

Machine Learning for Networks: Introduction -

Andrea Araldo

May 13, 2024



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

Section 1

Presentation

The allure of Artificial Intelligence (AI)

3 / 25

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

...

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

Accenture [website](#).



A consultant (or a researcher) today.

The AI allure

4 / 25



Common vision of AI



Common vision of AI



Your vision after this course



Common vision of AI



Your vision after this course

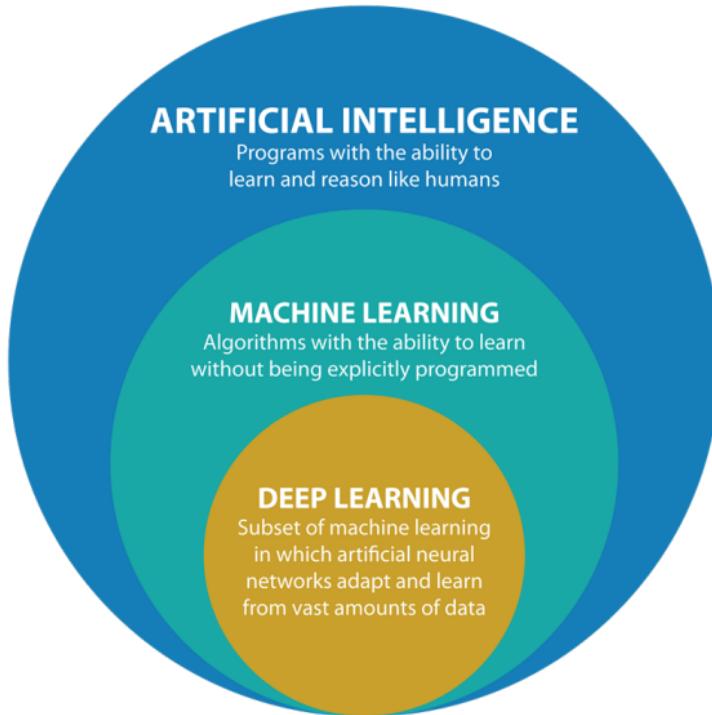
- No mysterious “intelligence” or “understanding” in machines
- Machines are still **stupid**.
 - They can only minimize functions.
- Why is AI powerful
 - Smart combination of statistics and optimization
 - Evolution of computing architectures

Figures from Pixabay and [Savage Rodent](#).

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.**



Sometimes it's better to avoid deep learning
(see [Akamai: predicting users' behavior](#))

Picture from [Victoria Holt's blog](#)

Definitions

7 / 25

The term “Artificial Intelligence” was coined in the Dartmouth workshop (1957).



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

7 / 25

The term “Artificial Intelligence” was coined in the Dartmouth workshop (1957).



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.

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

Challenges:

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

⇒ The classic “Model and control” approach is infeasible
⇒ 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

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

- One topic per class.
- Groups of 4 students
 - 1 dataset \iff 1 group.
- Cours intégré: 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.

- 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

14 / 25

Supervised L (label)

Unsupervised L

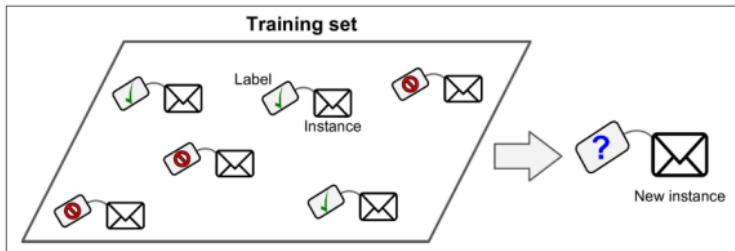
(no label, insights
from features)

Semi-supervised L

Reinforcement L

Recommender Sys

Pictures from [Gér17].



Classification, e.g. traffic class



Regression, e.g. QoE

Taxonomy of Learning Methods

14 / 25

Supervised L (label)

Unsupervised L

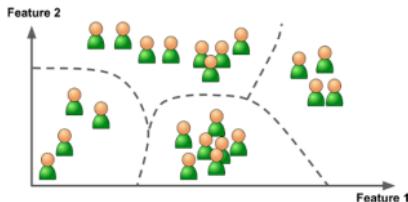
(no label, insights
from features)

Semi-supervised L

Reinforcement L

Recommender Sys

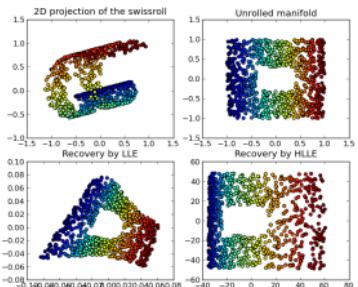
Pictures from [Gér17].



Clustering (e.g. users for targeted ads)



Anomaly Detection (e.g. attack or errors)



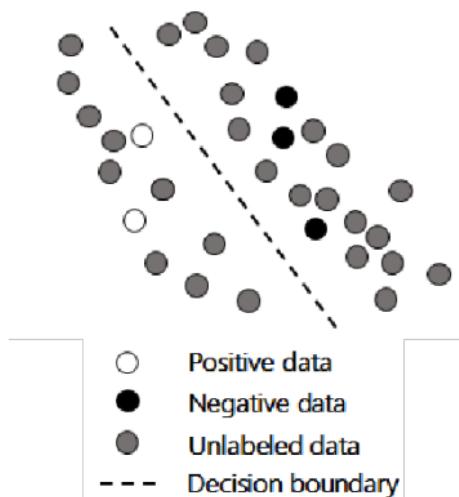
Dimensionality reduction
(reduce num of features)

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 clusters the prevalent label - src: [Wikipedia](#))

Taxonomy of Learning Methods

14 / 25

Supervised L (label)

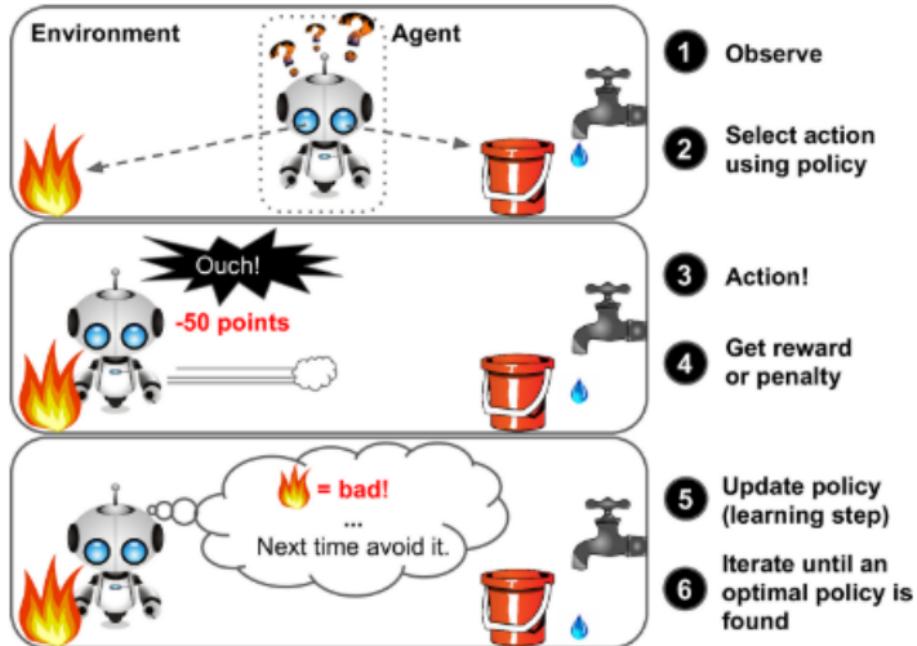
Unsupervised L
(no label, insights
from features)

Semi-supervised L

Reinforcement L

Recommender Sys

Pictures from [Gér17].

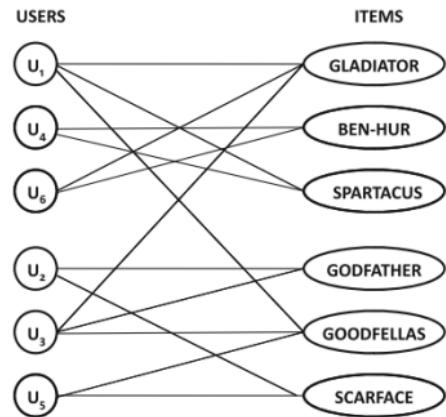


Taxonomy of Learning Methods

14 / 25

- 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

15 / 25

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

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

Common concepts

16 / 25

Correct “protocol”:

- Train / test

BufferHealth	BufferProgress	BufferValid	label	label_num
10.241165	0.015357	true	q340p	360
4.446780	0.007103	true	q144p	144
3.989780	0.006537	true	q144p	144
3.700462	0.005697	true	q360p	360
4.512780	0.007156	true	q360p	360
9.454708	0.016805	true	q360p	360
4.806780	0.009646	true	q144p	144
5.301853	0.007990	true	q720p	720
3.638107	0.005493	true	q240p	240
5.314732	0.010000	true	q144p	144
8.554780	0.010000	true	q480p	480
4.189780	0.007576	true	q360p	360
3.633641	0.005897	true	q480p	480
1.495941	0.002473	true	q720p	720
8.802211	0.014076	true	q1440p	1080
4.611142	0.009260	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

X y

Common concepts

16 / 25

Correct “protocol”:

- Train / test

BufferHealth	BufferProgress	BufferValid	label	label_num
10.241163	0.015307	true	q160p	360
4.446780	0.007103	true	q144p	144
3.989780	0.006509	true	q144p	144
3.700462	0.005897	true	q160p	360
4.512780	0.007106	true	q160p	360
9.451290	0.016605	true	q160p	360
4.656780	0.008004	true	q144p	144
5.301157	0.010790	true	q720p	720
3.638107	0.006509	true	q160p	360
5.314732	0.009400	true	q240p	240
8.554790	0.011688	true	q160p	360
4.189780	0.007516	true	q160p	360
3.633641	0.005897	true	q480p	480
1.495841	0.002473	true	q720p	720
8.802211	0.014007	true	q160p	360
4.611142	0.009203	true	q144p	144
5.596071	0.009913	true	q480p	480
4.940166	0.008851	true	q160p	360
4.940166	0.007521	true	q160p	360
9.239532	0.016550	true	q720p	720

X train Y train

X test Y test

Common concepts

16 / 25

Correct “protocol”:

- Train / test

BufferHealth	BufferProgress	BufferValid	label	label_num
10.241183	0.015357	true	q160p	360
4.446780	0.007103	true	q144p	144
3.985780	0.006509	true	q144p	144
3.700462	0.005697	true	q160p	360
4.912780	0.007156	true	q160p	360
9.454210	0.011605	true	q160p	360
4.696780	0.008044	true	q144p	144
5.301910	0.007290	false	q120p	720
3.636107	0.006250	true	q120p	240
5.314732	0.009400	true	q240p	240
8.554780	0.011668	true	q160p	480
4.180780	0.007519	true	q160p	360
3.633641	0.005697	true	q160p	480
1.495641	0.002473	true	q720p	720
8.802211	0.014076	true	q160p	1080
4.611142	0.009283	true	q144p	144
5.590370	0.009113	true	q160p	360
4.940168	0.008851	true	q160p	360
4.941780	0.008551	false	q160p	360
9.239532	0.016825	false	q720p	720

X train Y train

X test Y test



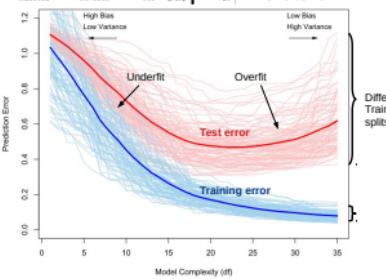
Common concepts

16 / 25

Correct “protocol”:

- Train / test

BufferHealth	BufferProgress	BufferValid	label	label_num
10.241163	0.015357	true	q160p	360
4.446780	0.007103	true	q144p	144
3.985790	0.006509	true	q144p	144
3.700462	0.005697	true	q160p	360
4.912780	0.007156	true	q160p	360
9.454212	0.011605	true	q160p	360
4.696780	0.008050	true	q144p	144
5.301912	0.007290	false	q120p	72
3.638107	0.006250	false	q120p	240
5.314732	0.009400	true	q240p	240
8.554780	0.011668	true	q160p	480
4.180780	0.007575	true	q160p	360
3.633641	0.005697	true	q160p	480
1.495641	0.002473	true	q720p	720
8.802211	0.014076	true	q160p	1080
4.611142	0.009283	true	q144p	144
5.590370	0.009113	true	q160p	360
4.940165	0.008851	true	q160p	360
4.941780	0.008851	false	q160p	360
9.239532	0.016105	false	q720p	720



Parsimony:

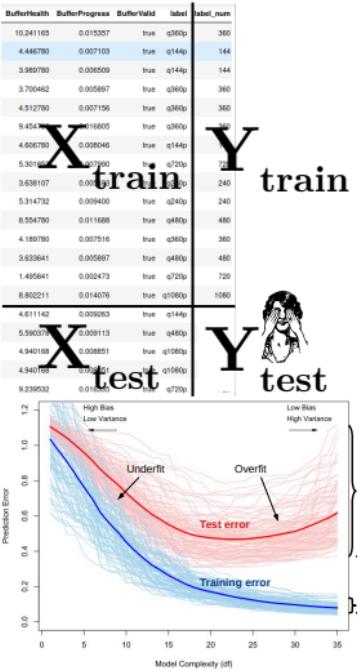
- Many parameters \Rightarrow Complex Model \Rightarrow Overfit

Common concepts

16 / 25

Correct “protocol”:

- Train / test



Parsimony:

- Many parameters \Rightarrow Complex Model \Rightarrow Overfit

Randomness

- Why: it avoids overfitting.

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

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

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

Section 3

Data exploration

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.



Go to notebook 01.exploration

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

- Video: Intelligence artificielle en 1h [[Par](#)]

- [Agg16] Charu C. Aggarwal, *Recommender Systems*, 2016.
- [AVC19] *Uber crash shows complexities of training self-driving vehicles*, 2019.
- [Bri19] Naveen Sundar Bringsjord, Selmer and Govindarajulu, *Artificial Intelligence*, Stanford Encyclopedia of Philosophy (2019).
- [CDCs19] Corinne Cattekwaad, Roel Dobbe, and Corinne Cath-speth, *Politicians and Administrators : Don't expect miracles from Artificial Intelligence*, <https://www.oi.ox.ac.uk/blog/politicians-and-administrators-dont-expect-miracles-from-artificial-intelligence/>, 2019.
- [Cou18] Rachel Courtland, *The bias detectives*, Nature News (2018).
- [Fun20] Jeffery Funk, *AI: Expect Evolution, Not Revolution*, IEEE Spectrum (2020), no. 3.

- [Gér17] Aurélien Géron, *Hands-on machine learning with scikit-learn and tensorflow*, 2017.
- [Ger19] Aurelien Geron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly, 2019.
- [Li19] Michael Li, *Another Self-Driving Car Accident, Another AI Development Lesson*, 2019.
- [Par] Jean-Luc Parouty, *Deep Learning : Je t'aime moi non plus..*,
<https://replay.jres.org/videos/watch/ab4a07f4-d209-4427-88a2-78231c341464>.
- [SM18] Christos Saltapidas and Ramin Maghsoud, *Financial Risk: The fall of Knight Capital Group*, Tech. report, Chalmers University, 2018.