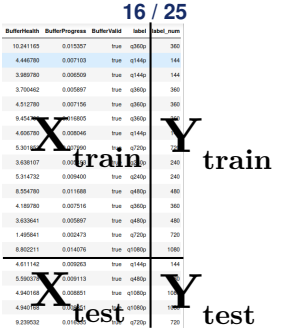
- Train / test

$\mathbf{X}$ $\mathbf{y}$

Common concepts

Correct **“protocol”**:

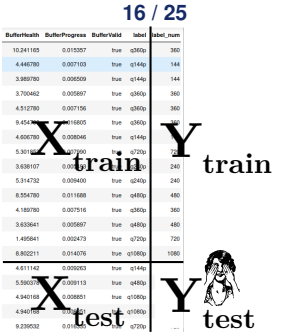
- Train / test



Common concepts

Correct **“protocol”**:

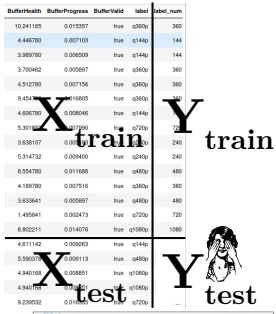
- Train / test



Common concepts

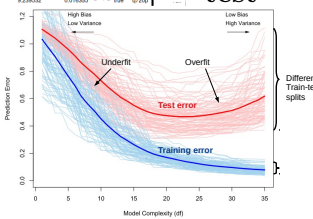
Correct **“protocol”**:

- Train / test



Parsimony:

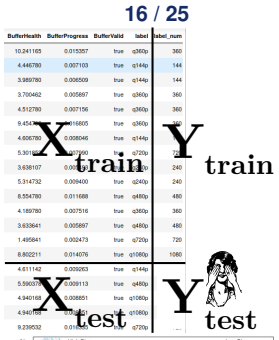
- Many parameters $\Longrightarrow$ Complex Model $\Longrightarrow$ Overfit



Common concepts

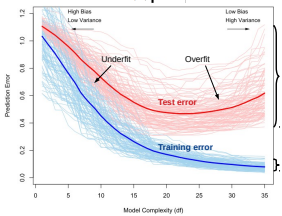
Correct **“protocol”**:

- Train / test



Parsimony:

- Many parameters $\Longrightarrow$ Complex Model $\Longrightarrow$ Overfit



Randomness

- Why: it avoids overfitting.

Notation

Symbol type	Notation
Scalar	a
Vector	$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$
Vector transpose	$\mathbf{a}^T = (a_1, \ldots, a_n)$
Matrix	$\mathbf{U} = \begin{pmatrix} u_{1,1} & \ldots & u_{1,n} \\ \vdots & \ddots & \vdots \\ u_{m,1} & \ldots & u_{m,n} \end{pmatrix}$
Set	$\mathscr{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots)$

We will try to follow the notation of [Ger19] (**Available online**).

- Pre-processing
- Data exploration
- Model selection
- Performance evaluation

Section 3

Data exploration

Statistics

Before performing any supervised or unsupervised learning task, it is better to explore the dataset.
Exercise: Write the definition or the formula of the following quantities (from the notebook) here:

- Mean, Variance, Standard Deviation
- Percentiles, Median
- Boxplot
  (Write just what are the lines that compose and how you can compute them)
- Histogram
- Covariance
- Pearson's correlation coefficient

Note: The formulas of the quantities above may be asked at the exam.

Let's code ...

Go to notebook 01.exploration

Recap

In this lesson:

- AI and ML
- Applications to Communication Networks
- Taxonomy of ML
- Google Colab
- Data exploration
- Basic statistics
- Pre-processing
  - Logarithmic Transformation
  - Missing values
  - One-Hot Encoding

In next lesson:

- Supervised Learning:
  - Linear Regression
  - Polynomial Regression

To know more

- Video: Intelligence artificielle en 1h [Par]

References I

- Charu C. Aggarwal, *Recommender Systems*, 2016.
- *Uber crash shows complexities of training self-driving vehicles*, 2019.
- Naveen Sundar Bringsjord, Selmer and Govindarajulu, *Artificial Intelligence*, Stanford Ecnyclopedia of Philosophy (2019).
- Corinne Cattekwaad, Roel Dobbe, and Corinne Cath-speth, *Politicians and Administrators : Don't expect miracles from Artificial Intelligence*, https://www.oii.ox.ac.uk/blog/politicians-and-administrators-dont-expect-miracles-from-artificial-intelligence/, 2019.
- Rachel Courtland, *The bias detectives*, Nature News (2018).
- Jeffery Funk, *AI: Expect Evolution, Not Revolution*, IEEE Spectrum (2020), no. 3.

References II

- Aurélien Géron, *Hands-on machine learning with scikit-learn and tensorflow*, 2017.
- Aurelien Geron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly, 2019.
- Michael Li, *Another Self-Driving Car Accident, Another AI Development Lesson*, 2019.
- Jean-Luc Parouty, *Deep Learning : Je t'aime moi non plus..*, https://replay.jres.org/videos/watch/ab4a07f4-d209-4427-88a2-78231c341464.
- Christos Saltapidas and Ramin Maghsood, *Financial Risk: The fall of Knight Capital Group*, Tech. report, Chalmers University, 2018.