

Machine Learning

Master Degree in Computer Science

Curriculum in Data Science and Engineering

Who we are



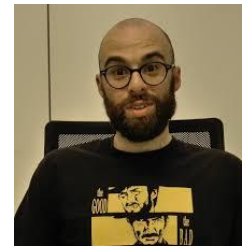
Lorenzo Rosasco (instr)
lorenzo.rosasco@unige.it



Nicoletta Noceti (instr)
nicoletta.noceti@unige.it



Jacopo Dapuesto (TA)
jacopo.dapuesto@edu.unige.it



Paolo Didier Alfano (TA)
paolo.alfano@iit.it

Today

*Course
aims,
logistics
and rules*

*Welcome
to ML*

Course aims, logistics and rules

The role of ML in our study course

Artificial Intelligence Track

Machine Learning (pillar)

*Deep Learning (track specific)
Computational Vision (track specific)*

*Sequential prediction and
Reinforcement Learning (elective)*

Trustworthy AI (elective)

AI and law (elective)

Data Analytics Track

Machine Learning (pillar)

Predictive Analytics Projects (elective)

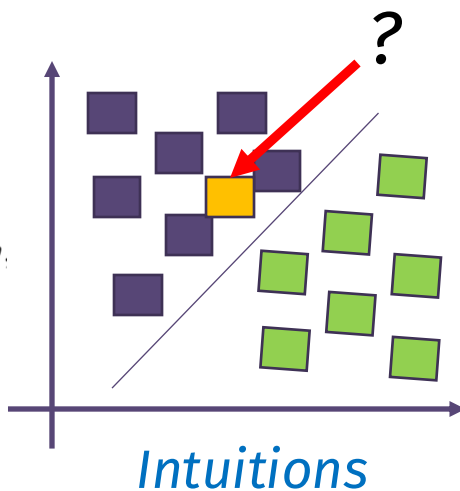
AI and law (elective)

Aims of the course

- This course provides an introduction to the **fundamental methods** at the **core** of modern machine learning
- It covers **theoretical foundations** as well as **essential algorithms** for machine learning complemented by practical lab sessions (with Python and Jupyter Notebook)

Math

$$\min_{w \in \mathbb{R}^D} \frac{1}{n} \sum_{i=1}^n (y_i - w^\top x_i)^2 + \lambda w^\top w$$



Practice



Program

Supervised methods	A bit of theory	Statistical learning Bias and variance and cross-validation
	Local methods	KNNs
	Global methods	Regularized networks Neural networks Convolutional Neural Networks
	Interpretable models	Feature selection Decision Trees
	Clustering	KMeans
Unsupervised methods	Dimensionality reduction	PCA

Classes time and place

- Class times

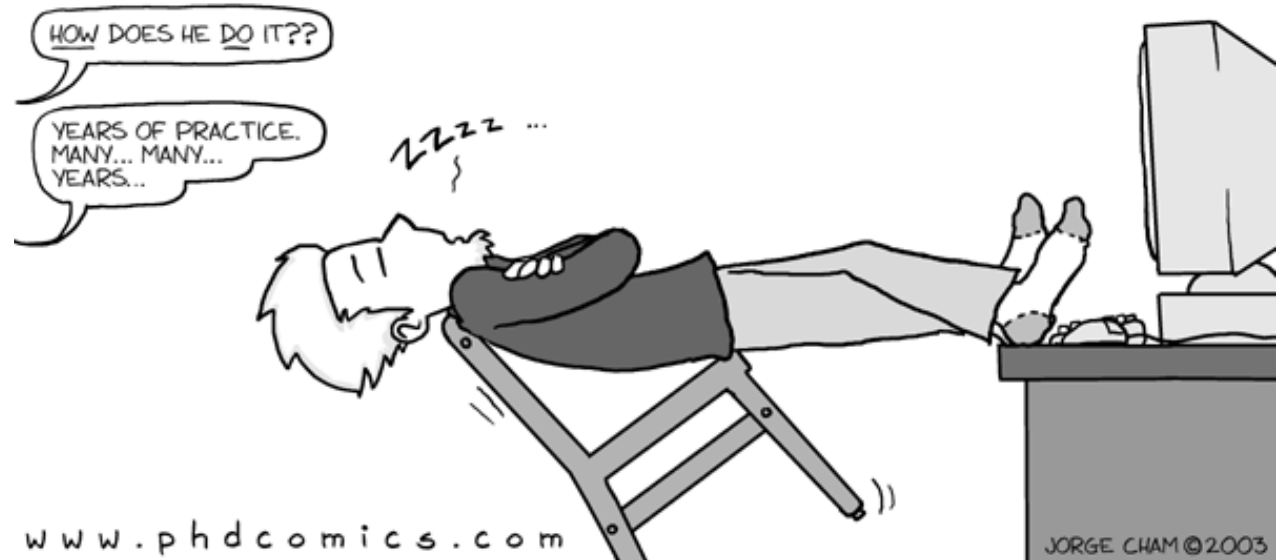
- Monday 9-11 Room 710
- Wednesday 11-13 Room 710
- Friday 14-16 Room 505/SW2

- 72+2 hours (~8 hours for the camp, ~38 hours of theory and algorithms, ~18 hours of lab activities, ~10 hours of **other activities**)

- *Mid-terms*
- *An example of written exam*

Rules of the game

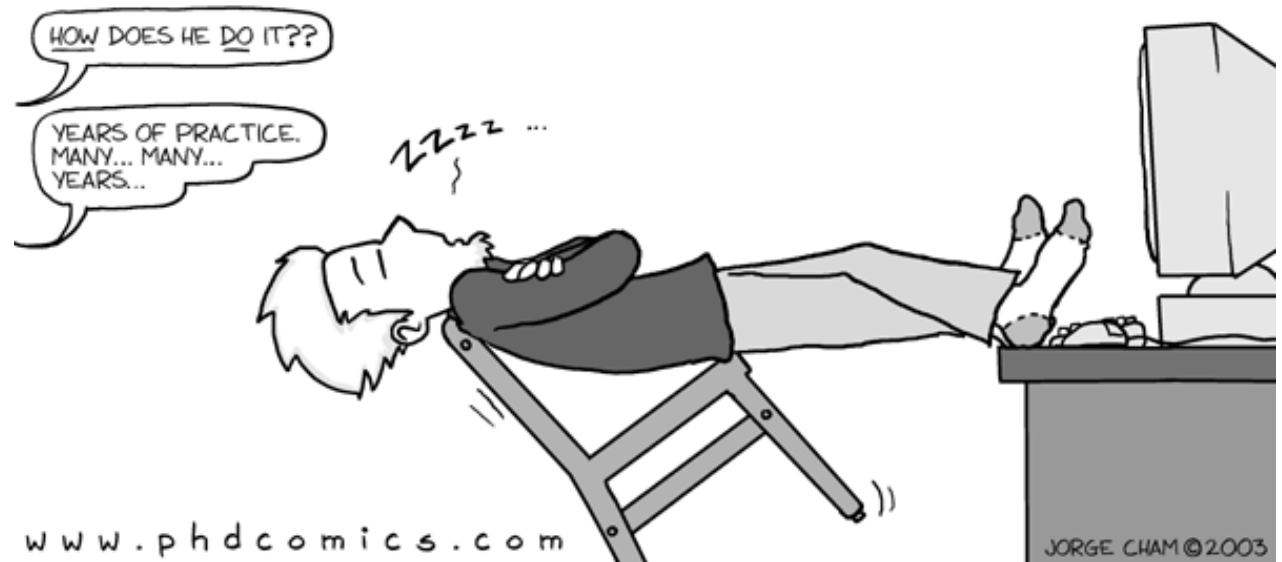
- Hands-on activities → Optional submission few days after the class
- A mid-term in two parts → «More structured» hands-on activities
- Quiz → A «filter» in view of the final exam
- Written exam → 5 exercises/ questions / coding



Rules of the game

*If you submit them all you gain
a +1 on the final grade →
Participation point*

- Hands-on activities → Optional submission few days after the class
- A mid-term in two parts → «More structured» hands-on activities
- Quiz → A «filter» in view of the final exam
- Written exam → 5 exercises/ questions / coding

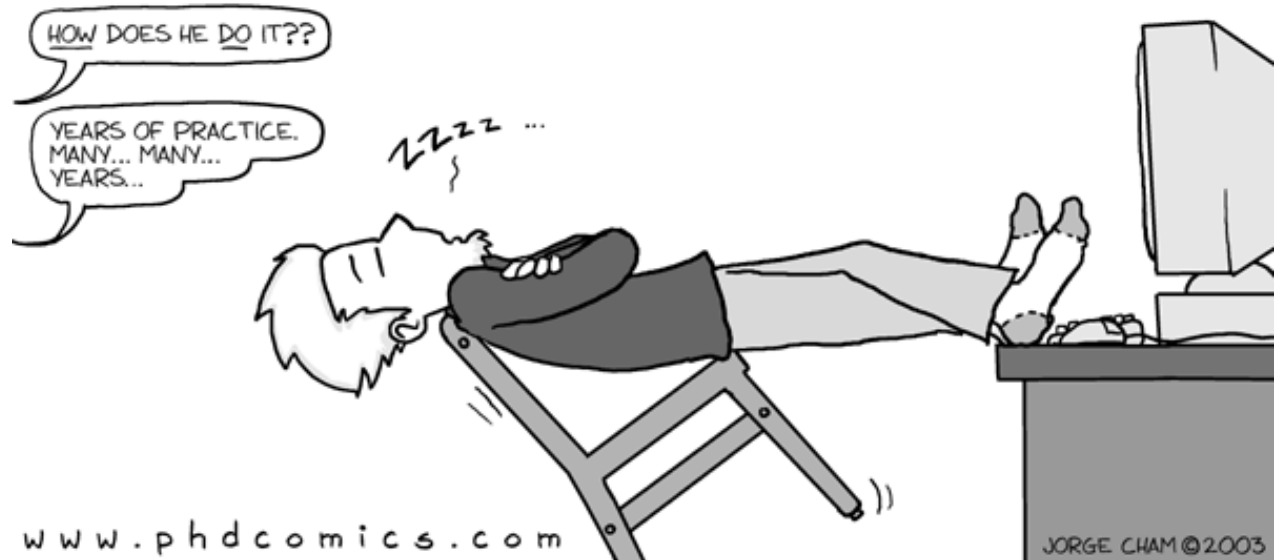


Rules of the game

- Hands-on activities → Optional submission few days after the class
- A mid-term in two parts → «More structured» hands-on activities
- Quiz → A «filter» in view of the final exam
- Written exam → 5 exercises/ questions / coding

You can work in groups of at most 3 people

If you submit them and the evaluation is good, you will have the opportunity of accessing to a shorter written exam (3 Q instead of 5)

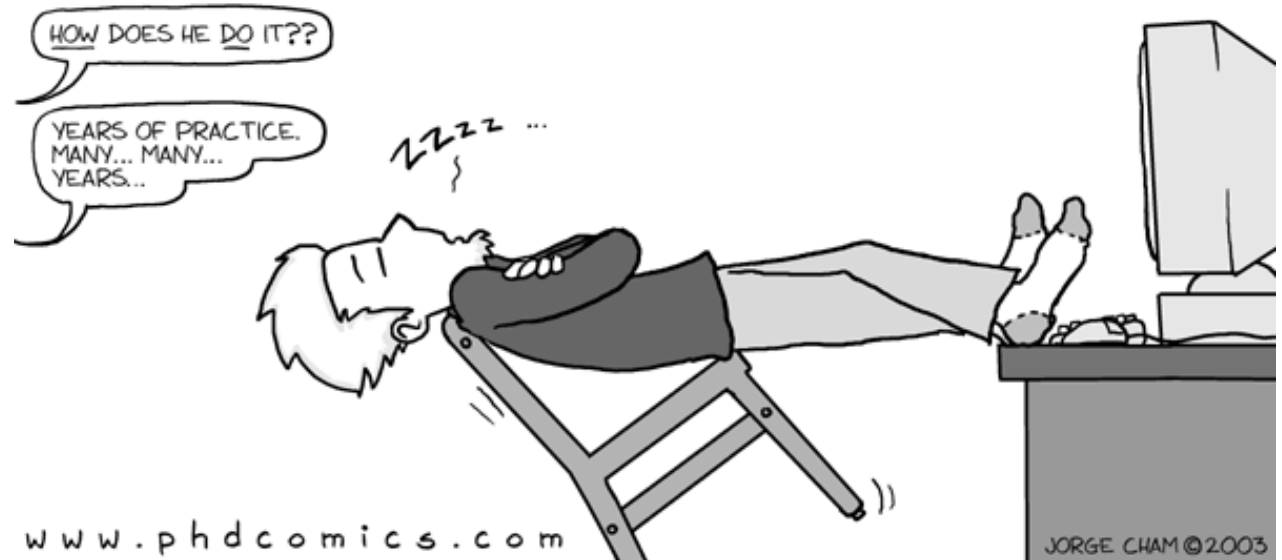


Rules of the game

- Hands-on activities → Optional submission few days after the class
- A mid-term in two parts → «More structured» hands-on activities
- Quiz → A «filter» in view of the final exam
- Written exam → 5 exercises/ questions / coding

You can work in groups of at most 3 people

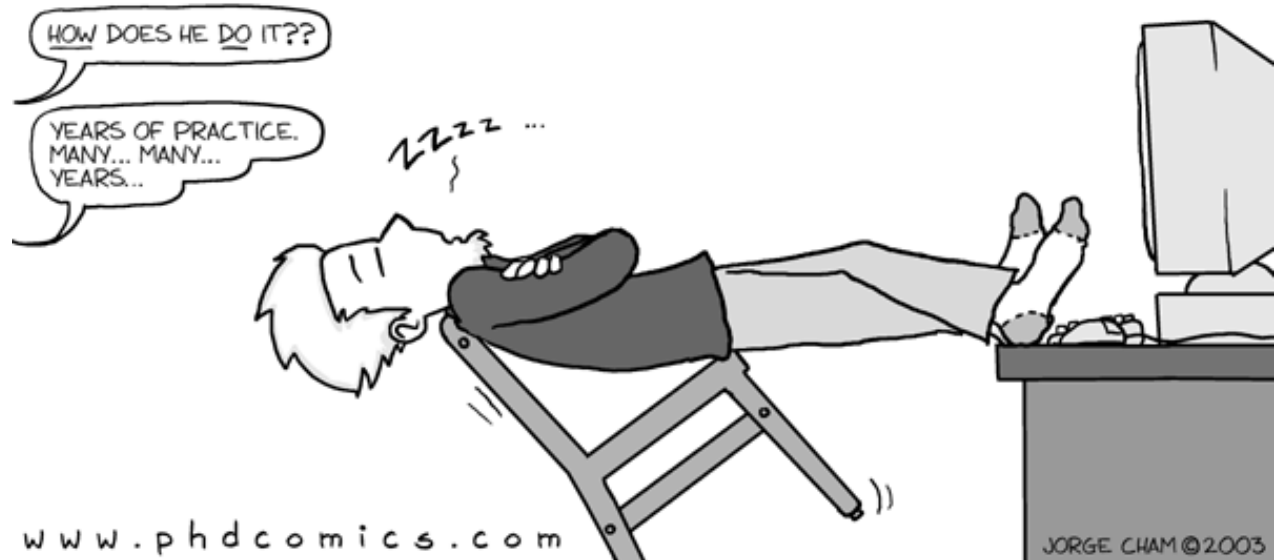
If you submit them and the evaluation is good, you will have the opportunity of accessing to a shorter written exam (3 Q instead of 5)



Rules of the game

- Hands-on activities → Optional submission few days after the class
- A mid-term in two parts → «More structured» hands-on activities
- Quiz → A «filter» in view of the final exam
- Written exam → 5 exercises/ questions / coding

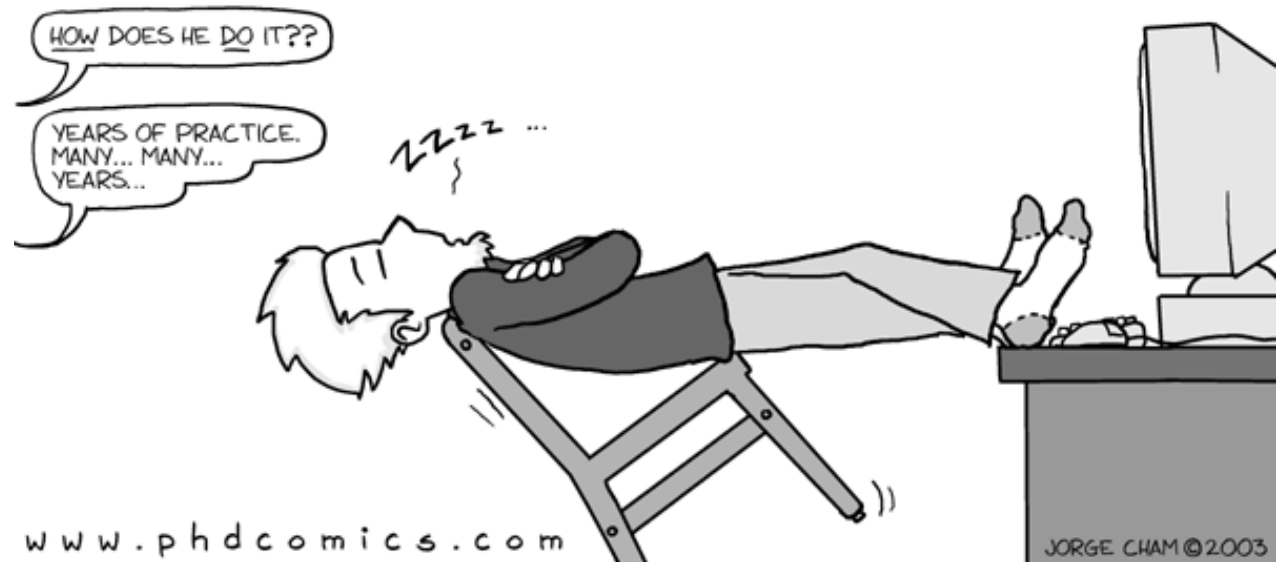
No grade, just passed/not passed



Rules of the game

To summarize, you can attend the course in two ways:

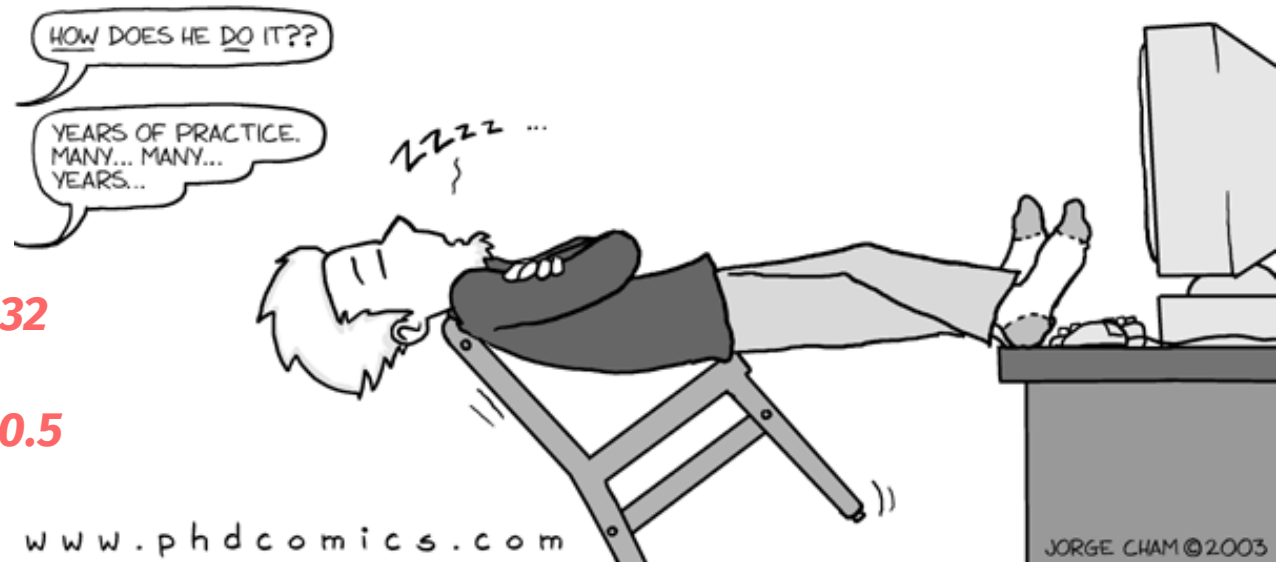
- Actively participating: 1 mid-term in 2 parts + 3 written exercises / questions / coding [+ weekly hands-on submissions]
- 5 written exercises/ questions / coding



Rules of the game

To summarize, you can attend the course in two ways:

- Actively participating: 1 mid-term in 2 parts + 3 written exercises / questions / coding [+ weekly hands-on submissions] **30%** **70%** **[+1]**
- 5 written exercises / questions / coding **100%**



Sum of the points max 32

Laude if final grade > 30.5

Prerequisites

- Basics of programming, linear algebra, calculus, probabilities, and statistics

References

Introductory Machine Learning Notes¹

Lorenzo Rosasco

DIBRIS, Università degli Studi di Genova

LCSL, Massachusetts Institute of Technology and Istituto Italiano di Tecnologia

`lrosasco@mit.edu`

December 21, 2017

Hastie, Tibshirani and Friedman

The Elements of Statistical Learning

- MIT 9.520: Statistical Learning Theory and Applications
- MLCC videos
- Stanford CS229 Machine Learning
- All references reported on our Aulaweb 2022 module **MACHINE LEARNING – 90498**

References

If you have suggestions or find any typos in the notes please fill this form:

<https://goo.gl/forms/OPJ9Ggk1aWYObYCB3>



Introductory Machine Learning Notes¹

Lorenzo Rosasco

DIBRIS, Università degli Studi di Genova

LCSL, Massachusetts Institute of Technology and Istituto Italiano di Tecnologia

lrosasco@mit.edu

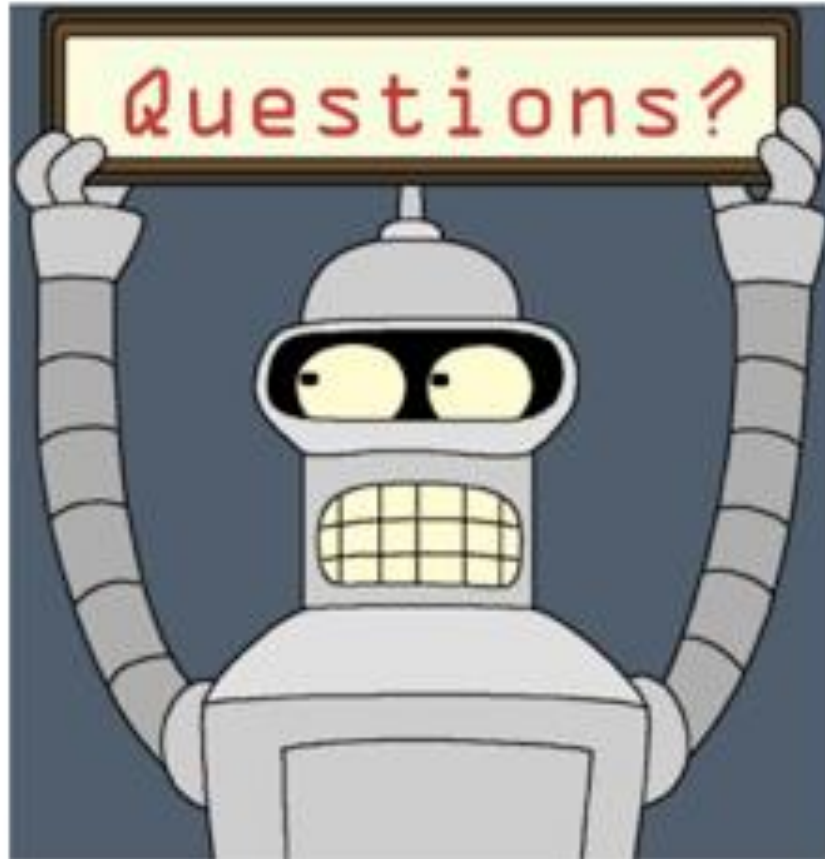
December 21, 2017

Hastie, Tibshirani and Friedman

The Elements of Statistical Learning

- MIT 9.520: Statistical Learning Theory and Applications
- MLCC videos
- Stanford CS229 Machine Learning
- All references reported on our Aulaweb 2024 module **MACHINE LEARNING – 90498**

Good time for questions



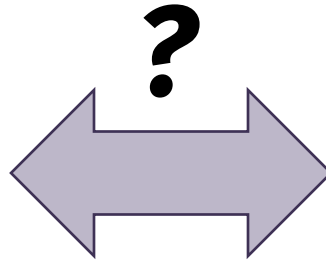


Welcome to Machine Learning

What is (Machine) Learning

Intelligent Systems

Technologically advanced machines that **perceive** and **respond to** the world around them. Intelligent systems can take many forms (automated vacuums, facial recognition programs, personalized shopping suggestions)



Data Science

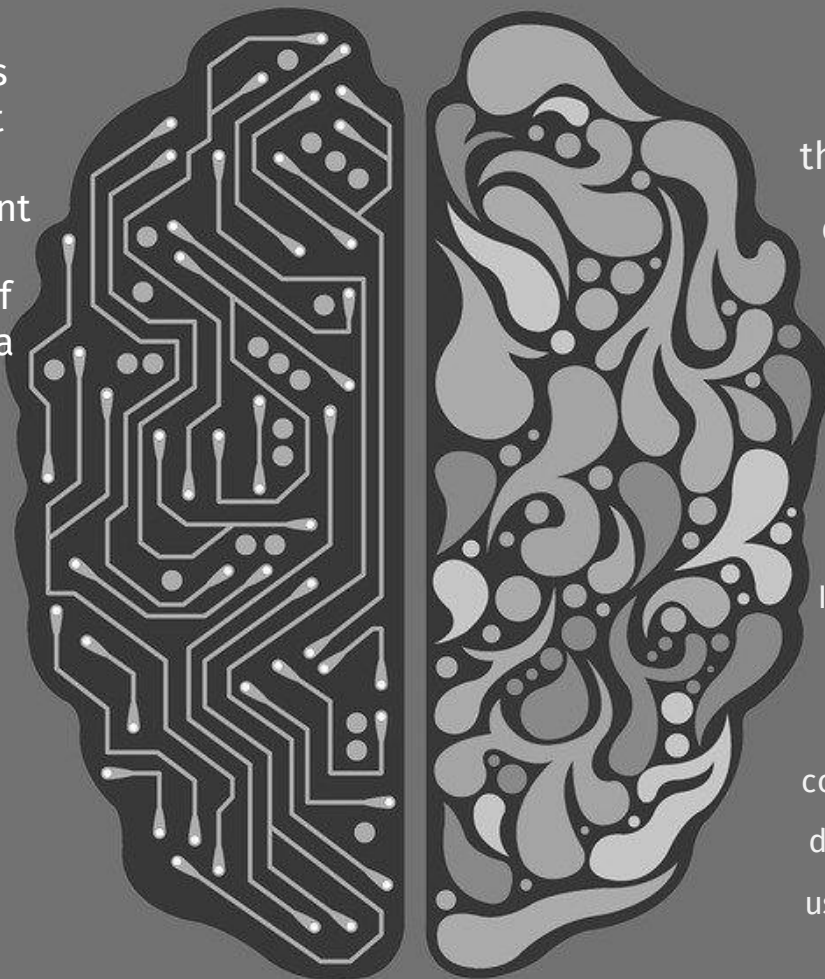
An inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data

BUT ALSO

a concept to unify statistics, data analysis and their related methods in order to understand and analyse actual phenomena with data

Intelligence

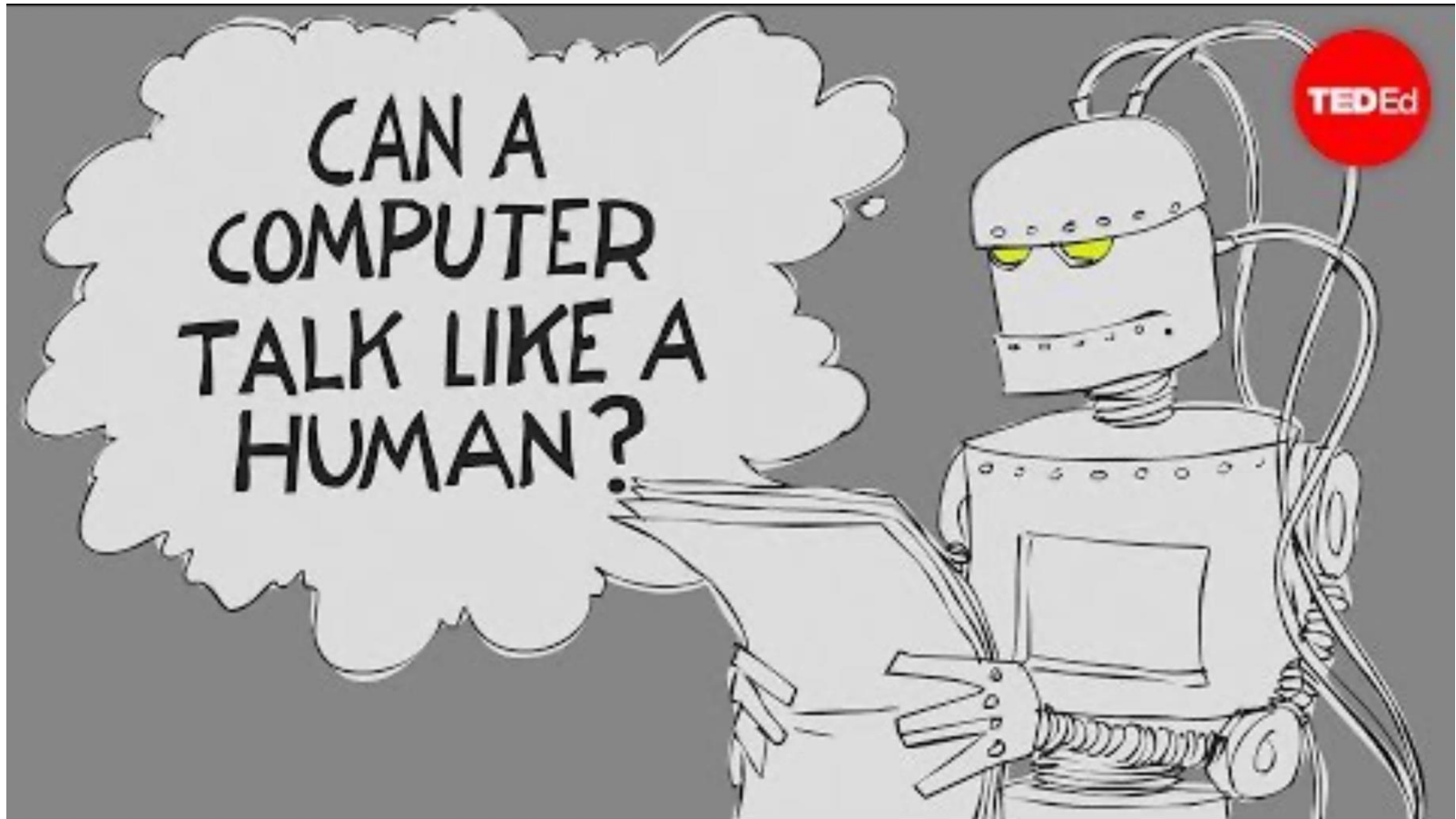
Artificial intelligence is the study of intelligent agents, devices that perceive the environment and takes actions to maximise the chance of successfully achieving a goal



Human intelligence is the intellectual process of humans, relying on complex cognitive feats and high levels of motivation and self-awareness

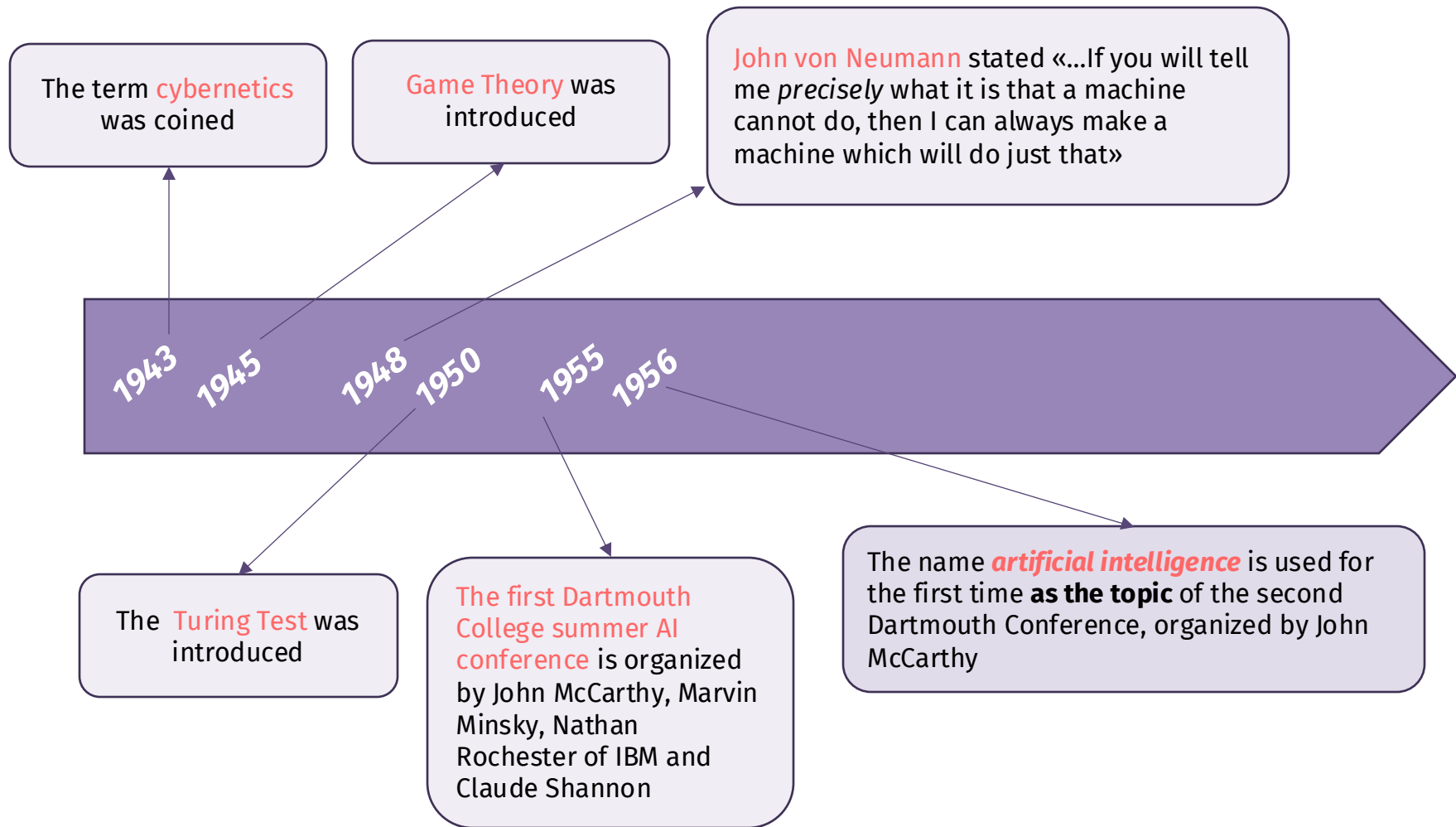
It enables cognitive abilities to learn, form concepts, understand, apply logic, and reason, including the capacities to recognize patterns, comprehend ideas, plan, solve problems, make decisions, retain information, and use language to communicate

The Turing test



<https://www.youtube.com/watch?feature=oembed&v=3wLqsRLvV-c>

Birth of a dream



A nice reading: *AS WE MAY THINK* by Vannevar Bush (1945), a **prescient vision of the future in which computers assist humans in many activities**

The birth of a dream... But how did it go?

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the **conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.** An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for **a summer**

• **Dartmouth Summer Research Conference on Artificial Intelligence** organised by John McCarthy and proposed by McCarthy, Marvin Minsky, Nathaniel Rochester and Claude Shannon.

Around the same time in Genova



P.A.P.A.

First optical character
recognition able to learn from
examples (Genova, Istituto di
Fisica, 1961)

Birth of a dream



Web crawlers and other AI-based information extraction programs become essential in widespread use of the World Wide Web.

Birth of a dream



The Deep Blue chess machine (IBM) beats the world chess champion, Garry Kasparov.

Birth of a dream



DARPA introduces the DARPA Grand Challenge requiring competitors to produce **autonomous vehicles** for prize money.

How are we doing now?



How are we doing now?



<https://youtu.be/pW6nZXeWlGM>

Living portraits



Gemini



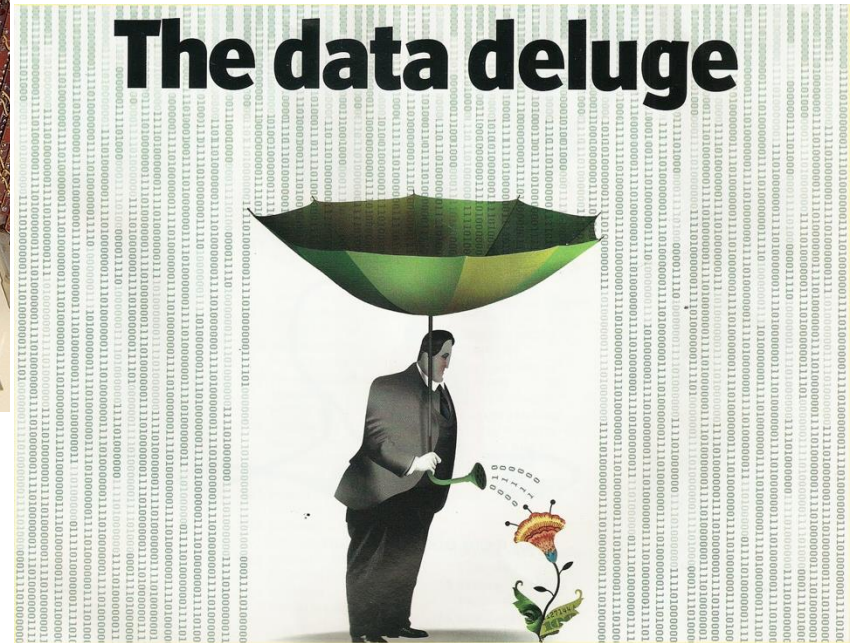
Two Pillars



In August 1961 NASA contracted the MIT Instrumentation Laboratory (later called the Charles Stark Draper Laboratory) to develop the Apollo guidance, navigation and control system. Eldon Hall (shown above) was selected to lead the development team, and astronaut David Scott was chosen as the interface between the designers and the users.

DATA

The data deluge



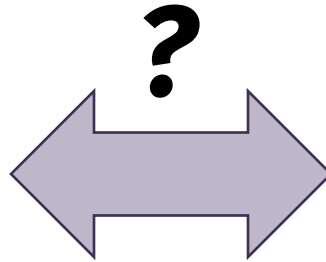
Machine Learning

“It is the field of study that gives computers the ability to learn without being explicitly programmed”

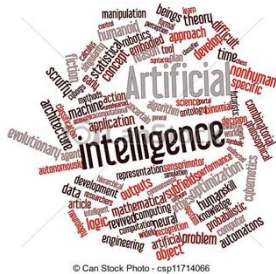
- Arthur L. Samuel, AI pioneer, 1959

What is (Machine) Learning

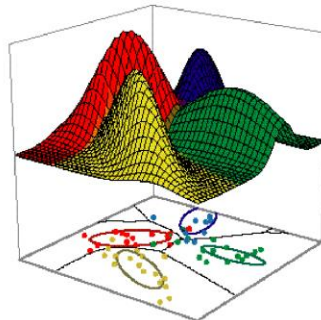
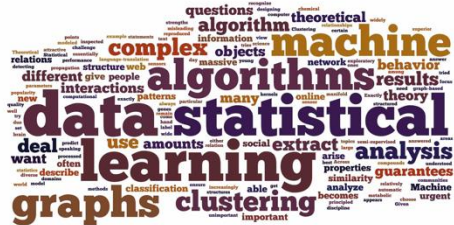
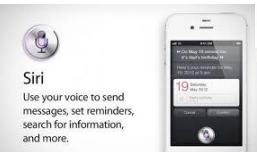
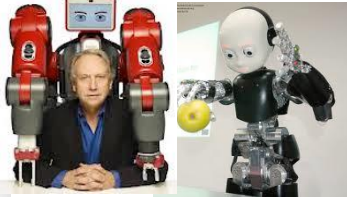
Intelligent Systems



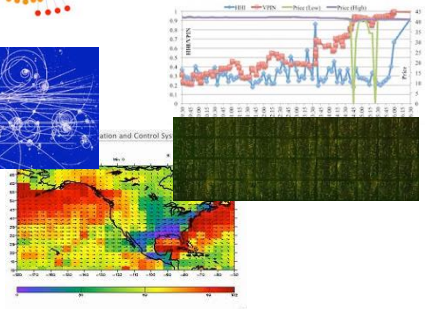
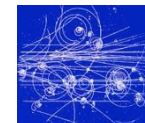
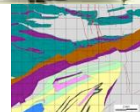
Data Science



© Can Stock Photo - csp11714066



The data deluge



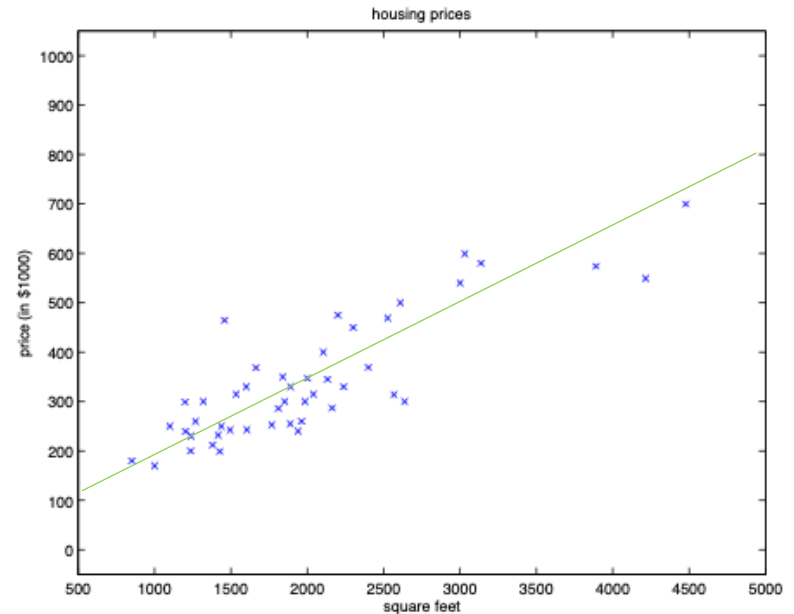
Basic setting: Regression (with an example)

DATA

Living area (feet ²)	Price (1000\$)
2104	400
1600	330
2400	369
1416	232
3000	540
⋮	⋮
2250	?
6000	?

$$(x_1, y_1), \dots, (x_n, y_n)$$

Living area (feet ²)	#bedrooms	Price (1000\$)
2104	3	400
1600	3	330
2400	3	369
1416	2	232
3000	4	540
⋮	⋮	⋮



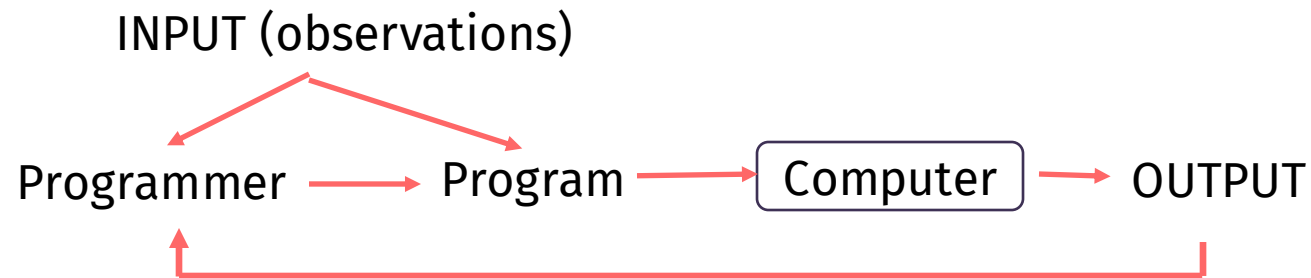
You are looking for a function
such that

$$y_i \simeq f(x_i)$$

And able to generalize to new
data

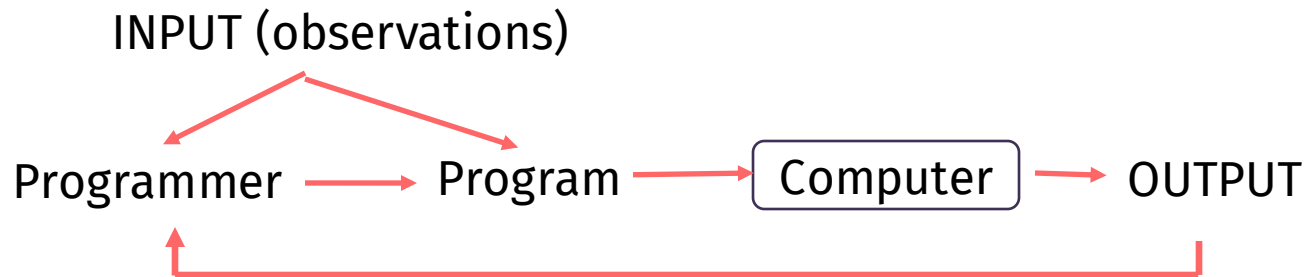
Machine Learning

- Traditional programming paradigm

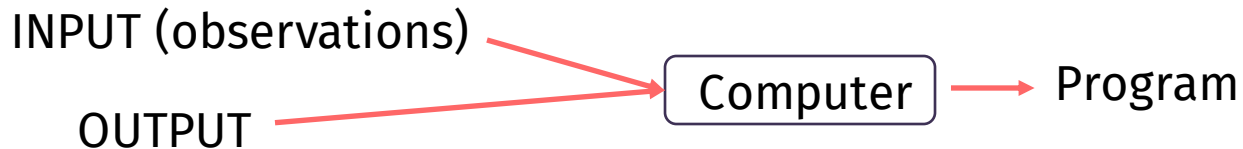


Machine Learning

- Traditional programming paradigm

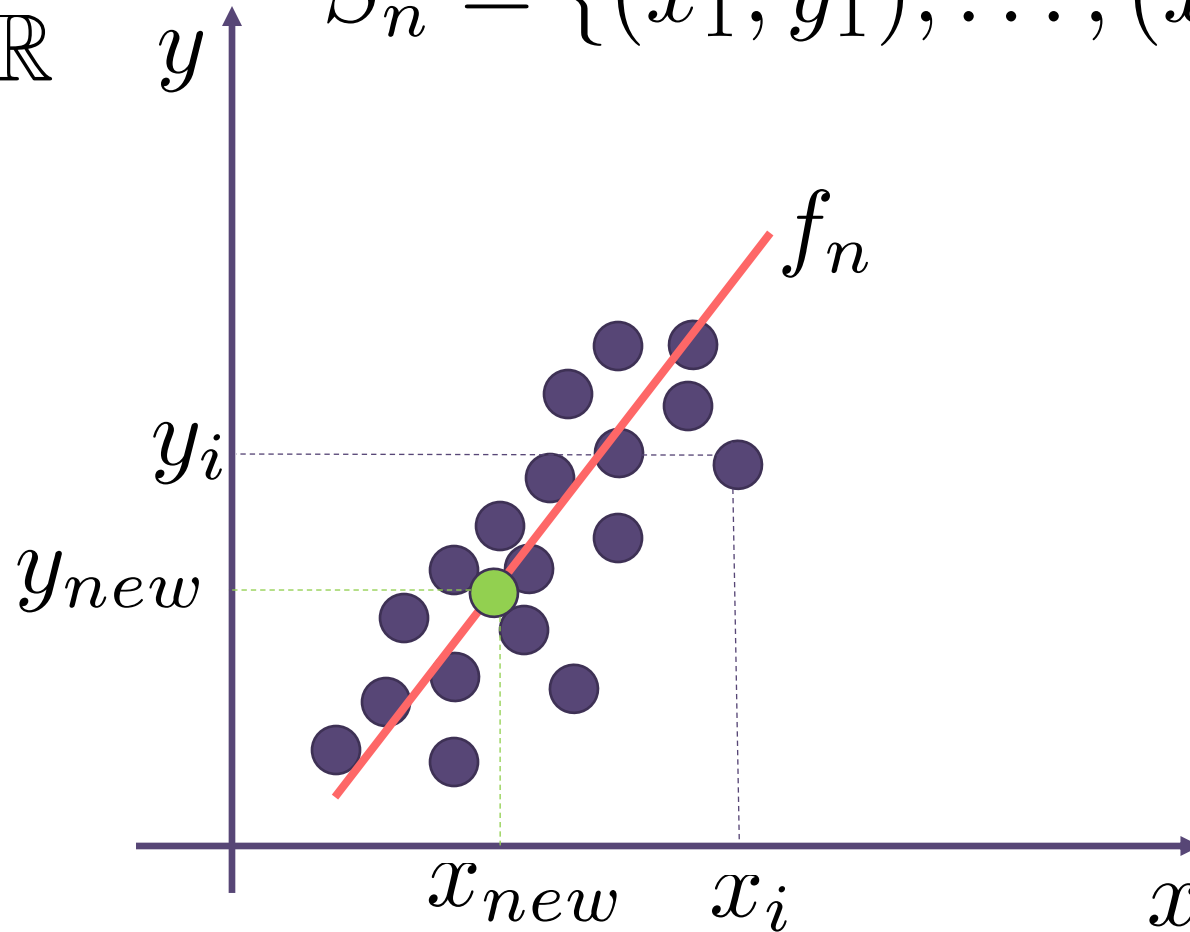


- Machine Learning paradigm



Basic setting: regression

$$Y \subseteq \mathbb{R} \quad S_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$$

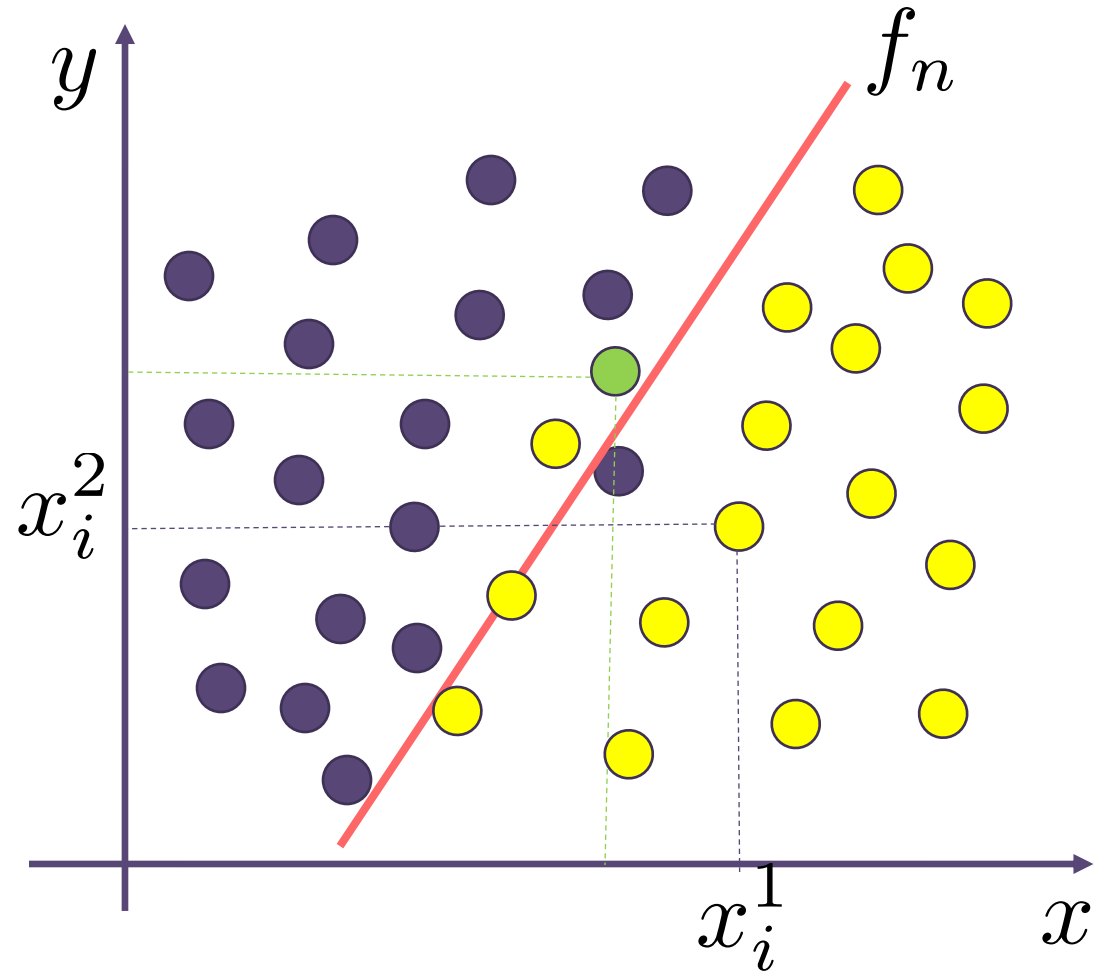


Basic setting: classification

$$Y = \{-1, 1\}$$

$$X \subseteq \mathbb{R}^2$$

$$x_i = [x_i^1, x_i^2]$$



$$S_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$$

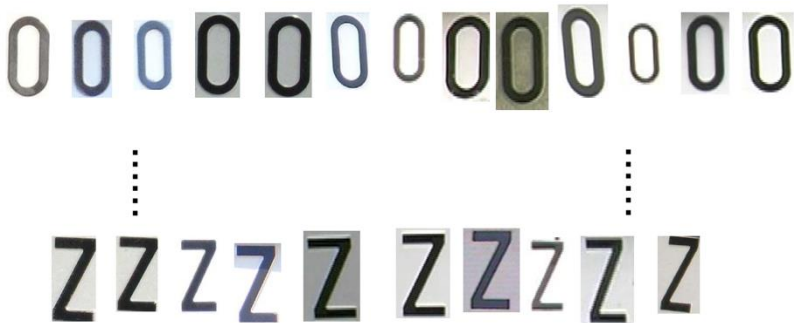
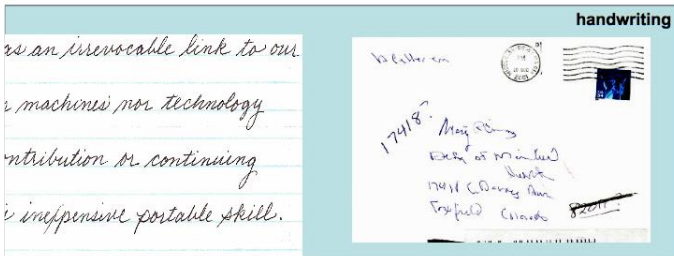
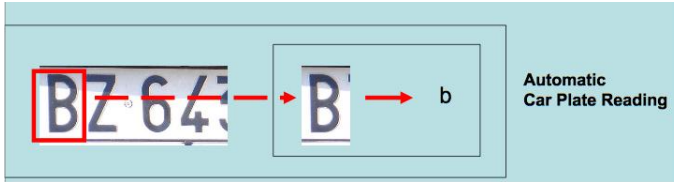
Text Classification



» 2007-05-22 18:33 - News in English		
 Soccer: Milan weighs attack options Gilardino tipped to lead the line, Inzaghi late card	← Sport	(football)
» 2007-05-22 17:29 - News in English		
 Wine wards off senile dementia A glass a day stops mild impairment worsening, Italians say	← Health	
» 2007-05-22 16:27 - News in English		
 Treasury may sell all of Alitalia Formal bids for carrier due by July 2	← Politics	(economy)
» 2007-05-22 15:58 - News in English		
 Afghanistan: Rome cold on Bush call 'We respond to parliament, no one else,' FM says	← Politics	(foreign)
» 2007-05-22 14:11 - News in English		
 Electronic nose sniffs asthma Device developed by Italian researcher in Netherlands	← Technology	

Subject	Date	Time	Body	Spam?
I has the viagra for you	03/12/1992	12:23 pm	Hi! I noticed that you are a software engineer so here's the pleasure you were looking for...	Yes
Important business	05/29/1995	01:24 pm	Give me your account number and you'll be rich. I'm totally serial	Yes
Business Plan	05/23/1996	07:19 pm	As per our conversation, here's the business plan for our new venture Warm regards...	No
Job Opportunity	02/29/1998	08:19 am	Hi II am trying to fill a position for a PHP ...	Yes
[A few thousand rows ommitted]				
Call mom	05/23/2000	02:14 pm	Call mom. She's been trying to reach you for a few days now	No

Image classification



(a) Sky

(b) Trees



(c) Grass

(d) Snow



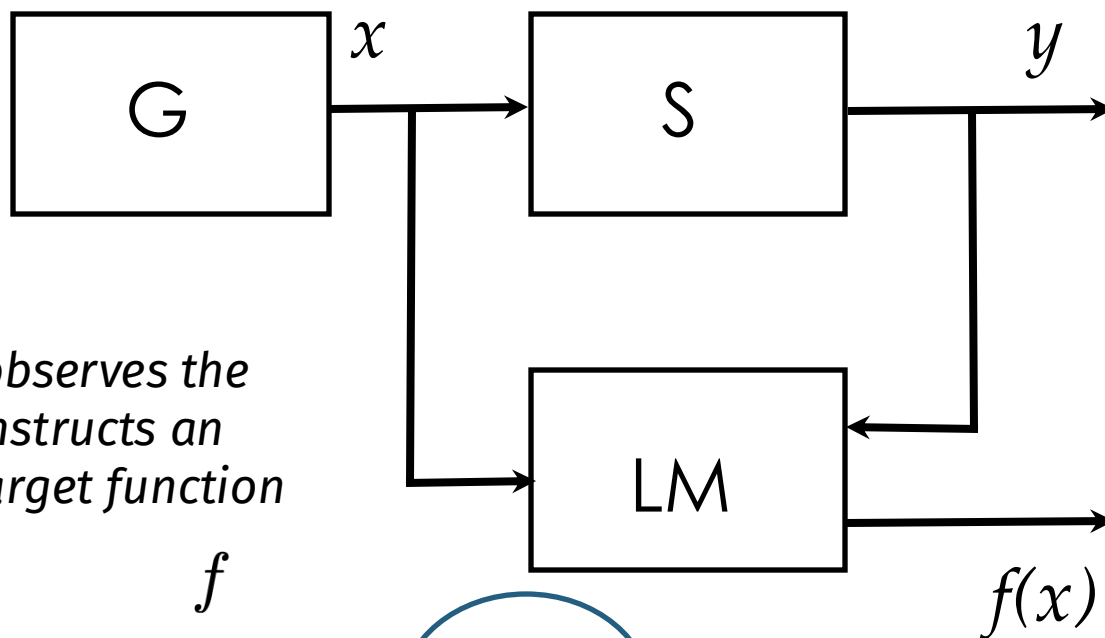
Supervised Learning

$$(x_1, y_1), \dots, (x_n, y_n)$$

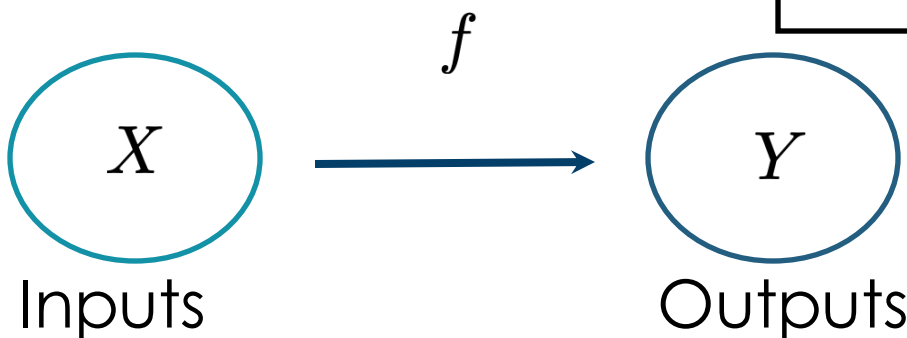
GENERATOR:

generates vectors according to “certain rules” which are **unknown** but **fixed**

SUPERVISOR: transforms the vectors into output values. It is **unknown** but **it exists** and **does not change**



LEARNING MACHINE: it observes the pairs of x and y and constructs an approximation of the target function

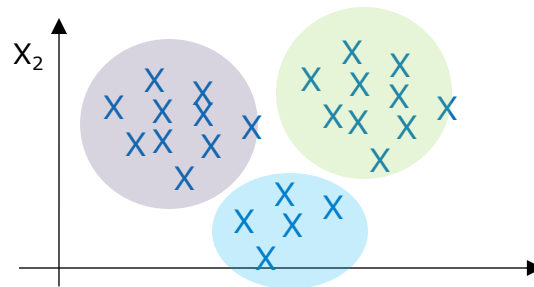


UnSupervised Learning

$$(x_1, \cancel{y_1}), \dots, (x_n, \cancel{y_n})$$

Metaphor: *learning without a teacher*

The goal of unsupervised learning is to directly infer **properties** of the data without the help of a supervisor or teacher providing correct answers or degree-of-error for each observation



A Turing test – Code LYANAZ



A



B

Which one is real?

UniGe

MaL**Ga**