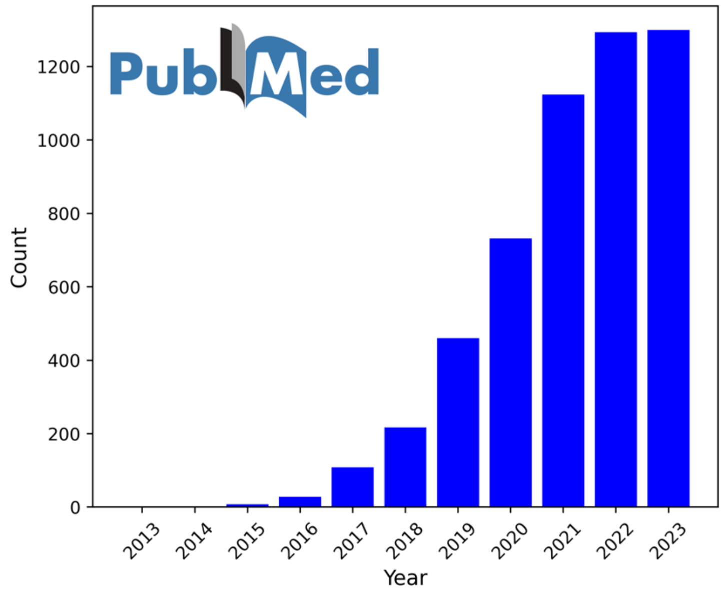
Machine Learning for Medical Images Processing

Michel Dojat

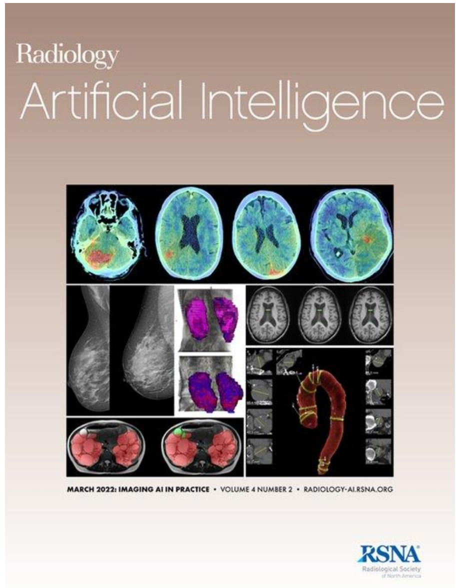
Part I

michel.dojat@inserm.fr

Automatic analysis of medical images

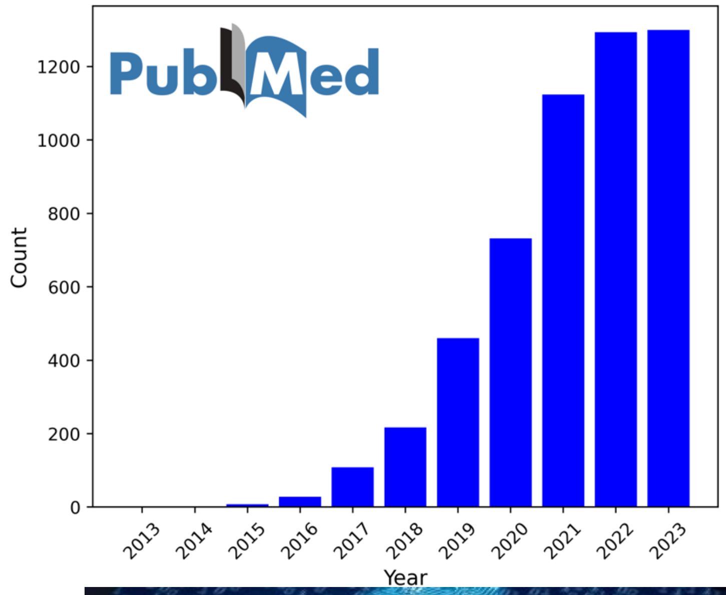


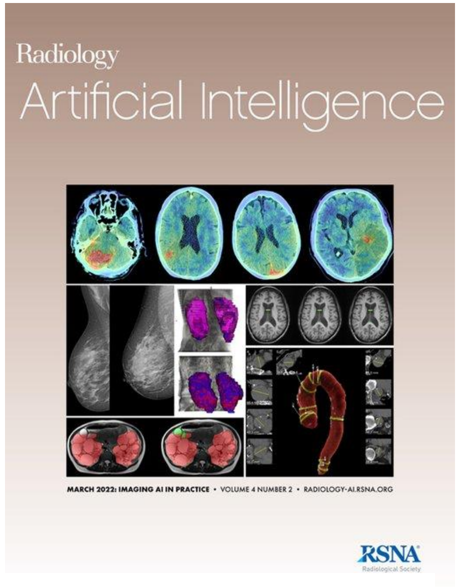
MIDL
Paris 2024

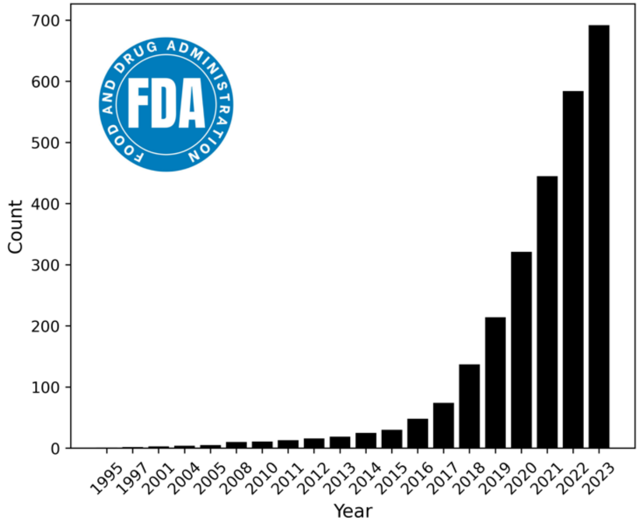


Kw: « Deep Learning » and « Medical Image Analysis »

Automatic analysis of medical images

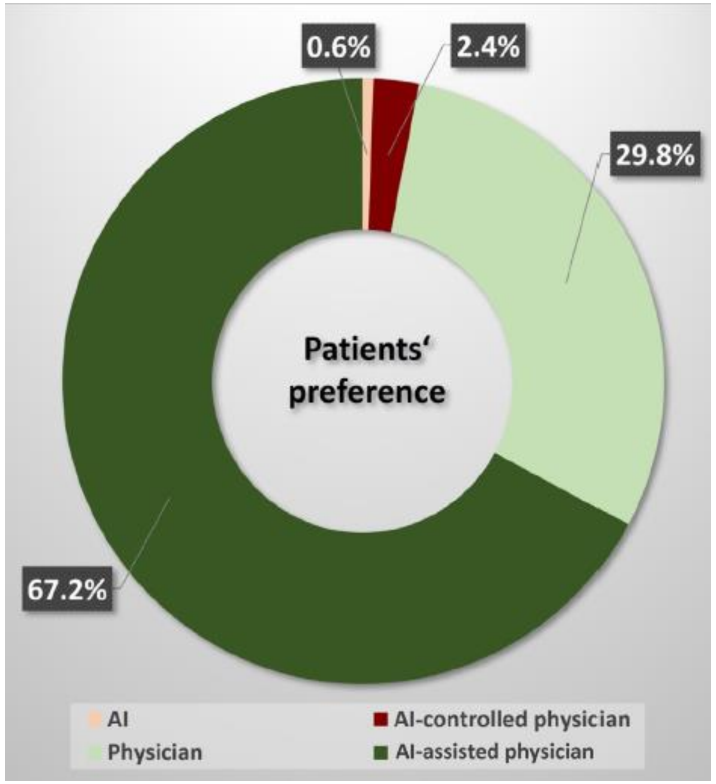






N= 466

Choice of optimal treatment setting between urologists and artificial intelligence (AI).
Prostate cancer patients were asked which type of consultation they would prefer in the current situation before magnetic resonance imaging, biopsy, or radical prostatectomy.



Rodler et al. Eur Urology Focus 2023

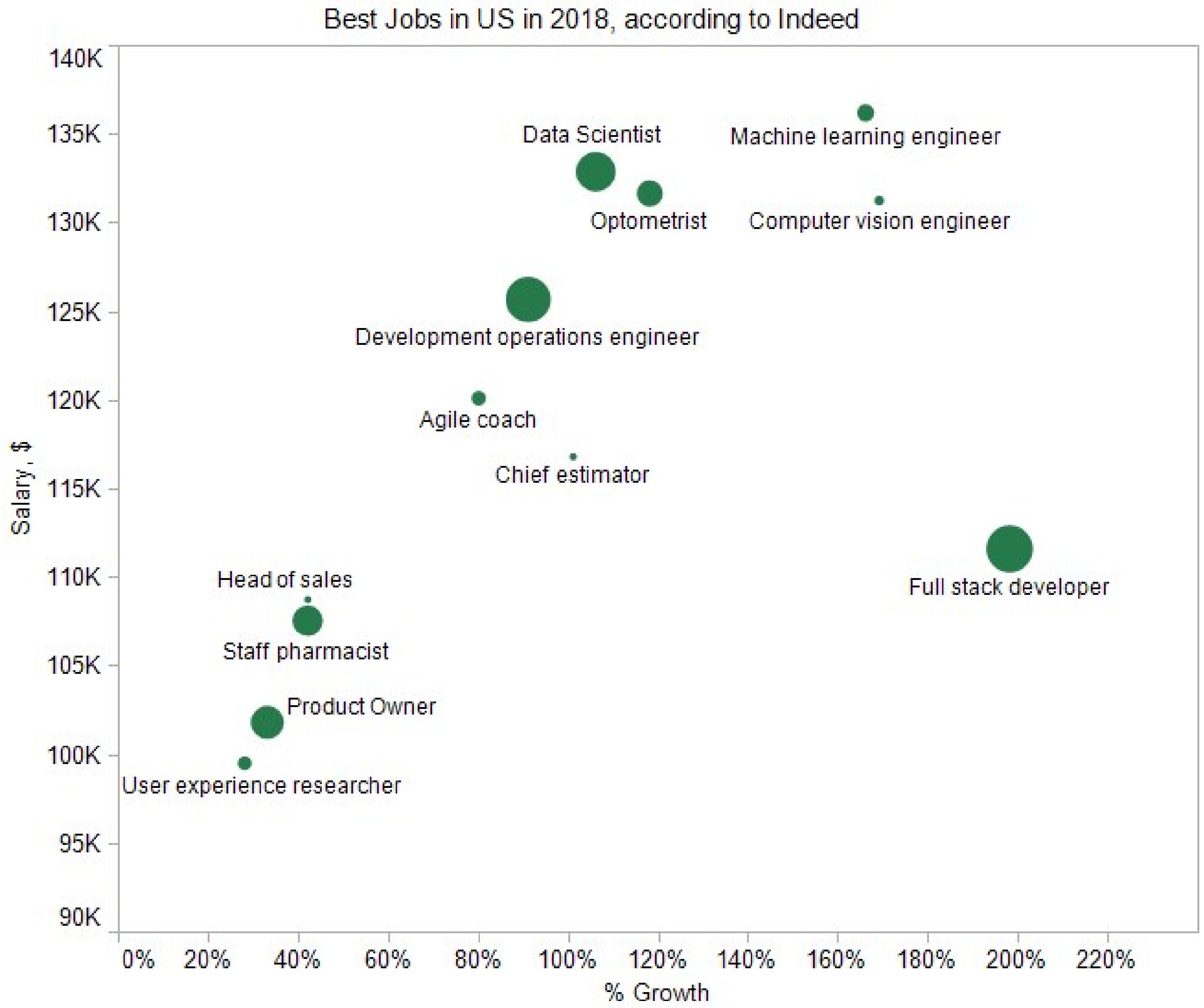
harris interactive

Inria

2021 2 046 pers

1.x. La santé : Identifiée à la fois comme le secteur où le numérique permet le plus de progrès (86%) et celui où ces progrès sont les plus prioritaires (89%).

Machine Learning & Data Scientist



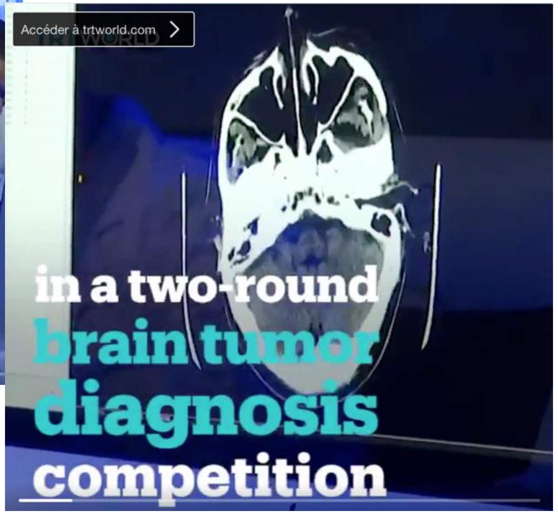
AI in Healthcare …

In 2022, the AI focus area with the most investment was medical and healthcare ($6.1 billion); followed by data management, processing, and cloud ($5.9 billion); and Fintech ($5.5 billion).

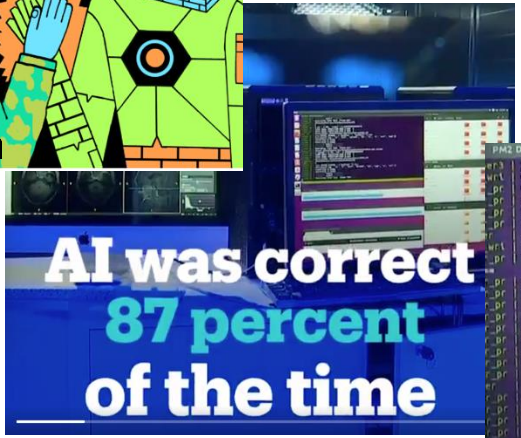
AI Index Report 2023 Stanford Univ

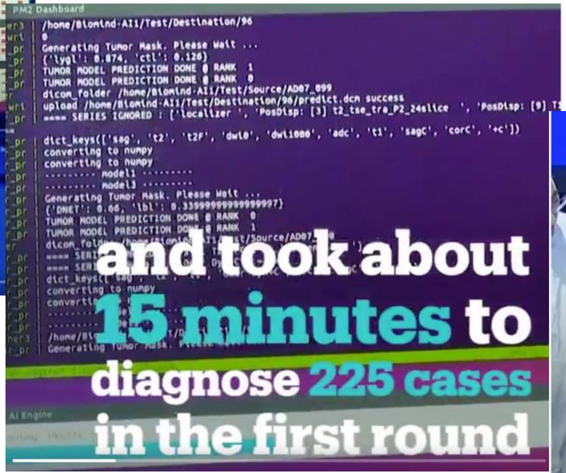

An Artificial Intelligence (AI) system


defeated elite doctors in China

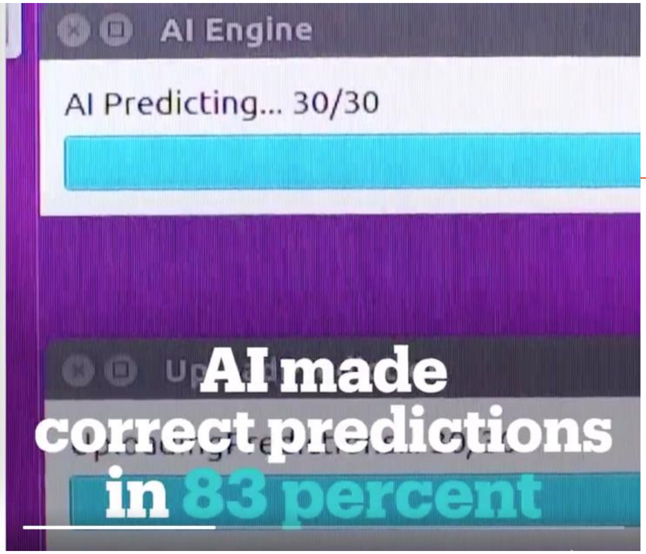

Accéder à trtworld.com
in a two-round brain tumor diagnosis competition

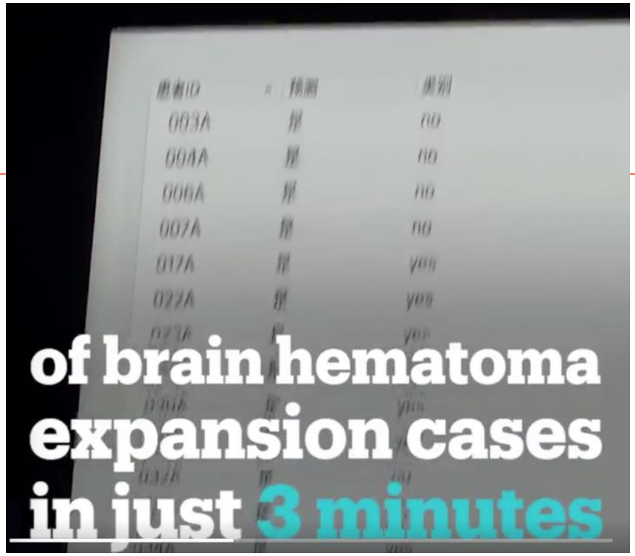

J. Malta

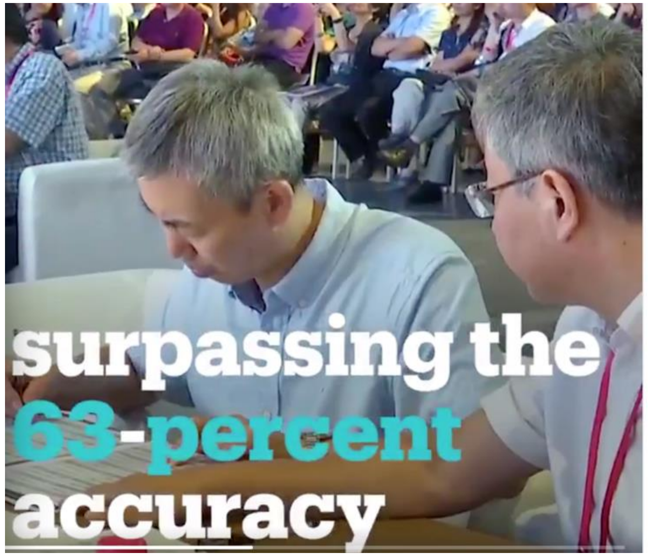

AI was correct 87 percent of the time


and took about 15 minutes to diagnose 225 cases in the first round


while the doctors achieved 66 percent accuracy in 30 minutes

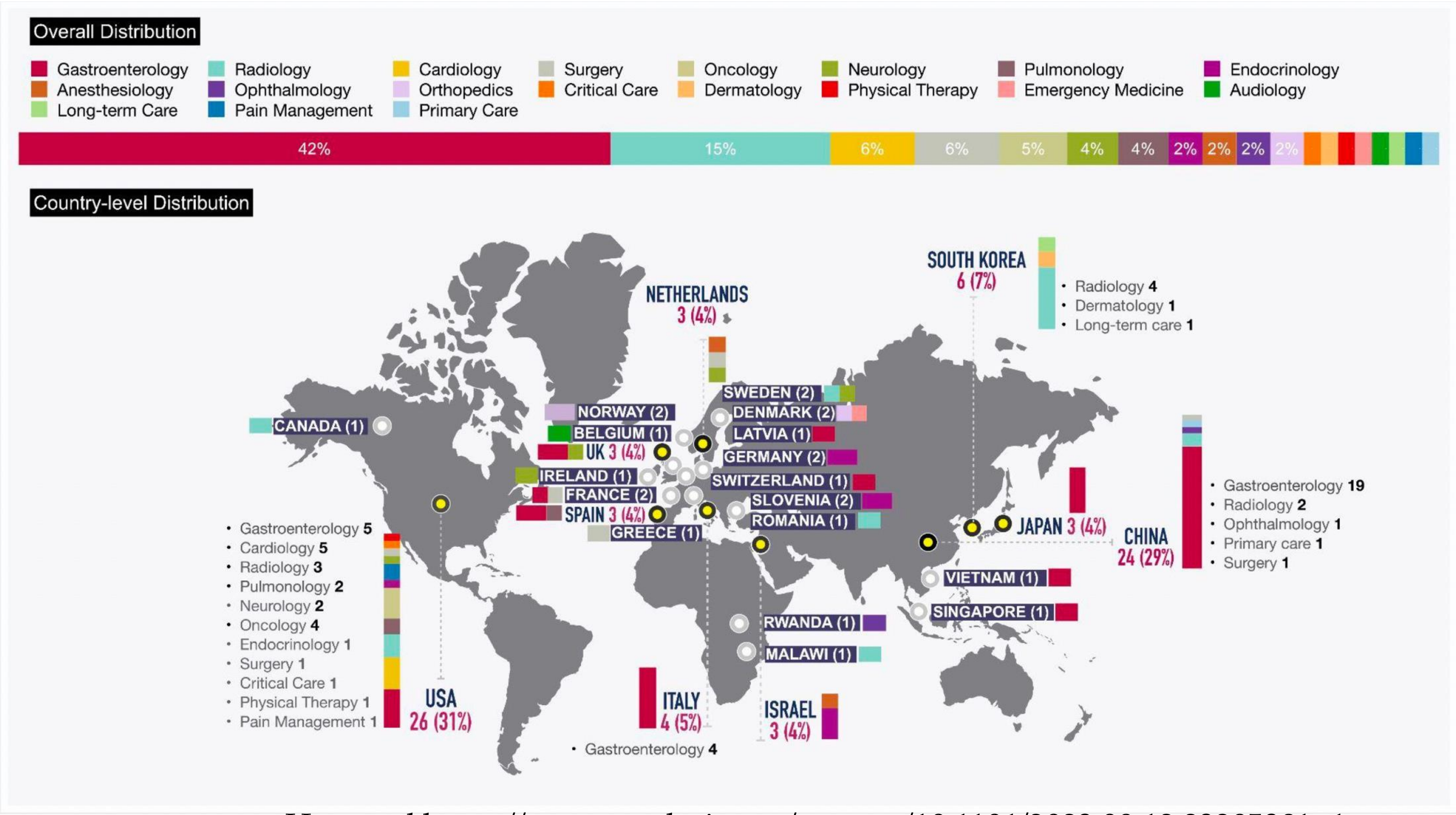






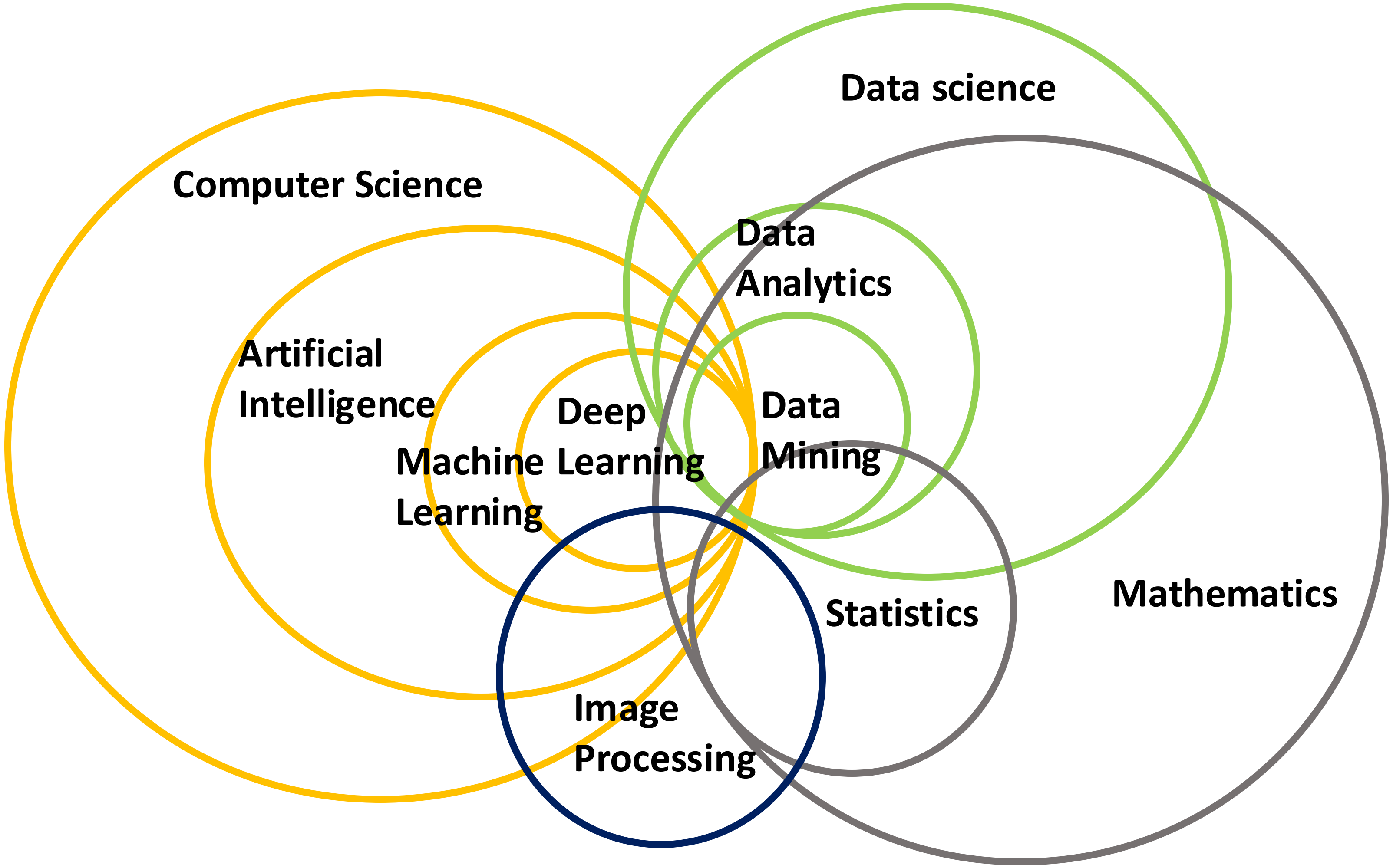
[July 2018 twitter: /AsiaNews_FR/status/1014275512695382017]

A Review of Medical A.I Randomized Trials



Han et al https://www.medrxiv.org/content/10.1101/2023.09.12.23295381v1

Convergence of different domains



AI

- Impact all clinical fields that use imaging data

AI

- Impact all clinical fields that use imaging data (radiology, dermatology, ophtalmology, anatomo-pathology, gastroenterology, neurology, oncology …)

- Psychiatry, Psychopathology developmental
- Pandemic management
- Brain-Computer Interface
- Computer-Assisted Surgery
- Epidemiology
- Health care organisation

AI

- Impact all clinical fields that use imaging data (radiology, dermatology, ophtalmology, anatomo-pathology, gastroenterology, neurology, oncology …)

- Psychiatry, Psychopathology developmental
- Pandemic management
- Brain-Computer Interface
- Computer-Assisted Surgery
- Epidemiology
- Health care organisation

→ **Ethical and Societal impacts**

What AI is?

What AI is?

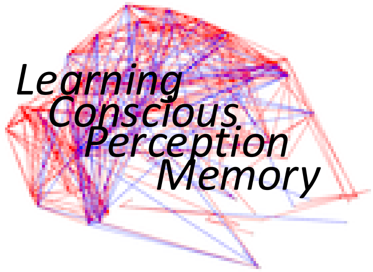
- Systems that act like humans
- Systems that think like humans
- Systems that think rationnally
- Systems that act rationnally

[Russell & Norvig 1995]

What AI is?

- Systems that act like humans
  => **Turing Test**

- Systems that think like humans



- Systems that think rationnally
  => **formal logic**

- Systems that act rationnally
  rational agents (limited rationally)

AI systems: Engineer Approach

- Système d’IA : Un système d’intelligence artificielle (ou système d’IA) est un système automatisé qui, pour un ensemble donné d'objectifs définis par l'homme, est en mesure d’établir des prévisions, de formuler des recommandations, ou de prendre des décisions influant sur des environnements réels ou virtuels.
- Les systèmes d'IA sont conçus pour fonctionner à des degrés d’autonomie divers.

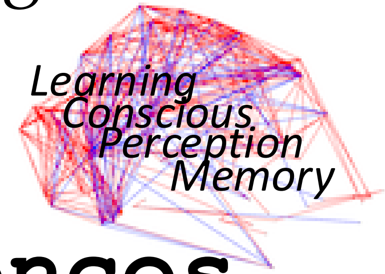
OCDE 2019

Organisation de coopération et de développement économiques

Artificial Intelligence

Intelligent Agent: an entity that takes the best possible action in a situation

- How to build an artificial intelligent agent?
- Test our models of natural intelligent agents?



Computer science

Neurosciences

Integration of heterogeneous datasets

Management of large repositories of data & knowledge

Knowledge discovery

Two main approaches

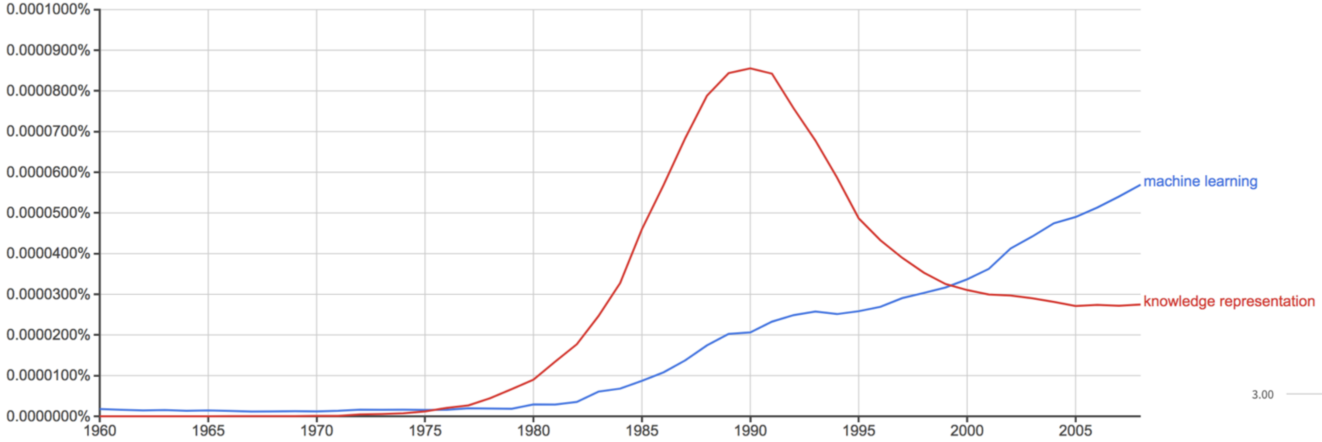
- Machine Learning
  - Bio-inspired
    - Artificial life
    - Neural Networks
      - W McCulloch & W Pitts (1943)
        Artificial neurons
      - D Hebb (1949)
        Learning by modification of connections
      - F Rosenblatt (1963)
        Convergence theorem
      - M Minsky & S Paper
        Perceptrons (1969)
  - Classification (SVM,…)

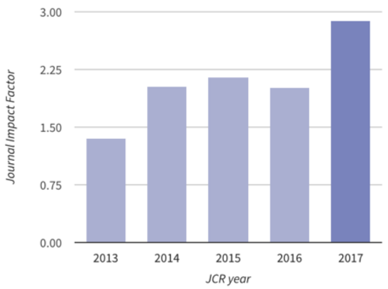
Operations on large vectors

- Symbolic Processing
  - Problem-solving
  - Planning
  - Logic
  - Knowledge representation
    - Common knowledge
    - Meta-knowledge
    - Ontology
  - Multi-agents
  - Co-construction

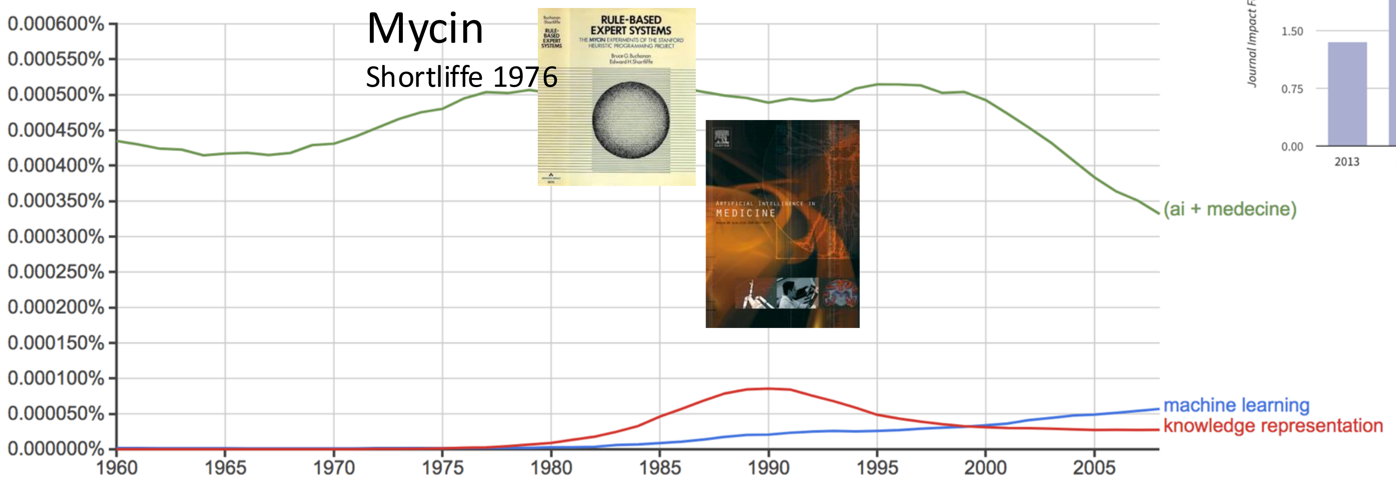
Rule-based manipulation of symbols

But …

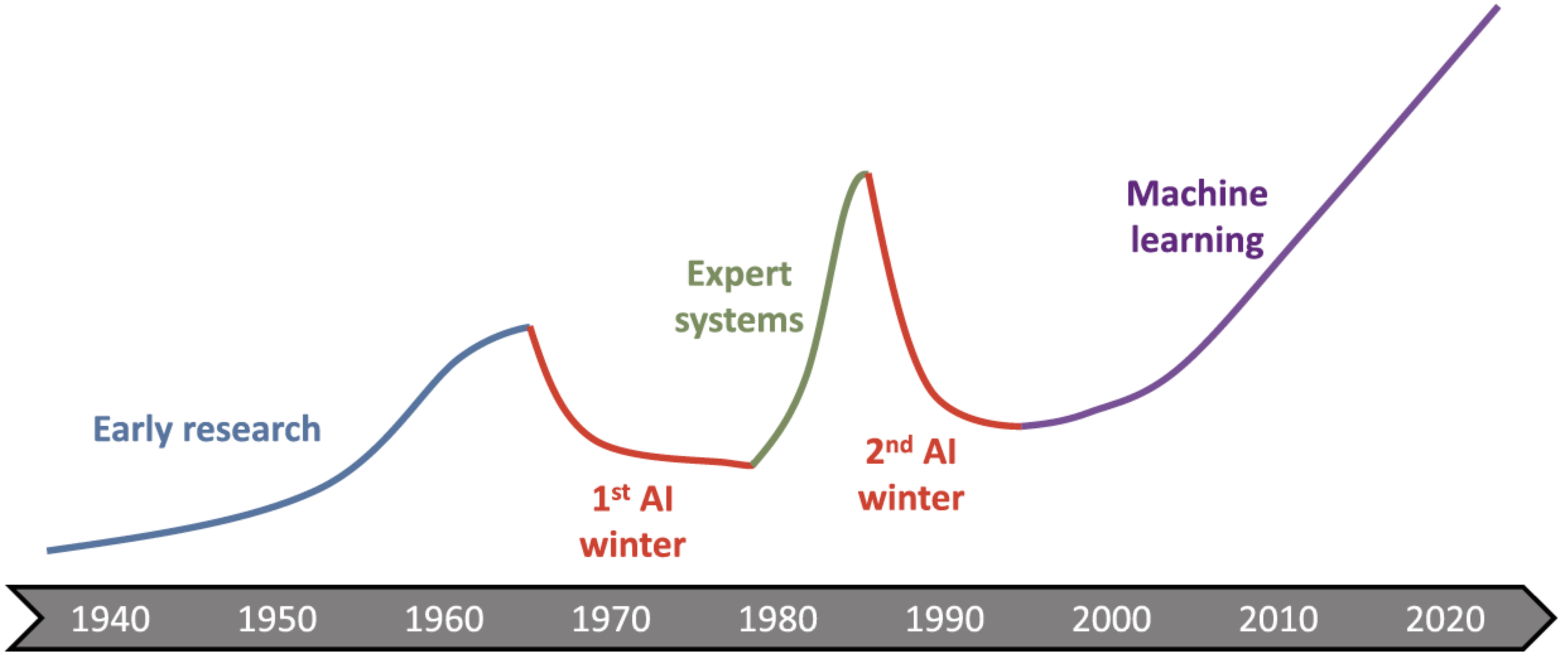




Mycin
Shortliffe 1976



Winters of AI



Colliot 2023 Neuromethods 197 Springer

Evolution connectionist/symbolic

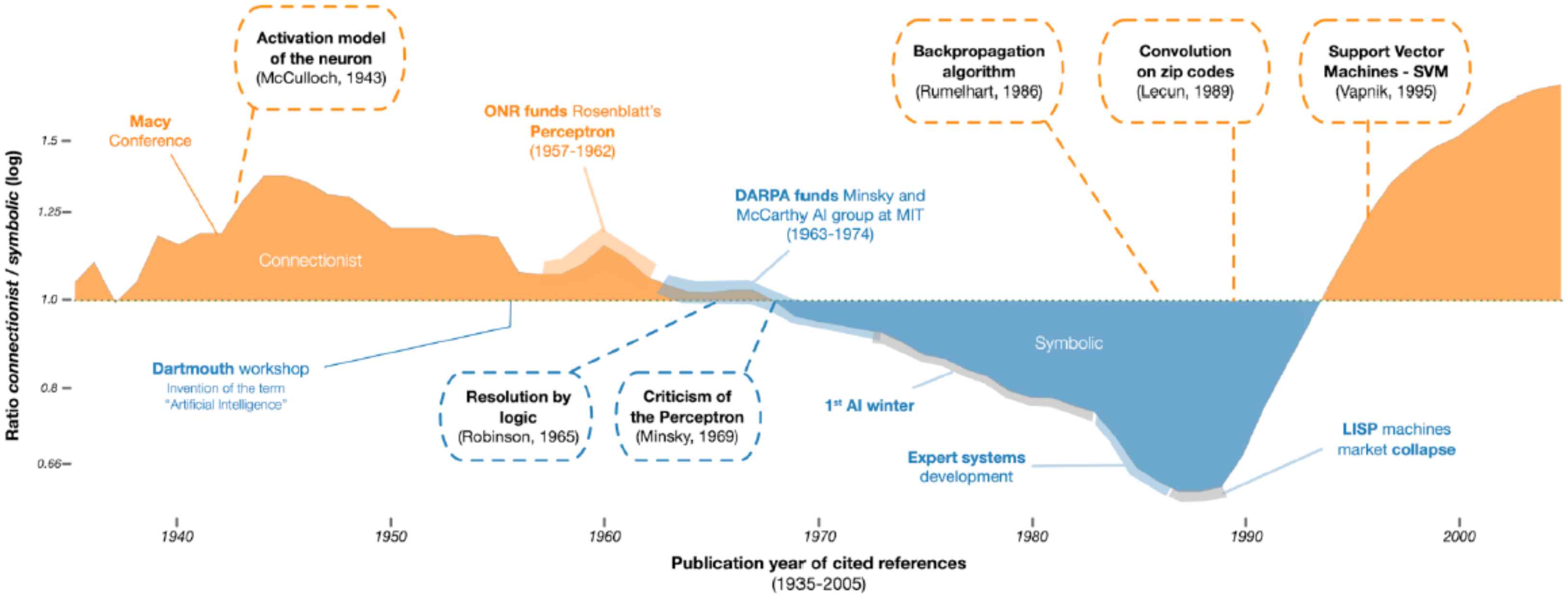
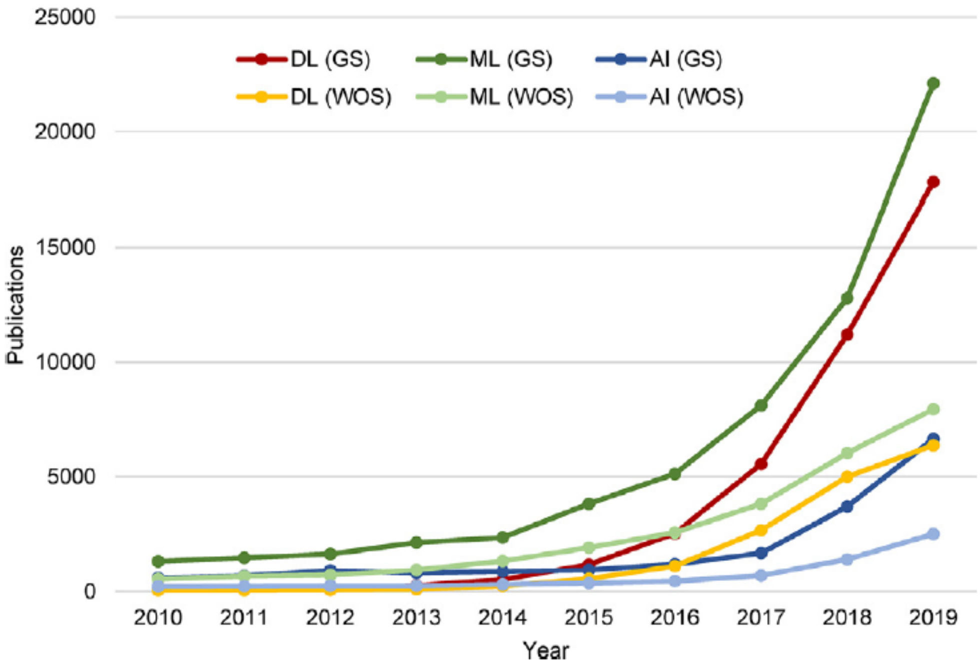


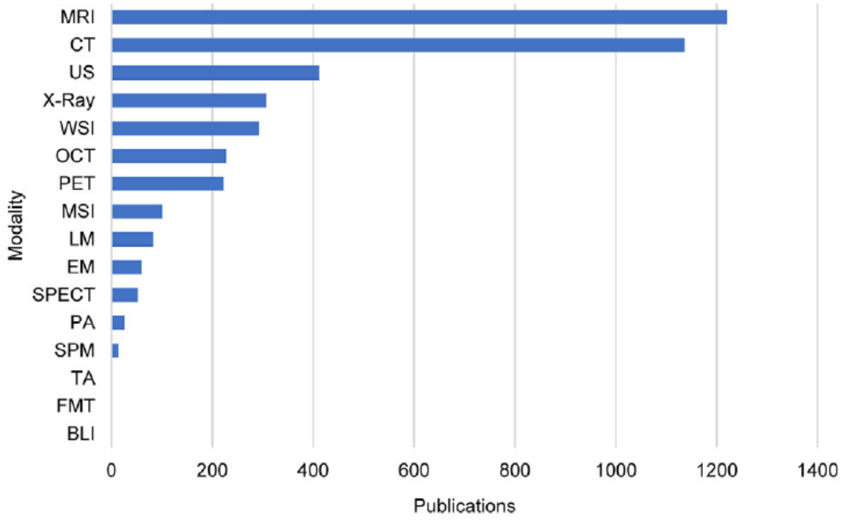
Table 1. The four ages of predictive machines

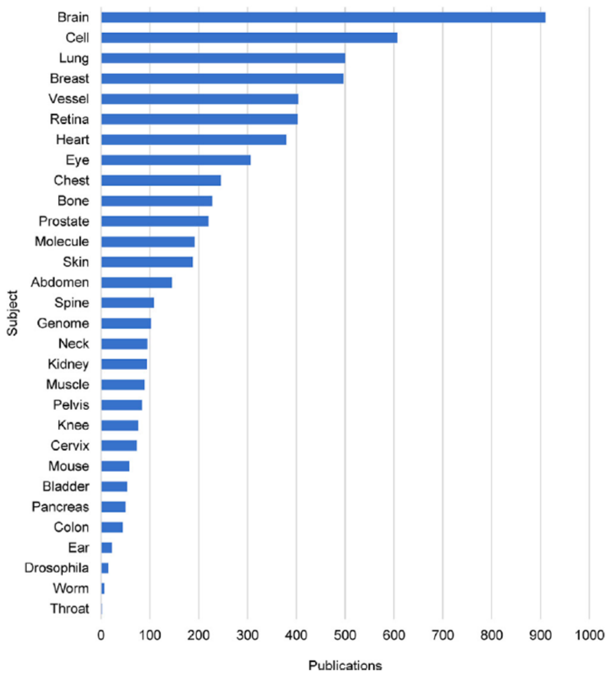
Machine	World	Calculator	Target
Cybernetics (connectionist)	*Environment*	*"Black box"*	*Negative feedback*
Symbolic AI (symbolic)	*"Toy" world*	*Logical reasoning*	*Problem-solving*
Expert systems (symbolic)	*World of expert knowl-edge*	*Selection of hypotheses*	*Examples/counterex-amples*
Deep learning (connectionist)	*The world as a vector of big data*	*Deep neural network*	*Objective-based error optimization*

Cardon et al
Réseaux, 211,
2018, La Découverte

DL&ML&AI





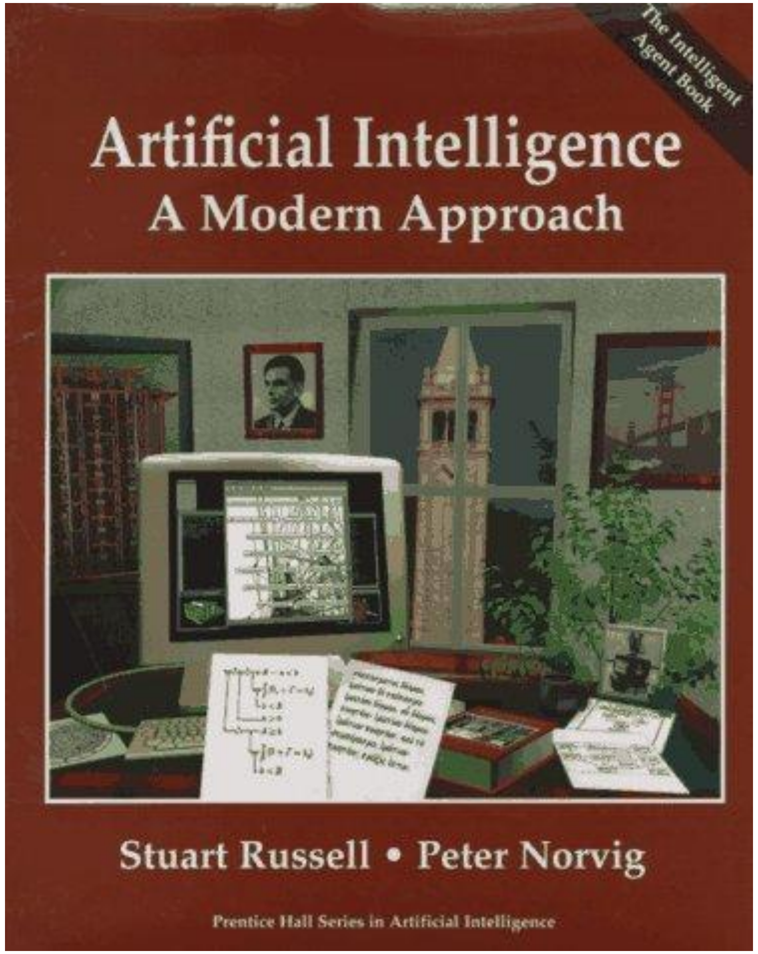


Meijering Comput Struct Biotech J 2020

Knowledge inside …

First ed
1995

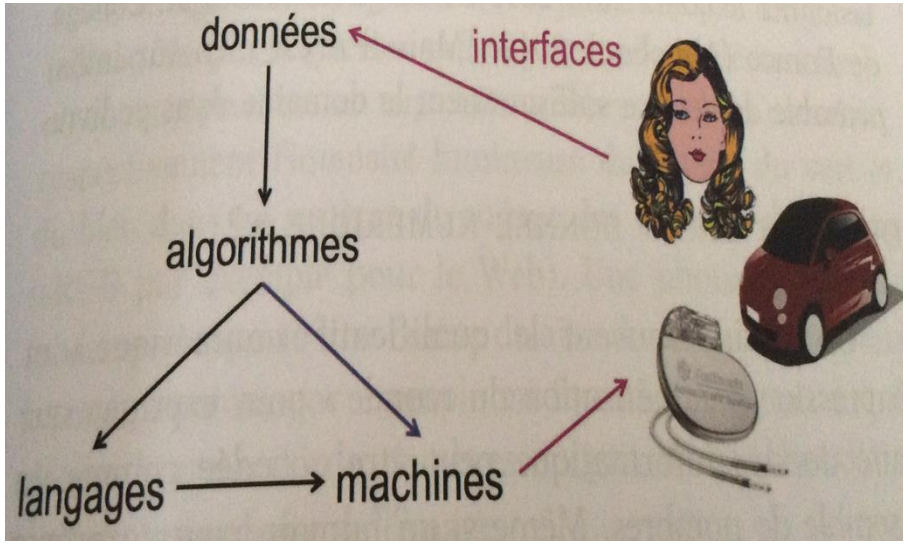
Third ed
2016



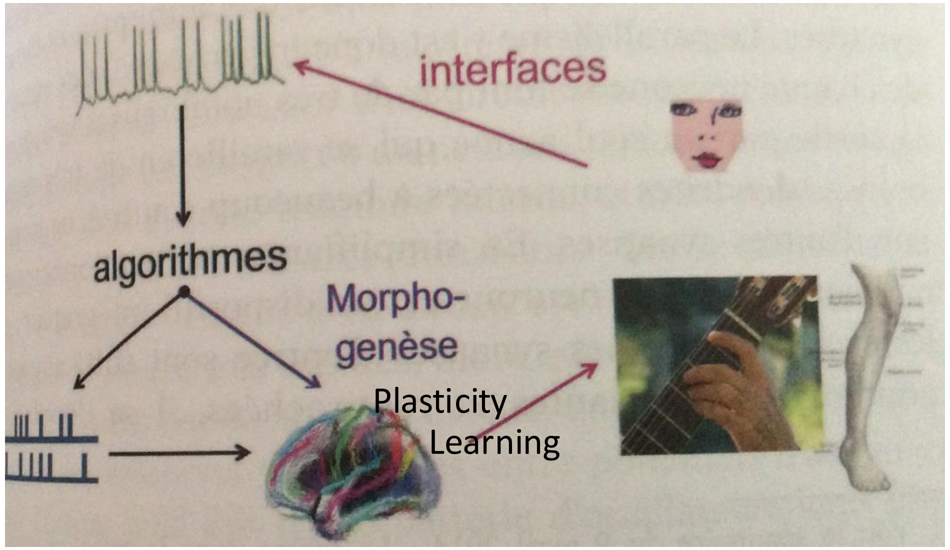
Summary of Contents

Computer & Brain

Computer science roots

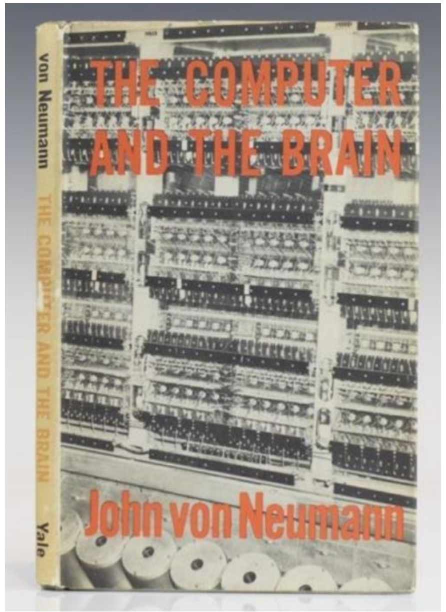


Neurosciences roots

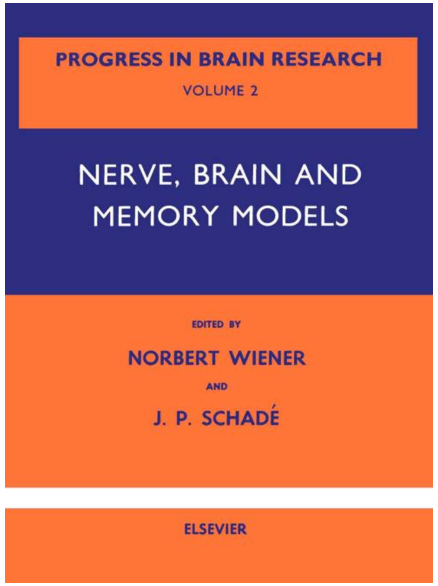


Berry G. L'hyperpuissance des ordinateurs O. Jacobs 2017

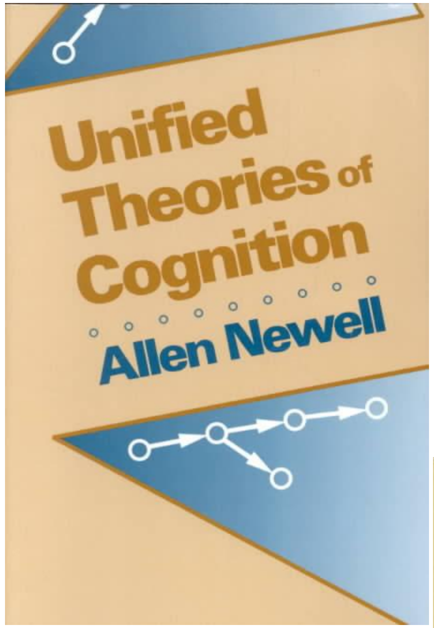
Computer & Brain



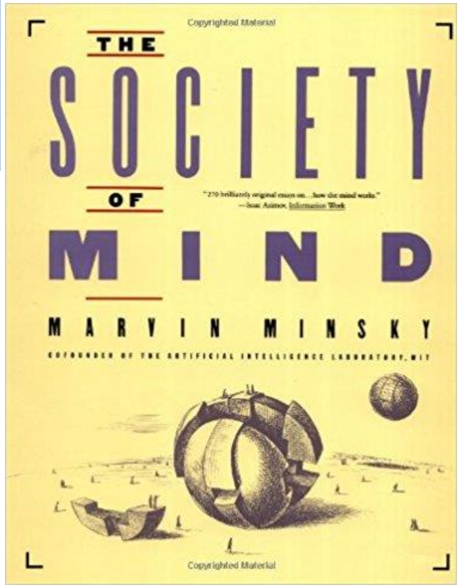
1958
Yale University Press, New Haven



1963

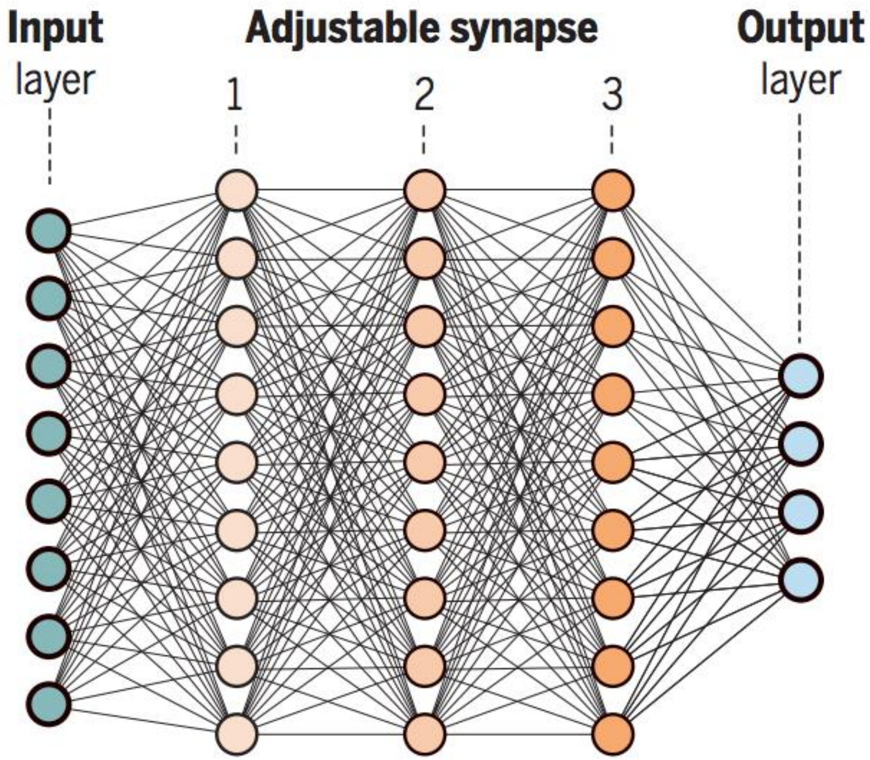


1982



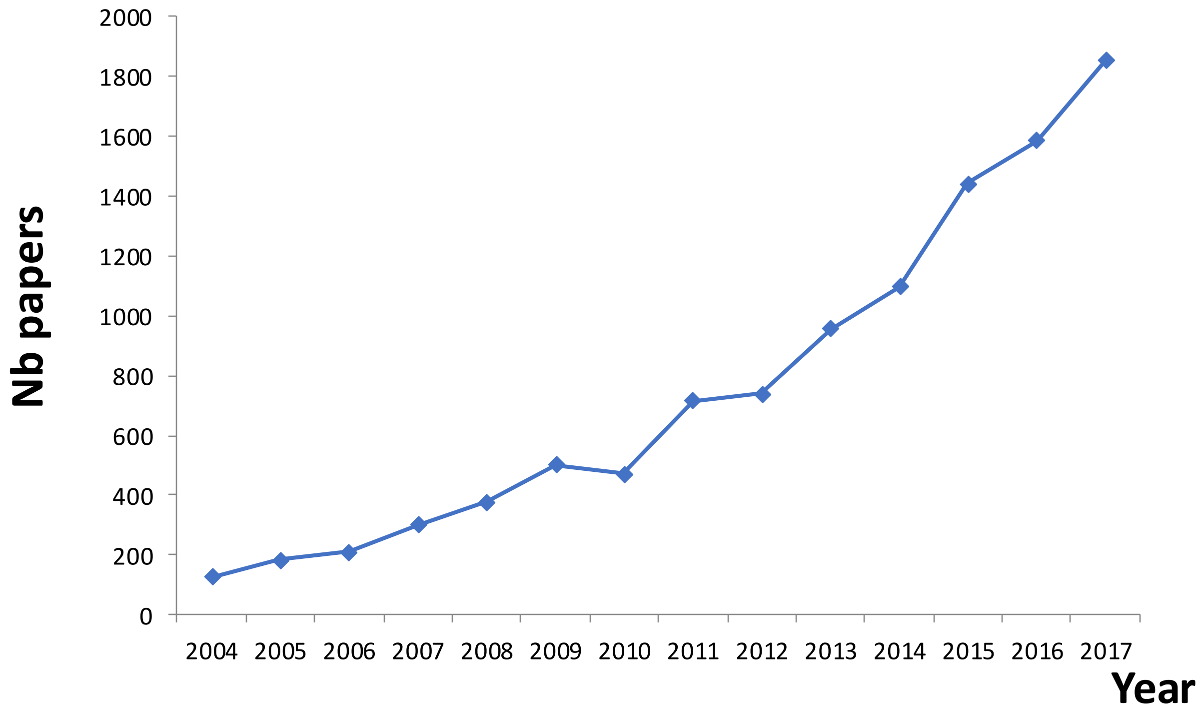
1988

Analogy …



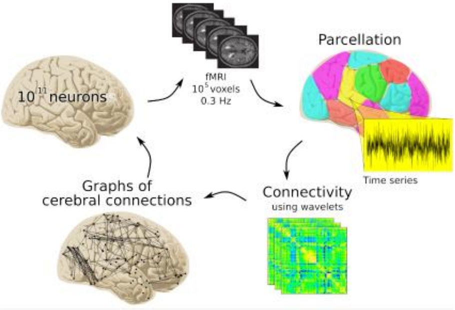
Ullman Science 2019

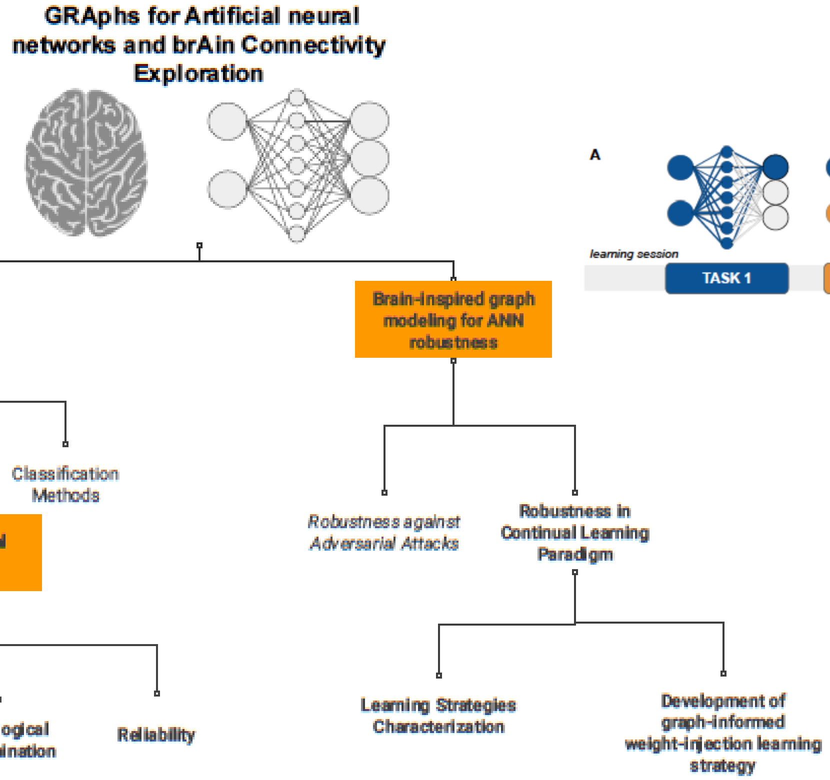
AI in Neurosciences



Wos: brain & (svm or multi-voxel or machine learning or decoding or classifier)

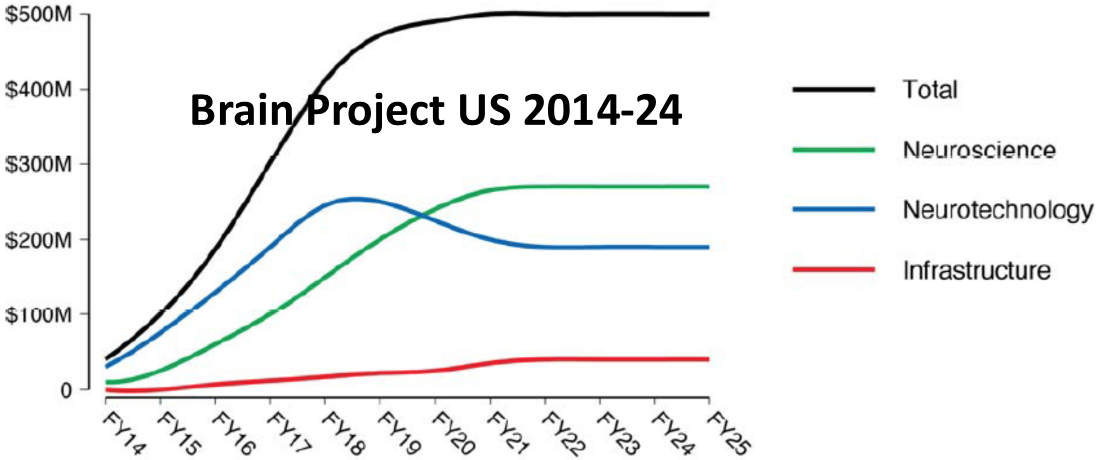
Analogy …





Carboni PhD 2023

Projects



Brain Project US 2014-24

Humain Brain Project EC 2013-23

Neuroinformatics Platform
Brain Simulation Platform
HPAC Platform
Medical Informatics Platform
Neuromorphic Computing Platform
Neurorobotics Platform

54 M€ (2013-16)
89 M€ (2016-18)
88 M€ (2018-20)

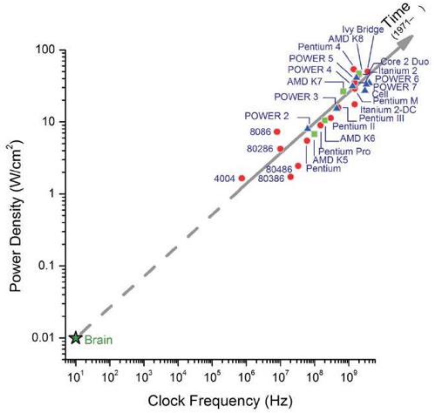
2018-2021: 40-80 M€
2024-2028: 70 M€

Grenoble, Nice, Paris, Toulouse

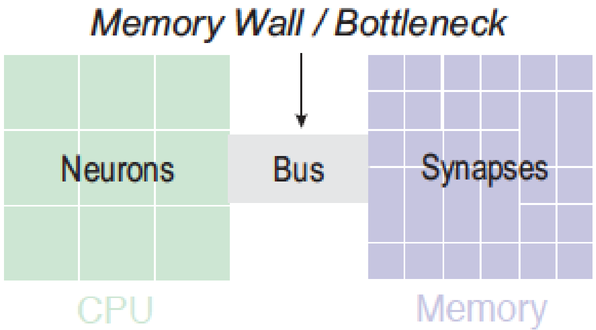
Computer & Brain: architecture -I

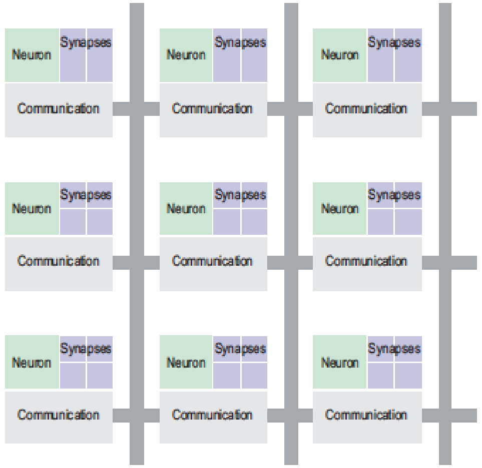
	Computer	Brain
Computational units	Core, 10^5 transistors	10^{11} neurons
Storage units	10^{10} ram, 10^{12} HD	10^4 synapses/ neurons => 10^{15}
Memory - Computation	Separated	Embedded

Architecture highly wired
& reconfigurable

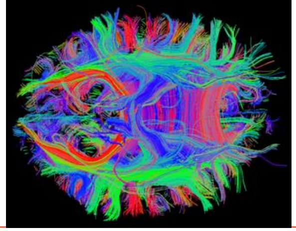


Merolla et al Science 2014





Computer & Brain: architecture-II



	Frontier supercomputer (June 2020)	Human brain
Speed	1.102 exaFLOPS	~1 exaFLOPS (estimate)
Power requirements	21 MW	10–20 W
Dimensions	680 m^2 (7,300 sq ft)	1.3–1.4 kg (2.9–3.1 lb)
Cost	$600 million	Not applicable
Cabling	145 km (90 miles)	850,000 km (528,000 miles) of axons and dendrites
Memory	75 TB/s read; 35 TB/s write; 15 billion IOPS flash storage system, along with the 700 PB Orion site-wide Lustre file system	2.5 PB (petabyte)
Storage	58 billion transistors	125 trillion synapses, which can store 4.7 bits of information each

Smirnova et al Front Science 2023

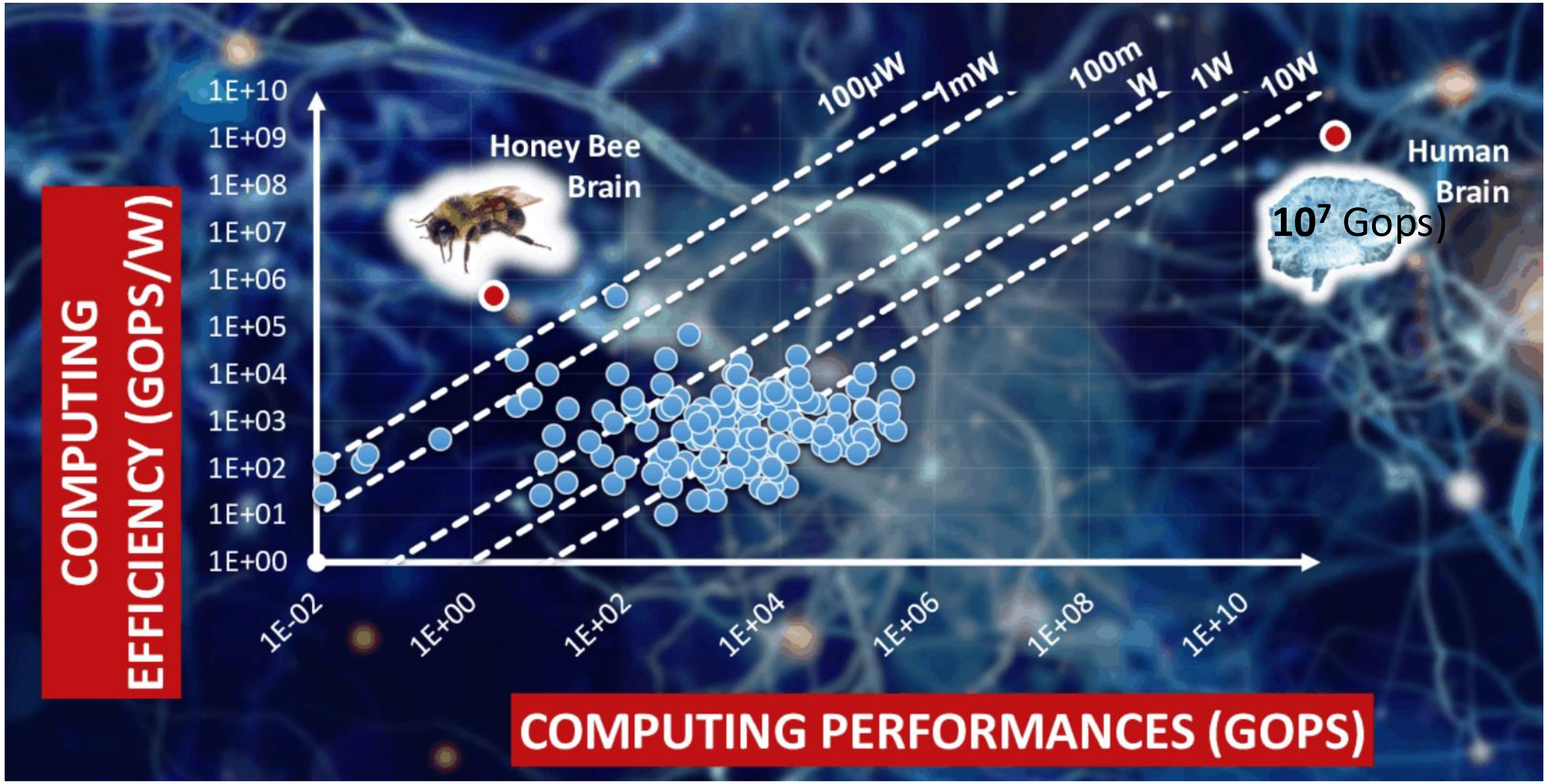
Visual system 10×10^9 bits/s on the retina

=> 6×10^6 bits/s optical nerve transmission

=> 10^4 bits/s reach V1

=> 100 bits/s for conscious visual perception

Consumation



Two main approaches

- Machine Learning
  - Bio-inspired
    - Artificial life
    - Neural Networks
      - W McCulloch & W Pitts (1943)
        Artificial neurons
      - D Hebb (1949)
        Learning by modification of connections
      - F Rosenblatt (1963)
        Convergence theorem
      - M Minsky & S Paper
        Perceptrons (1969)
  - Classification (SVM,…)

- Symbolic Processing
  - Problem-solving
  - Planning
  - Logic
  - Knowledge representation
    - Common knowledge
    - Meta-knowledge
    - Ontology
  - Multi-agents
  - Co-construction

Operations on large vectors (world2vec)
- Elementary data
- Different models
- Prediction on the world itself (inductive machines)

Manipulation of symbols
- Semantic attached to symbol
- One model
- Hypothetical-deductive machine

Two main approaches

- Machine Learning
  - Classification (SVM,…)
  - Bio-inspired
    - Artificial life
    - Neural Networks
      - WMcCulloch & W Pitts (1943)
      - Artificial neurons
      - D Hebb (1949) Learning by modification of connections
      - F Rosenblatt (1963) Convergence theorem
      - M Minsky & S Paper Perceptrons (1969)

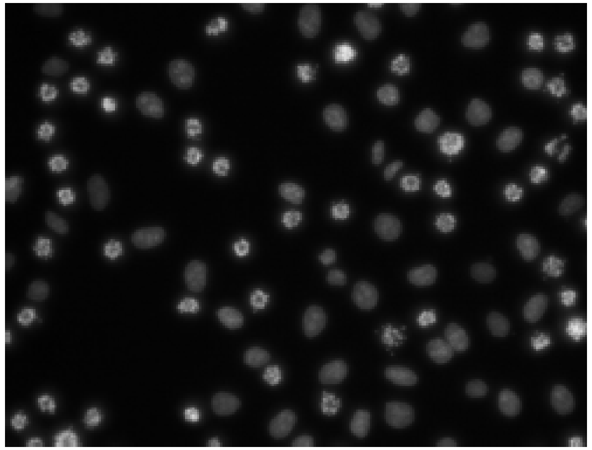
- Knowledge representation
  – Problem-solving
  – Planning
  – Logic
  – Common sense knowledge
  – Meta-knowledge
  – Ontology
  – Multi-agents
  – Co-construction

New algorithms
- **CNN (Y. Le Cun)**
- **GPU**
- **Performances (Speech, Vision)**

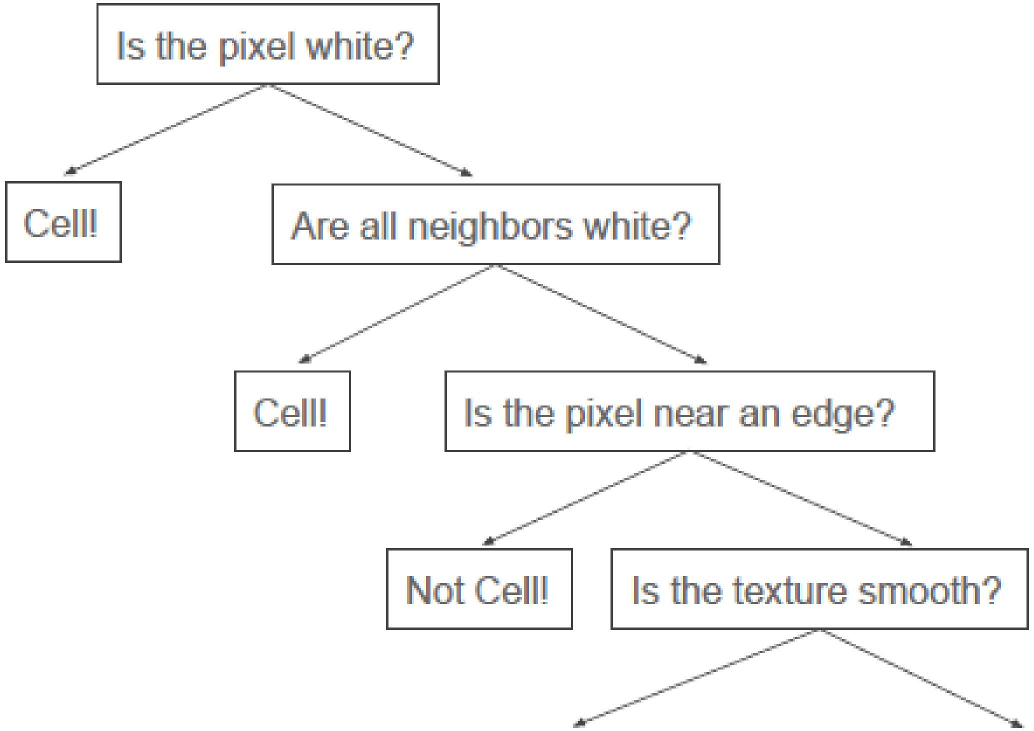
Limitations
- **Rules are not enough**
- **Frame problem**
- **Evolution**
- **Explanation**

Rule-based approach

Cells vs background segmentation



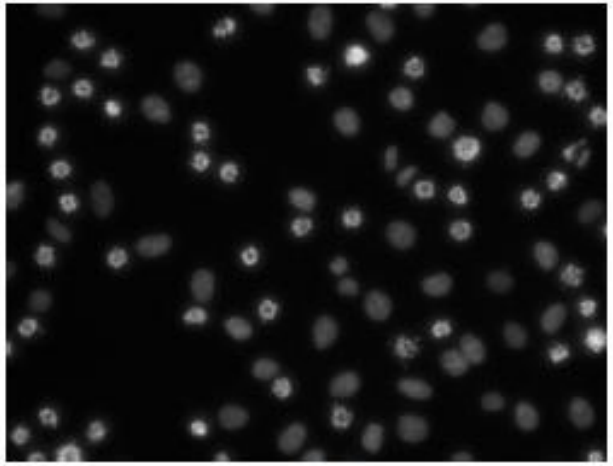
[Image: Gerlich Lab]



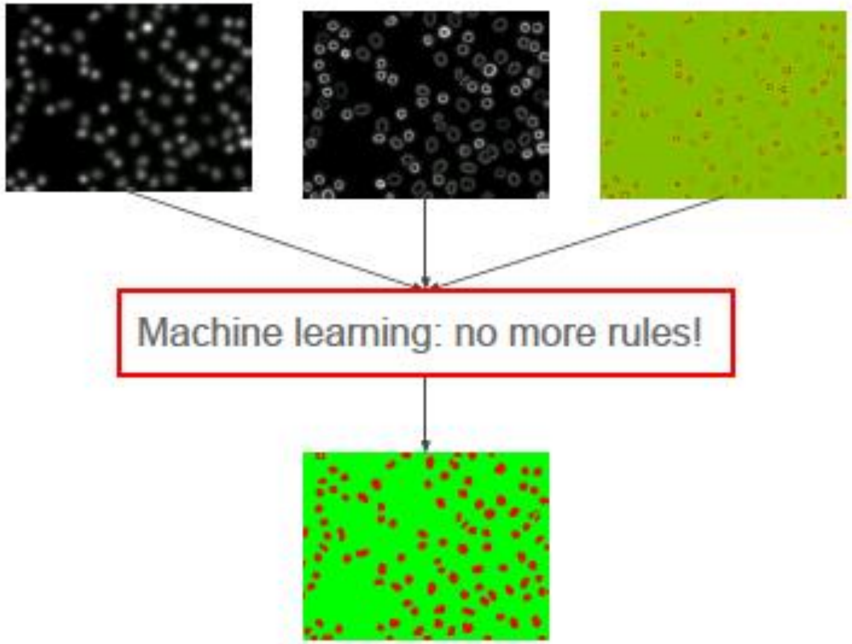
Courtesy A. Kreshuk

ML approach

Cells vs background segmentation

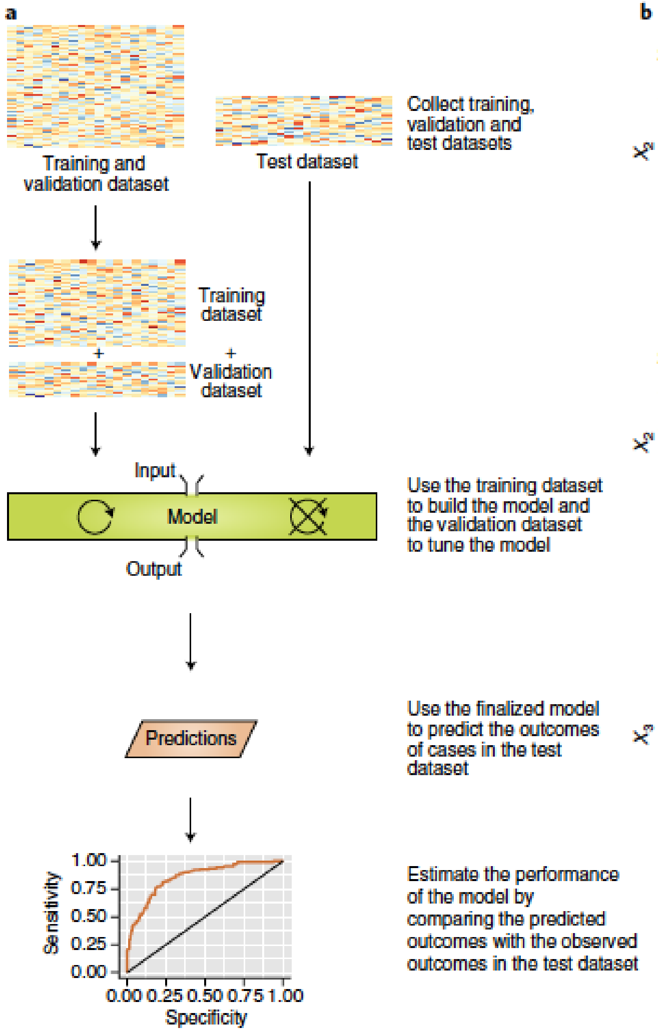


[Image: Gerlich Lab]

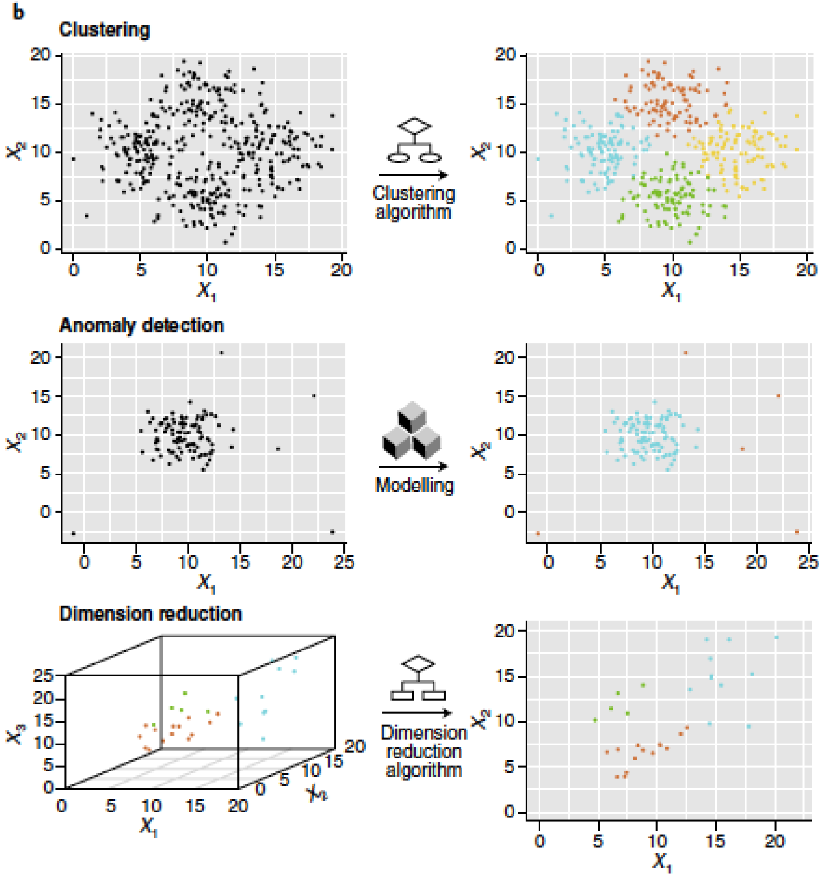


Courtesy A. Kreshuk

Supervised & Unsupervised ML



Supervised

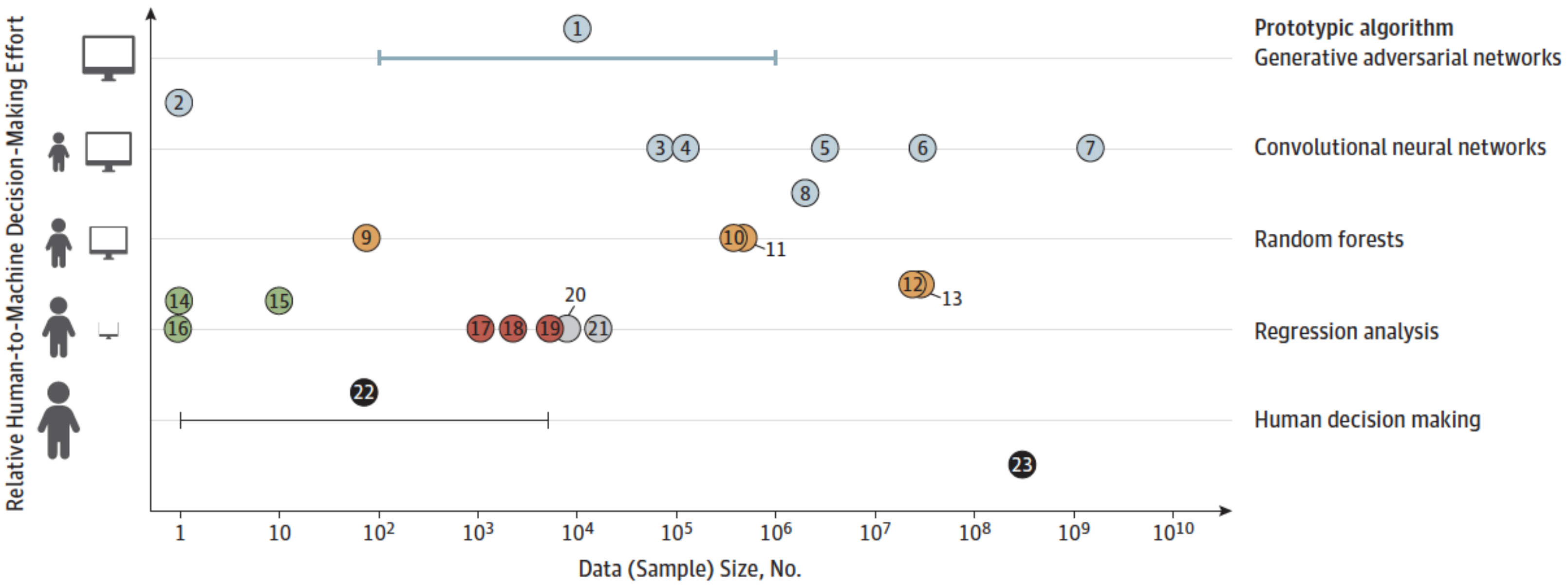


Unsupervised

Yu et Kohane Nat Biomedeng 2018

ML: Natural extension of traditional statistical approaches

Beam & Kohane Nature 2018



Deep learning

1. Generative adversarial networks (2014)
2. Google AlphaGo Zero (2017)
3. ATM check readers (1998)
4. Google diabetic retinopathy (2016)
5. ImageNet computer vision models (2012-2017)
6. Google AlphaGo (2015)
7. Facebook Photo Tagger (2015)
8. Prediction of 1-y all-cause mortality (2017)

Classic machine learning

9. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling (2002)
10. EHR-based CV risk prediction (2017)
11. Netflix Prize winner (2006)
12. Google Search (1998)
13. Amazon product recommendation (2003)

Expert AI systems

14. MYCIN (1975)
15. CASNET (1982)
16. DXplain (1986)

Risk calculators

17. CHA_2DS_2-VASc Score for atrial fibrillation stroke risk (2017)
18. MELD end-stage liver disease risk score (2001)
19. Framingham CV risk score (1998)

Randomized Clinical Trials

20. Celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis (2002)
21. Use of estrogen plus progestin in healthy postmenopausal women (2002)

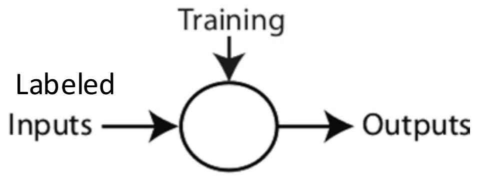
Other

22. Clinical wisdom
23. Mortality rate estimates from US Census (2010)

ML Approaches

Supervised learning

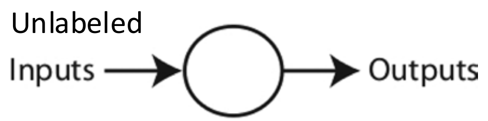
- Classification
  - Decision trees
  - K-nearest neighbours
  - Support vector machine
  - Random forests
  - Naive Bayes classifier
- Regression
  - Linear/non-linear regression
  - Local regression (LOESS)
  - Ordinary least squares regression
  - Neural networks



Learns known patterns
Predicts outcome

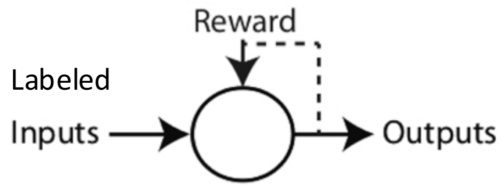
Unsupervised learning

- Cluster analysis
  - Hierarchical clustering
  - K-means clustering
  - Gaussian Mixture
- Dimension reduction
  - Linear discriminant analysis
  - Principal component analysis



Learns unknown patterns
Find hidden patterns

Reinforcement learning



Generates data
Interacts with environment

Transfert learning al

Applies learned patterns to a different but related task

[From Choy et al. Radiology 2017]

Neural networks take over other machine-learning methods

Percentage of papers that mention each method

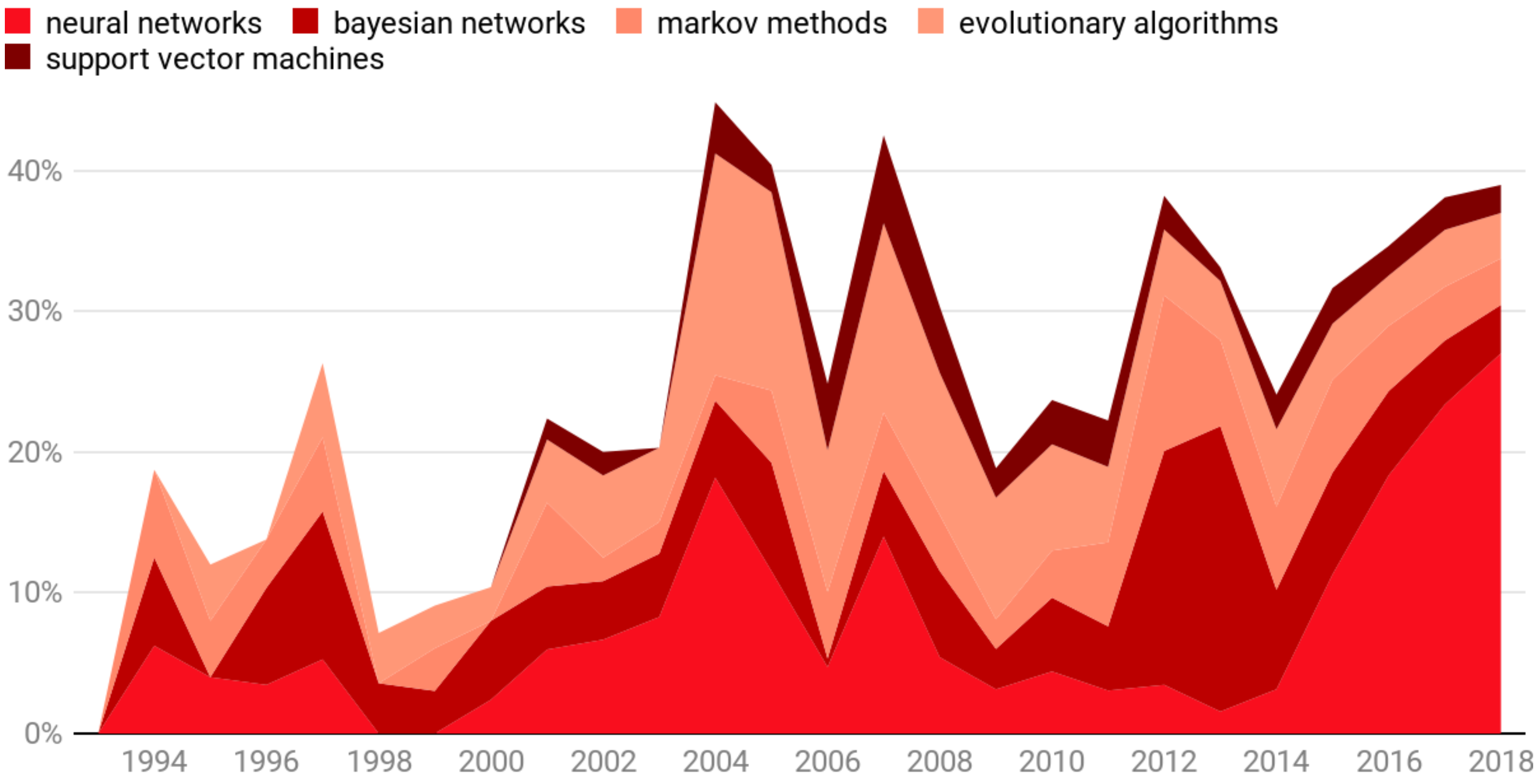
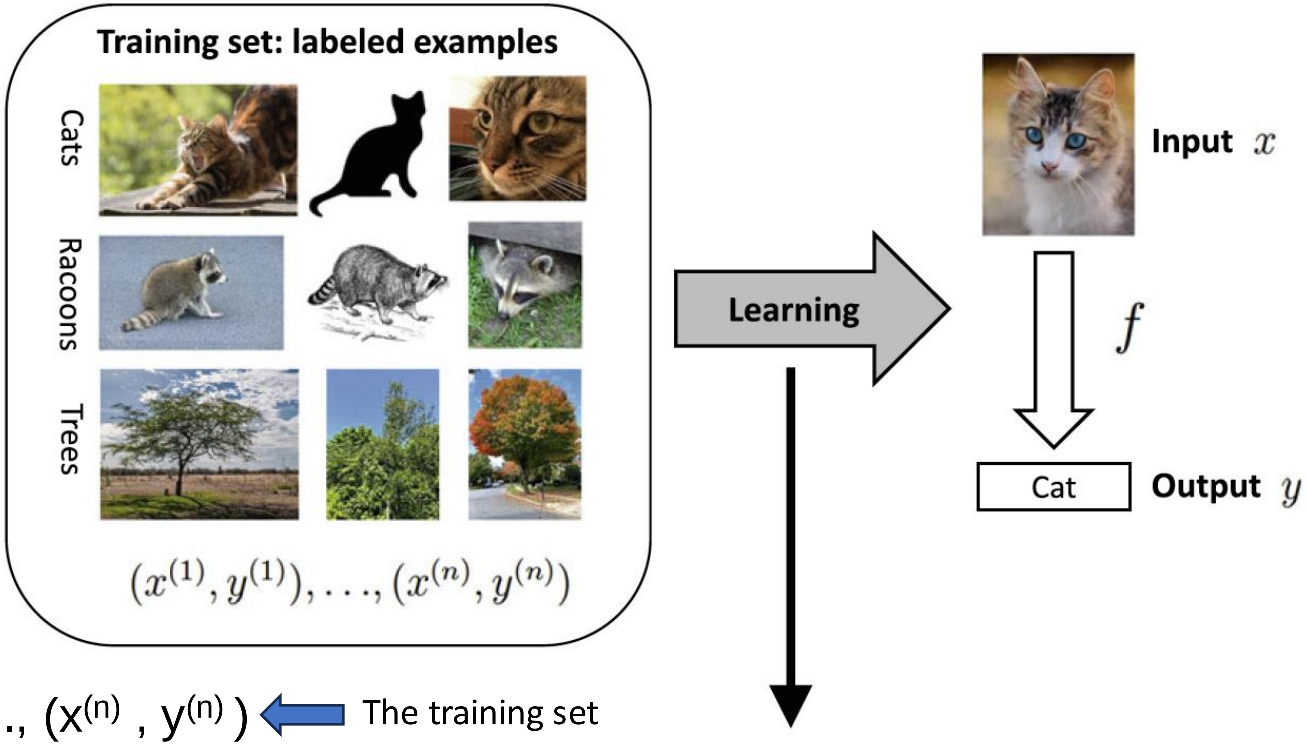


Chart: MIT Technology Review • Source: arXiv.org • Created with Datawrapper

ML

Algorithm	Prediction Speed	Training Speed	Memory Usage	Required Tuning	General Assessment
Logistic Regression (and Linear SVM)	Fast	Fast	Small	Minimal	Good for small problems with linear decision boundaries
Decision Trees	Fast	Fast	Small	Some	Good generalist, but prone to overfitting
(Nonlinear) SVM (and Logistic Regression)	Slow	Slow	Medium	Some	Good for many binary problems, and handles high-dimensional data well
Nearest Neighbor	Moderate	Minimal	Medium	Minimal	Lower accuracy, but easy to use and interpret
Naïve Bayes	Fast	Fast	Medium	Some	Widely used for text, including spam filtering
Ensembles	Moderate	Slow	Varies	Some	High accuracy and good performance for small- to medium-sized datasets
Neural Network	Moderate	Slow	Medium to Large	Lots	Popular for classification, compression, recognition, and forecasting

Supervised learning



$(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})$ The training set

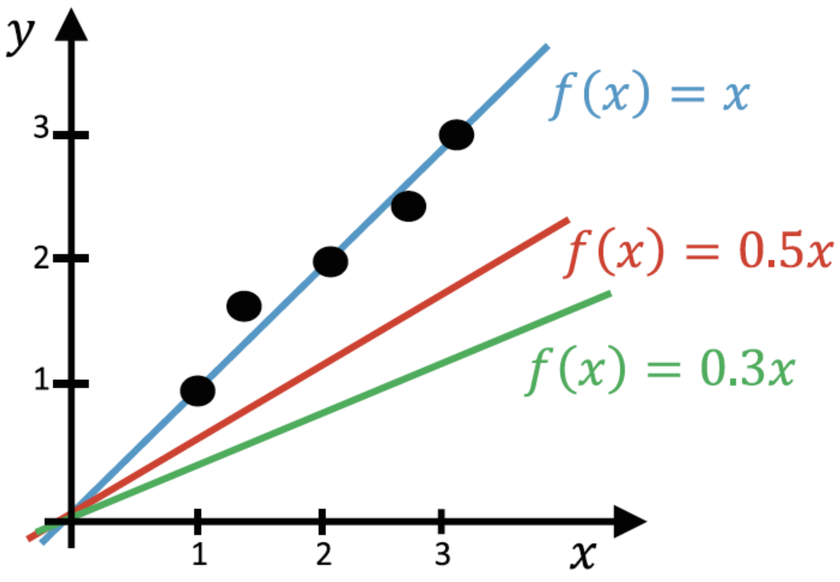
$$\hat{f} = \arg\min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} \ell\left(y^{(i)}, f(x^{(i)})\right)$$

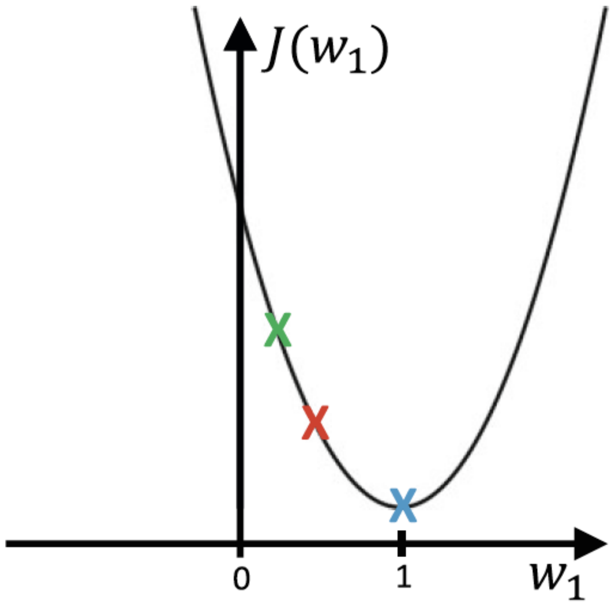
To minimize

Eg. Least squares loss $l(y,f(x))=(y-f(x))^2$

Colliot 2023 Neuromethods 197 Springer

An example





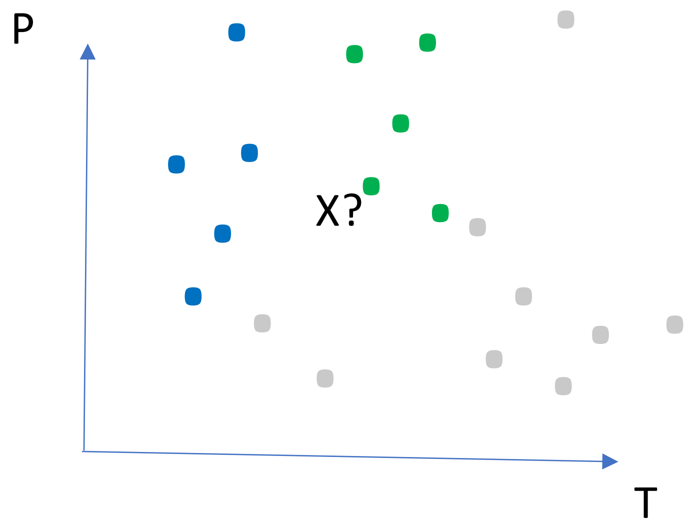
$f(x)=w_1x$

$(y,f(x))=(y-f(x))^2$

Colliot 2023 Neuromethods 197 Springer

More complex

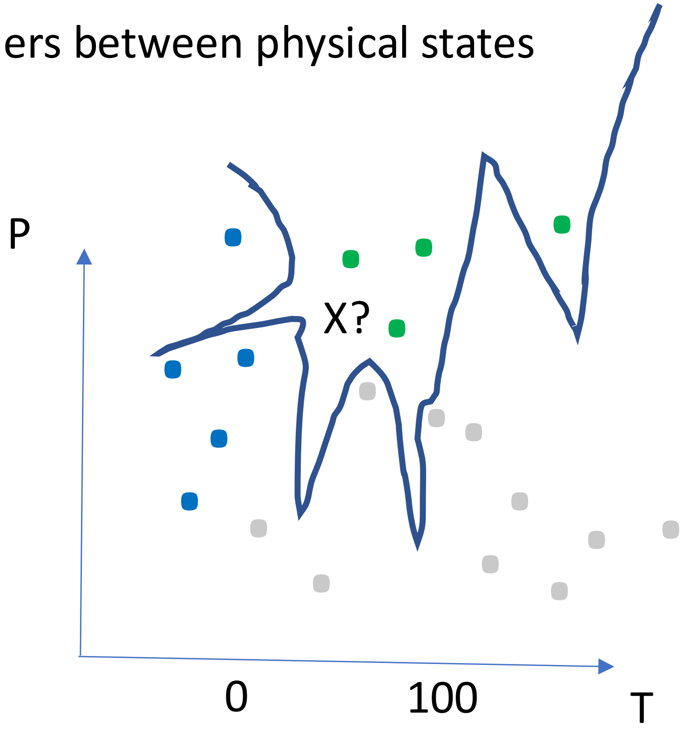
Water physical state function of P and T?



Example extracted from A. Maillard Coll. de France course

Classification

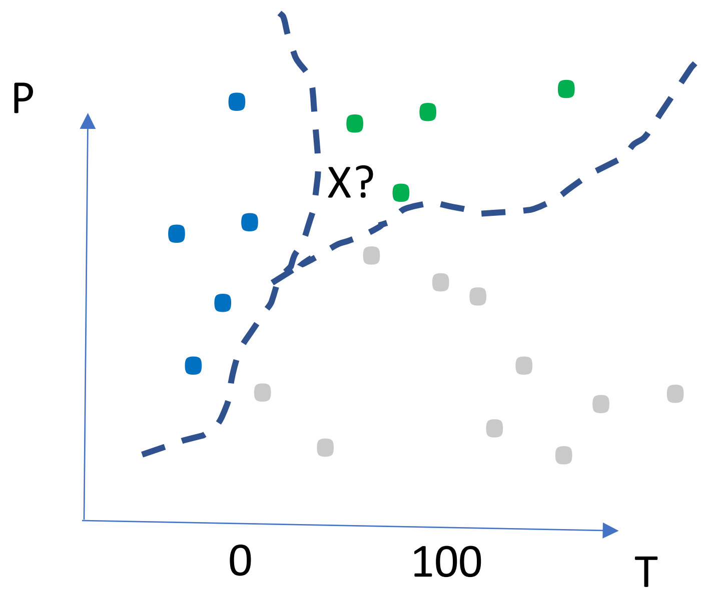
Estimate borders between physical states



Example extracted from A. Maillard Coll. de France course

Classification

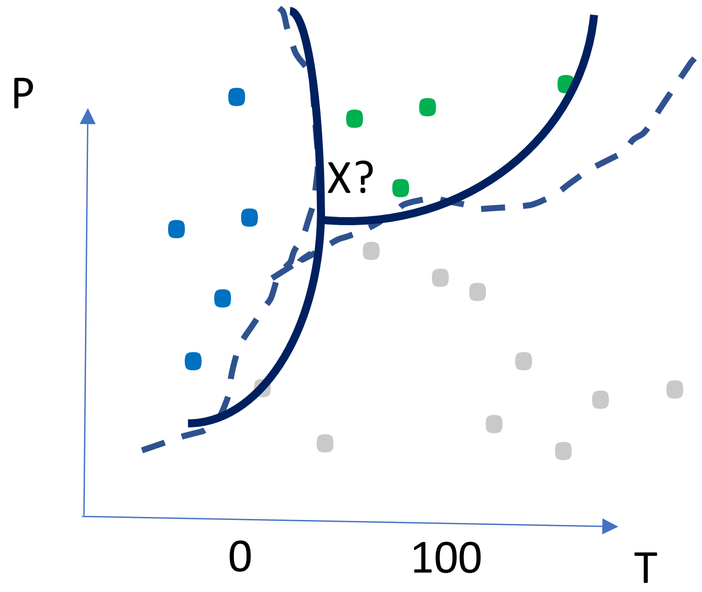
Estimate borders between physical states



Example extracted from A. Maillard Coll. de France course

Classification

Generalisation when sufficient number of examples is available; smooth frontiers



Phase diagram

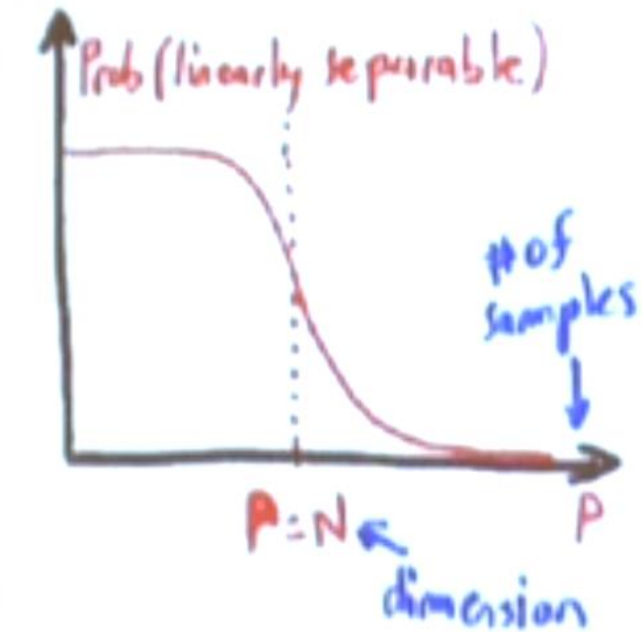
Example extracted from A. Maillard Coll. de France course

Curse of Dimensionality

P

T

- Many examples P to have one close to the target T



- Difficulty when P increases to linearly separate based on N dimensions (Cover's theorem 1966)

- Solution increase N for a large P set !!

Principle

Box 3: Summary of main concepts

- The input x
- The output y
- The training samples $(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})$
- The model: transforms the input into the output

$$f \text{ such that } y = f(x)$$

- The set of possible models $\mathcal{F}$
- The loss: measures the error between the predicted and the true output, for a given sample

$$\ell(y, f(x))$$

- The cost function: measures the average error across the training samples

$$J(f) = \frac{1}{n}\sum_{i=1}^{n} \ell\left(y^{(i)}, f\left(x^{(i)}\right)\right)$$

- Learning process: finding the model which minimizes the cost function

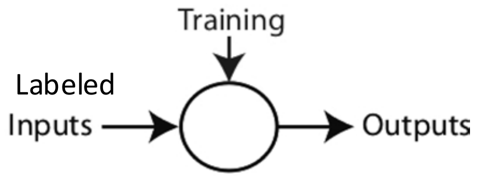
$$\hat{f} = \arg\min_{f \in \mathcal{F}} J(f)$$

Colliot 2023 Neuromethods 197 Springer

ML Approaches

Supervised learning

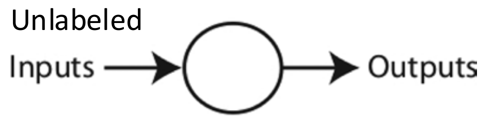
- Classification
  - Decision trees
  - K-nearest neighbours
  - Support vector machine
  - Random forests
  - Naive Bayes classifier
- Regression
  - Linear/non-linear regression
  - Local regression (LOESS)
  - Ordinary least squares regression
  - Neural networks



Learns known patterns
Predicts outcome

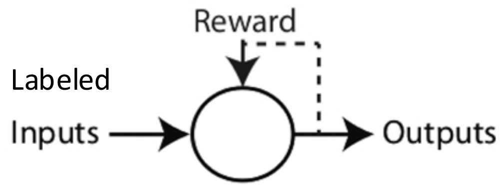
Unsupervised learning

- Cluster analysis
  - Hierarchical clustering
  - K-means clustering
  - **Gaussian Mixture**
- Dimension reduction
  - Linear discriminant analysis
  - Principal component analysis



Learns unknown patterns
Find hidden patterns

Reinforcement learning



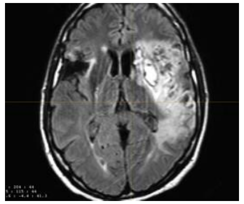
Generates data
Interacts with environment

Transfert learning al

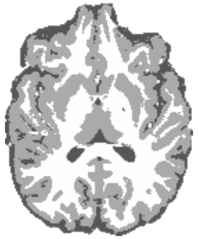
Applies learned patterns to a different but related task

[From Choy et al. Radiology 2017]

Image Segmentation

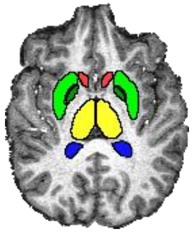


Observed intensities



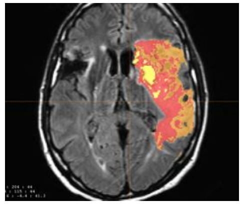
Tissue labels

White matter
Grey matter
Cerebrospinal fluid



Structure labels

Putamen
Ventricule
….



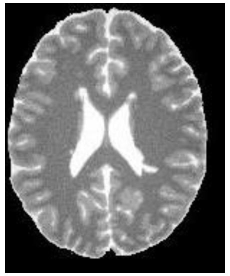
Lesion detection

Size and localisation

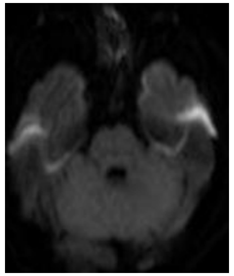
Difficulties

Inhomogeneity

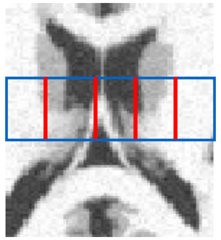
Noise



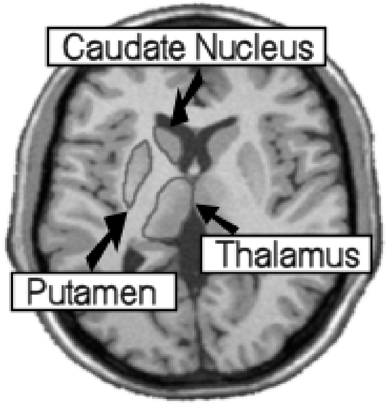
Low contrast

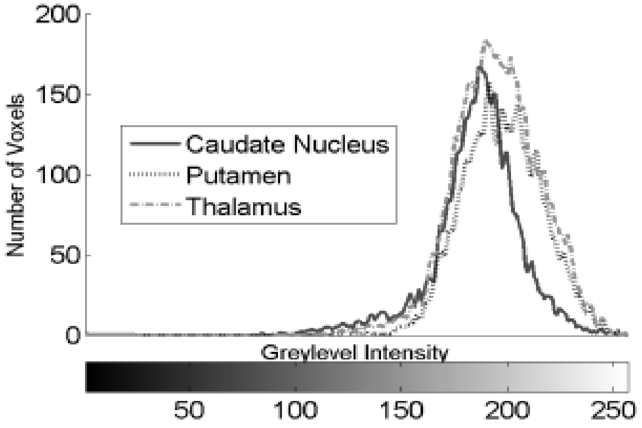


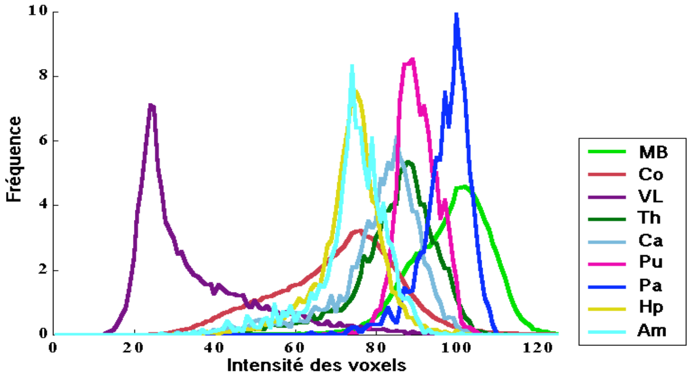
Artefacts



Partial volume effect







Inter-structures variations

(Fischl et al., 2002)

- Intensity distributions overlap
  - ⇨ Need for *a priori* anatomical knowledge

Bayes law

- **Modélisation du processus d'imagerie**

Image en niveaux de gris =
réalisation d'un champ aléatoire

$$\mathbf{y} = \{y_1, \ldots y_i, \ldots, y_N\}$$

Segmentation (« étiquettes ») =
réalisation d'un champ aléatoire

$$\mathbf{z} = \{z_1, \ldots z_i, \ldots, z_N\}$$

avec $z_i \in \{e_1, \ldots, e_k, \ldots, e_K\}$

Loi de Bayes : $p(\mathbf{z}|\mathbf{y},\Phi) = \dfrac{p(\mathbf{y}|\mathbf{z},\Phi_{\mathbf{y}})\,p(\mathbf{z}|\Phi_{\mathbf{z}})}{p(\mathbf{y})}$

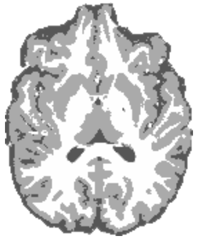
Estimation des modèles gaussiens avec prise en compte ou non du voisinage

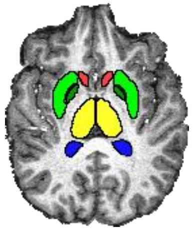
Terme d'attache aux données

Information a priori sur les étiquettes

Maximisation p(z|y,Φ)

Gaussian mixture

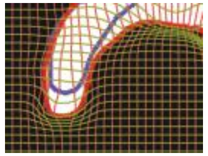


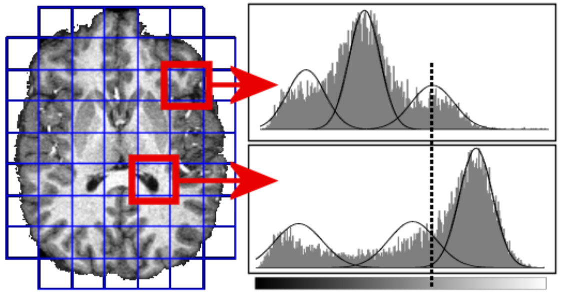


$$p(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y})$$

Parametric model estimation:

(1) intensity local distribution
(2) local atlas registration





Multi-agents approach
Distributed models

Scherrer et al. TMI (2009)

Model

Forbes et al. AISTAT conf 2010

Bayes Law :

$$p(\mathbf{z}|\mathbf{y},\Phi) = \frac{p(\mathbf{y}|\mathbf{z},\Phi_{\mathbf{y}})\,p(\mathbf{z}|\Phi_{\mathbf{z}})}{p(\mathbf{y})}$$

Data term: Gaussian models

Maximisation p(z|y,Φ)

$$p(\mathbf{z}|\mathbf{y},\Phi) = W^{-1}\exp\left(-H(\mathbf{z}|\mathbf{y},\Phi)\right) \text{ with } W = \sum_{z}\exp\left(-H(\mathbf{z}|\mathbf{y},\Phi)\right)$$

$$H(\mathbf{z}|\mathbf{y},\Phi) = \sum_{i=1}^{N}\left[{}^{t}z_i\alpha_i - \frac{\beta}{2}\sum_{j\in N(i)}{}^{t}z_i z_j\right] - \sum_{i\in S}\log p\left(y_i|z_i,\Phi_y\right)$$

A priori knowledge

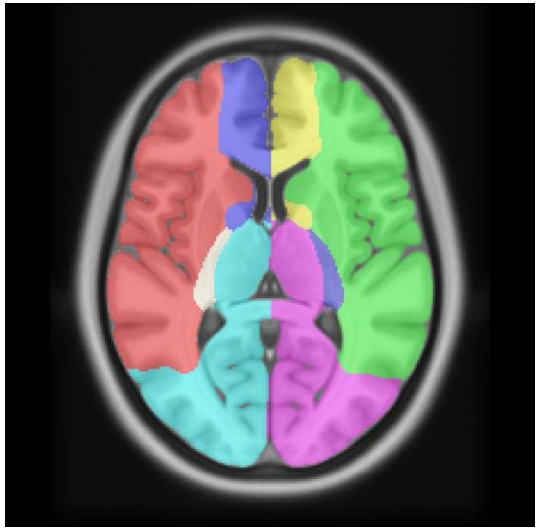
Regularisation term

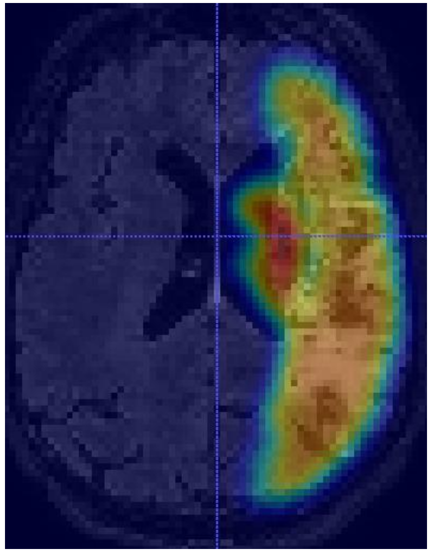
Data driven term based on intensities

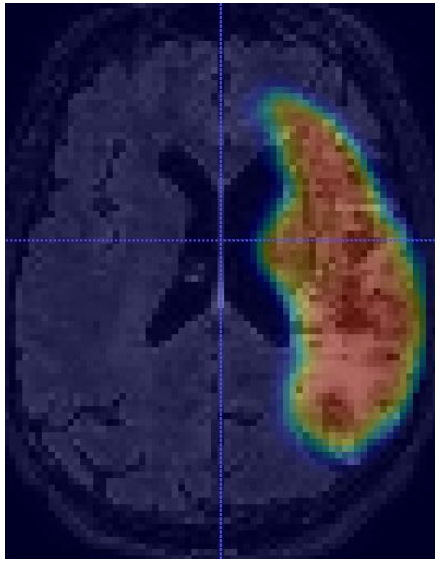
A priori knowledge

Our model considers:

- 4 normal tissue classes (GM, WM, CSF & other)
- 6 subclasses for the lesion class
- A probabilistic vascular territory atlas

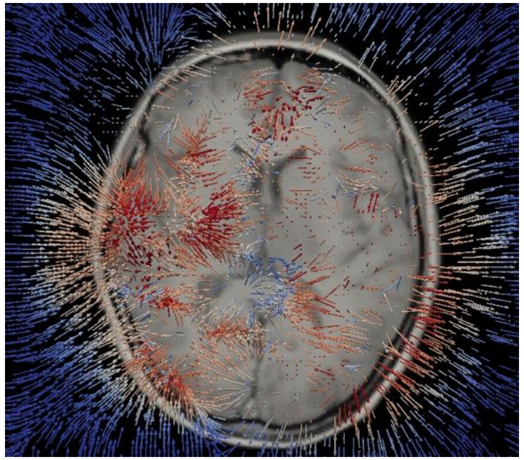




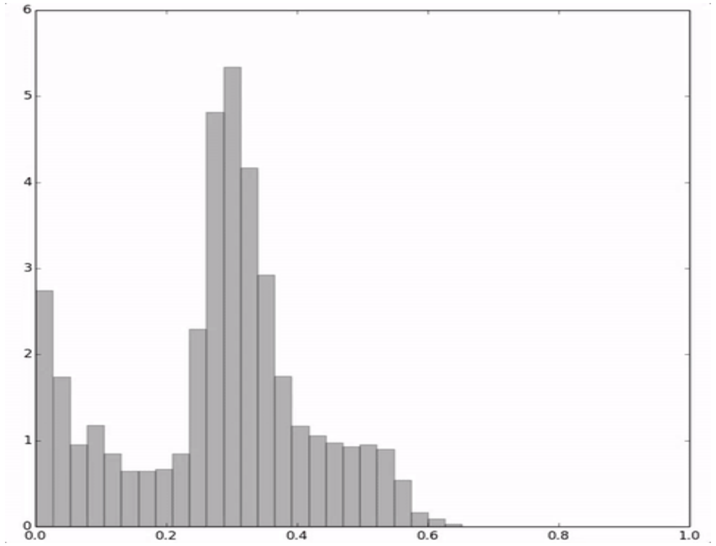


Tricks …

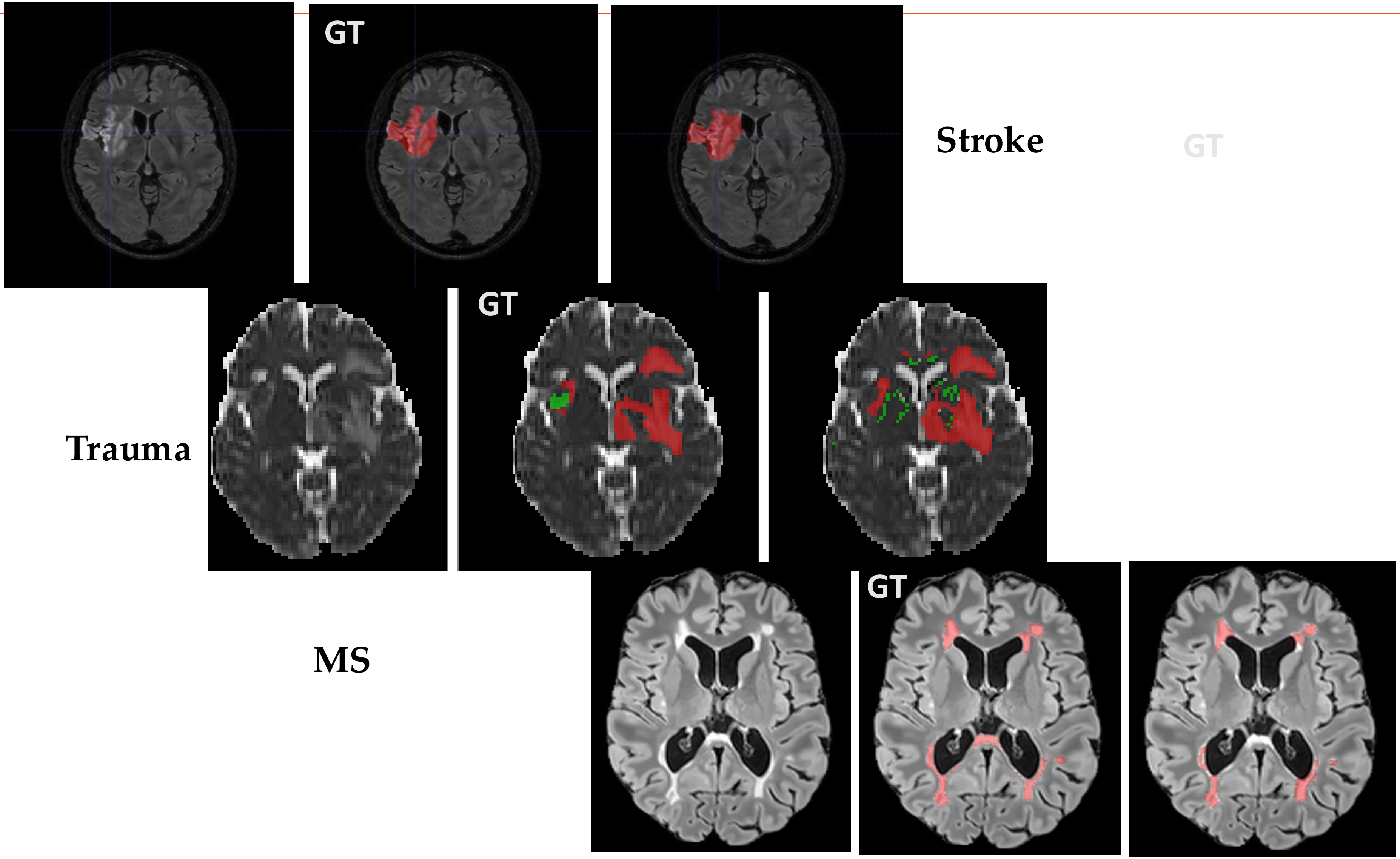
- Joint segmentation-registration
  - Robust atlas realignment



- Variational EM Approach
  - Iterative refinement of the lesion classes



Genericity: Pathological images



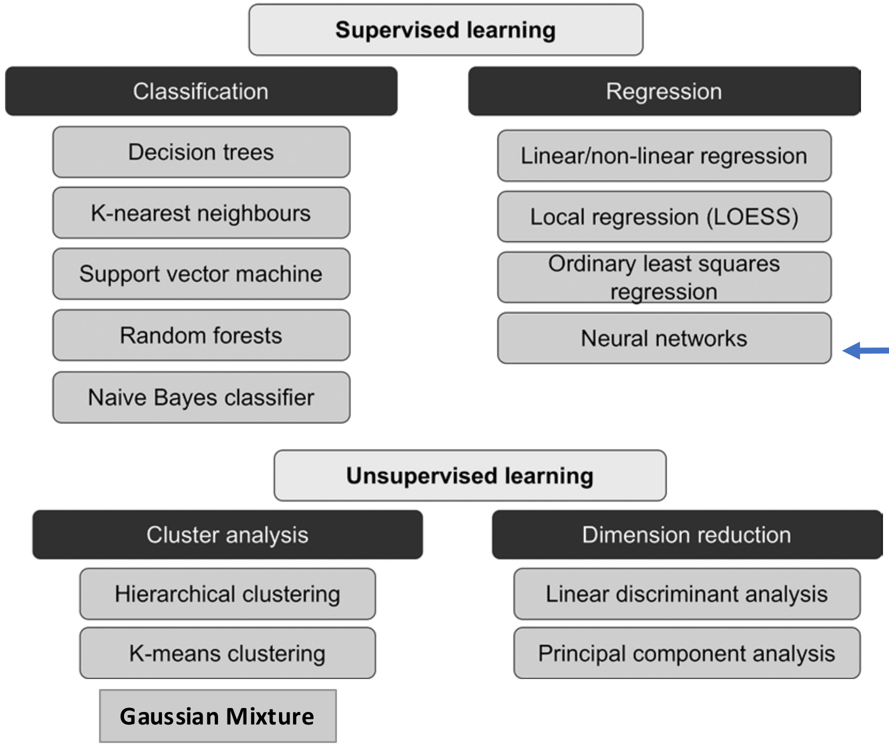
BUT …

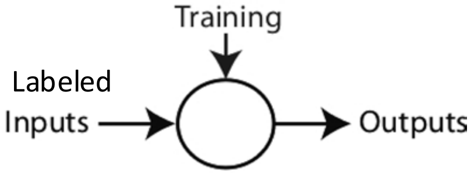
Machine learning supervised approaches

- ISLES 2017 challenge
    - Stroke : 14 participants, 14 NN ….
    - Trauma: 7 participants, 5 NN
    - Tumor: 22 participants, 19 NN, 1 SVM, 2 RF

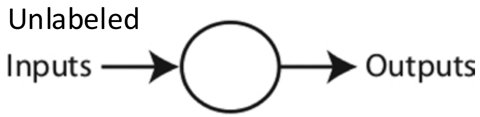
- MSSEG-2021 30 pipelines, 29 NN

ML Approaches

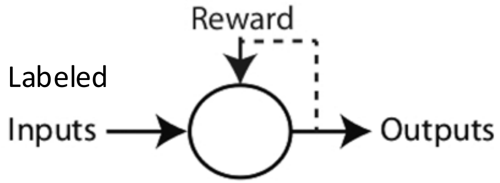




Learns known patterns
Predicts outcome



Learns unknown patterns
Find hidden patterns



Reinforcement learning

Generates data
Interacts with environment

Transfert learning al

Applies learned patterns to a different but related task

[From Choy et al. Radiology 2017]

International conference on

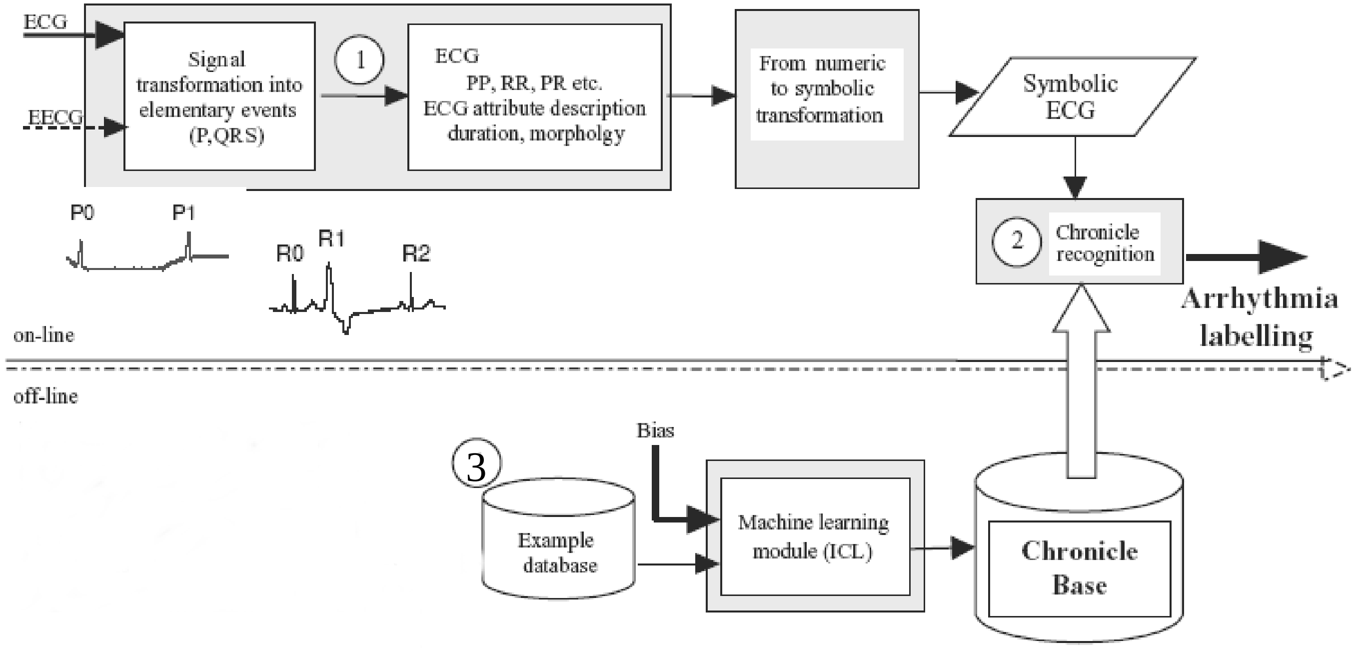
Medical Imaging with Deep Learning

Amsterdam, 4 – 6th July 2018

info@midl.amsterdam

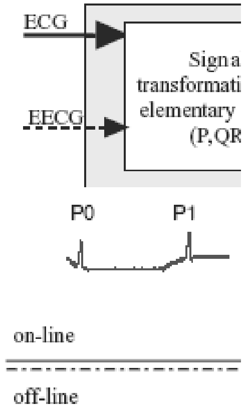
Cardia arrhythmia detection

- Symbolic learning



Cardia arrhythmia detection

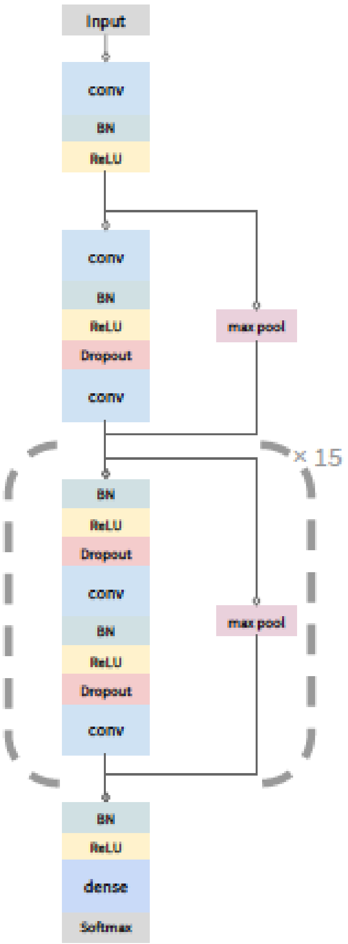
- Symbolic learning




```
class(bigeminy) :- % [15, 0, 0, 0, 0]
   qrs(R0, abnormal, _),
   p_wav(P1, normal, R0), qrs(R1, normal, P1),
   qrs(R2, abnormal, R1), rr1(R1, R2, short).
class(bigeminy) :- %[5, 0, 0, 0, 0]
   qrs(R0, normal, _),
   p_wav(P1, normal, R0), qrs(R1, abnormal, P1).
class(lbbb) :- % [0, 20, 0, 0, 0]
   qrs(R0, abnormal, _),
   p_wav(P1, normal, R0), qrs(R1, abnormal, P1).
class(mobitz2) :-% [0, 0, 17, 0, 0]
   p_wav(P0, normal, _), equal(P0, R0),
   p_wav(P1, normal, R0), qrs(R1, normal, P1).
class(mobitz2) :-,%[0, 0, 3, 0, 0]
   p_wav(P0, normal, _), equal(P0, R0),
   p_wav(P1, normal, R0), qrs(R1, abnormal, P1).
```

Carrault et al Artif Intell Med 2003

Cardiologist: Arrhythmia detection via CNN



64121 ECG from 29163 patients 30sec recordings

Training 90% of data, 10% validation

Test 336 records from 338 different patients

12 arrhythmias

	Seq		Set	
	Model	Cardiol.	Model	Cardiol.
Class-level F1 Score				
AFIB	**0.604**	0.515	**0.667**	0.544
AFL	**0.687**	0.635	**0.679**	0.646
AVB_TYPE2	**0.689**	0.535	**0.656**	0.529
BIGEMINY	**0.897**	0.837	**0.870**	0.849
CHB	**0.843**	0.701	**0.852**	0.685
EAR	**0.519**	0.476	**0.571**	0.529
IVR	**0.761**	0.632	**0.774**	0.720
JUNCTIONAL	0.670	**0.684**	**0.783**	0.674
NOISE	**0.823**	0.768	**0.704**	0.689
SINUS	**0.879**	0.847	**0.939**	0.907
SVT	**0.477**	0.449	**0.658**	0.556
TRIGEMINY	**0.908**	0.843	**0.870**	0.816
VT	0.506	**0.566**	0.694	**0.769**
WENCKEBACH	**0.709**	0.593	**0.806**	0.736
Aggregate Results				
Precision (PPV)	**0.800**	0.723	**0.809**	0.763
Recall (Sensitivity)	**0.784**	0.724	**0.827**	0.744
F1	**0.776**	0.719	**0.809**	0.751

32 layers + softMax
Filter: 64*16, 128*16, …

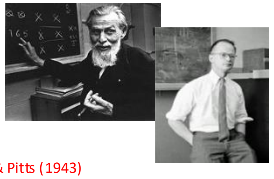
[Rajpurkar et al 2017]

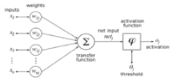
A nice bioinspired story

NS => AI

bioinspired

1940-1970



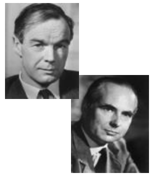


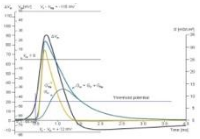
McCulloch & Pitts (1943)
Neurology & Psychology
The first formal neuron

Donald Hebb (1949)
Neuropsychology
Learning=Synaptic modification

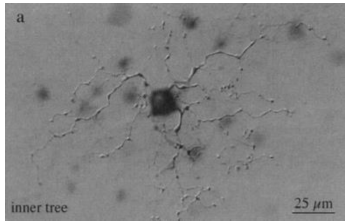
$$w_{ij} = \frac{1}{p}\sum_{k=1}^{p} x_i^k x_j^k,$$

Hodgkin, A. L., & Huxley, A. F. (1952)
Neuroscientists
Temporal dynamic in synapses modification





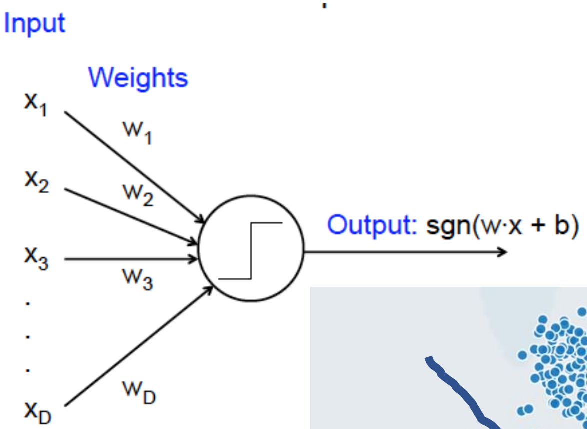
$$i_m = i_{mI} + c_m \frac{\partial V_m}{\partial t} = \frac{1}{r_i + r_o}\frac{\partial^2 V_m}{\partial x^2}$$

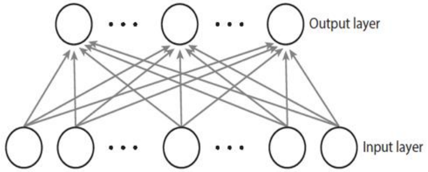


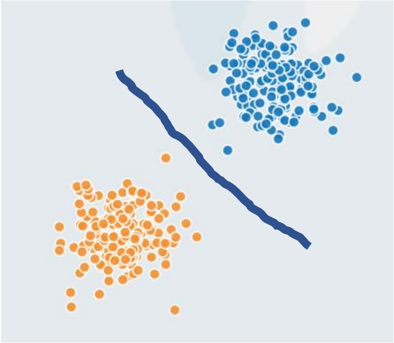
Frank Rosenblatt – Psychology

The Perceptron, the first Artificial Neural Network

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.



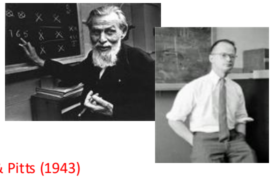


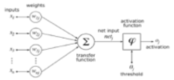


A nice bioinspired story… The pioneers

1940-1970

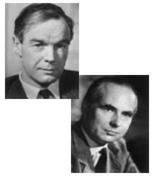
Donald Hebb (1949)
Neuropsychology
Learning=Synaptic modification

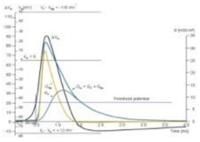




McCulloch & Pitts (1943)
Neurology & Psychology
The first formal neuron

$$w_{ij} = \frac{1}{p}\sum_{k=1}^{p} x_i^k x_j^k,$$



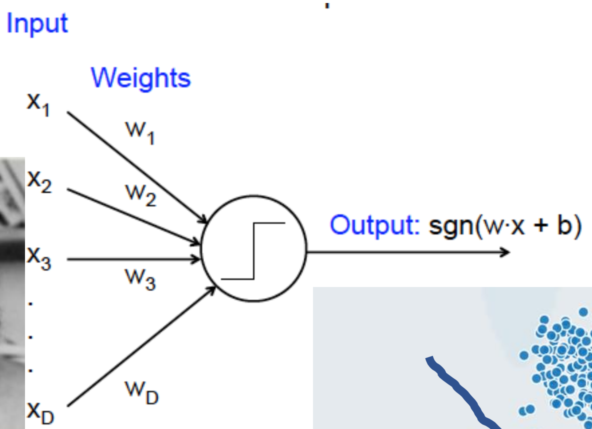


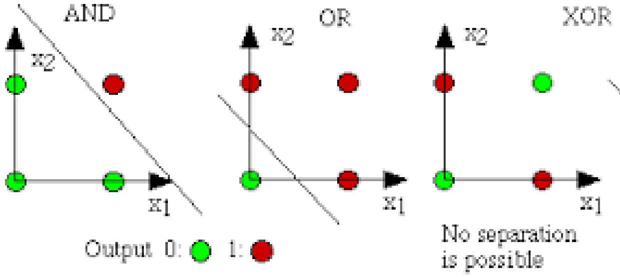
$$i_m = i_{ml} + c_m \frac{\partial V_m}{\partial t} = \frac{1}{r_i + r_o}\frac{\partial^2 V_m}{\partial x^2}$$

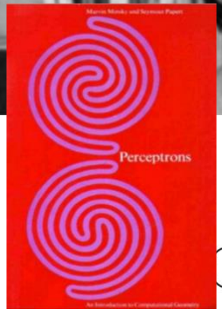
Hodgkin, A. L., & Huxley, A. F. (1952)
Neuroscientists
Temporal dynamic in synapses modification

Frank Rosenblatt – Psychology

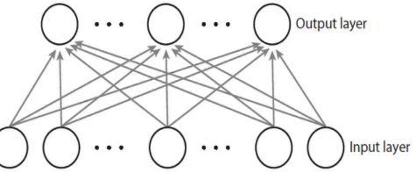
The Perceptron, the first Artificial Neural Network
Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

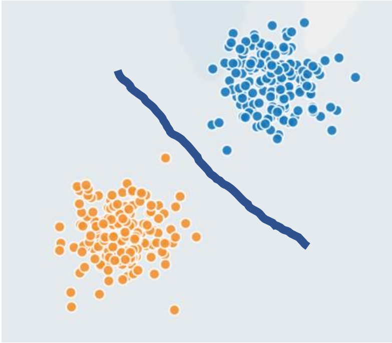






M. Minsky & S Papert (1967)
Perceptron



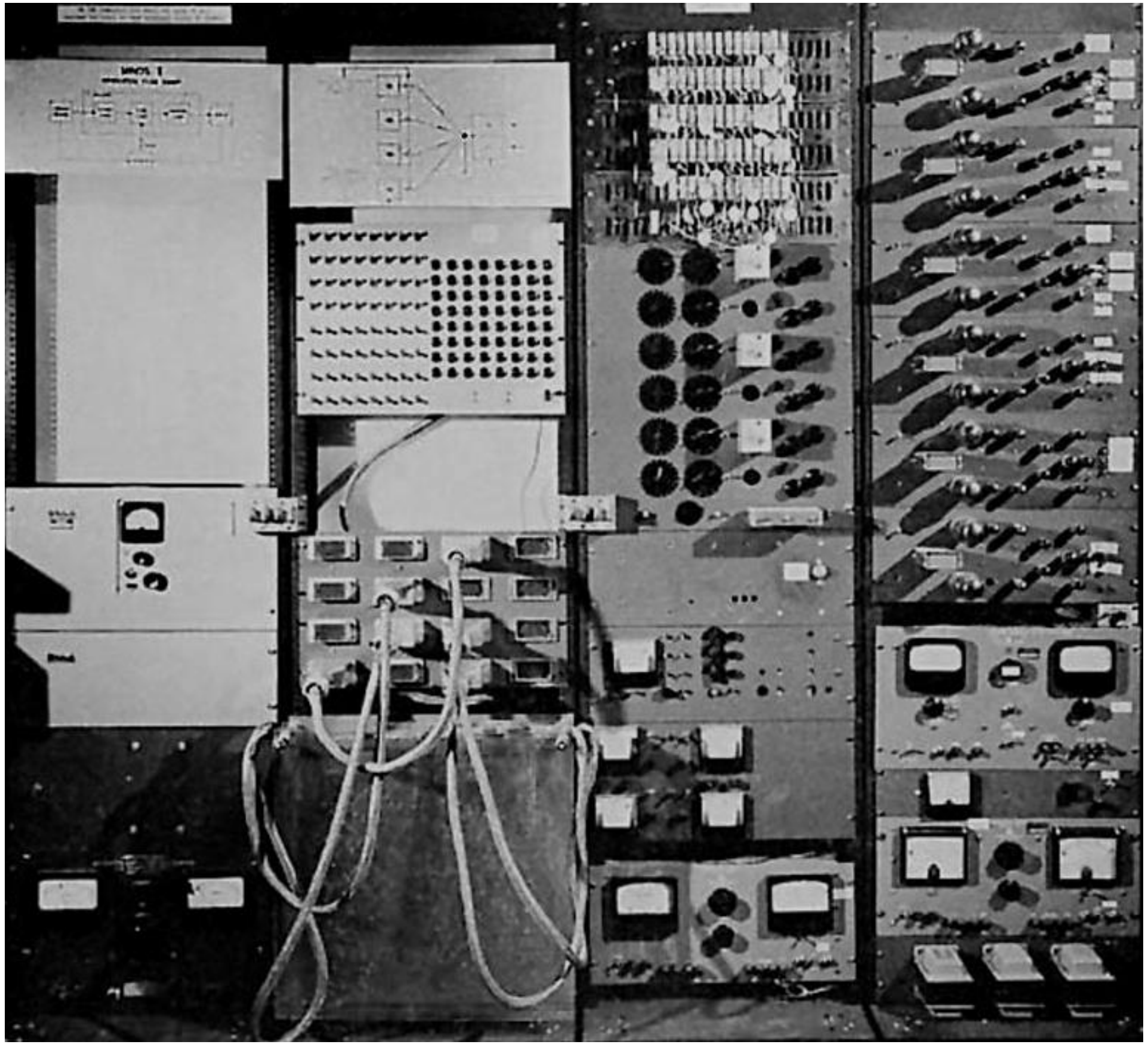


The fall of Perceptron

A nice bioinspired story… The pioneers

Minos

1960
Stanford
Research
Institute



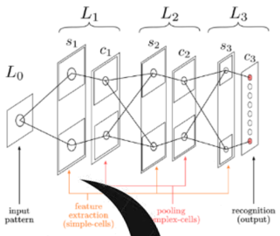
A nice bioinspired story… Renewal

1980-….

Kunihiko Fukushima- Bio-inspired Computer Science

Fukushima, K. (1980). Neocognitron: A self-organizing neural network for a mechanism of pattern recognition unaffected by shift in position. *Bio Cyb 46-193-202*

Filtre convolutionnel



Deep Learning: the direct offspring of the Multi-Layer Perceptron

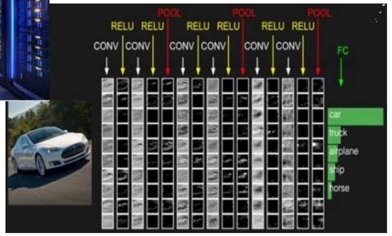
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436.

Turing Prize 2018

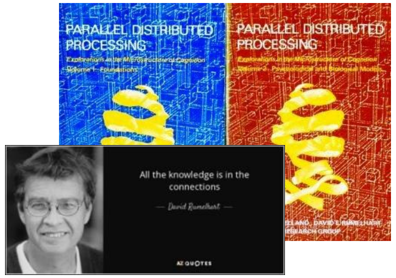
Yann LeCun
Computer Science

Yoshua Bengio Computer Science

Geoffrey Hinton
Cognitive Psychology & Computer Science



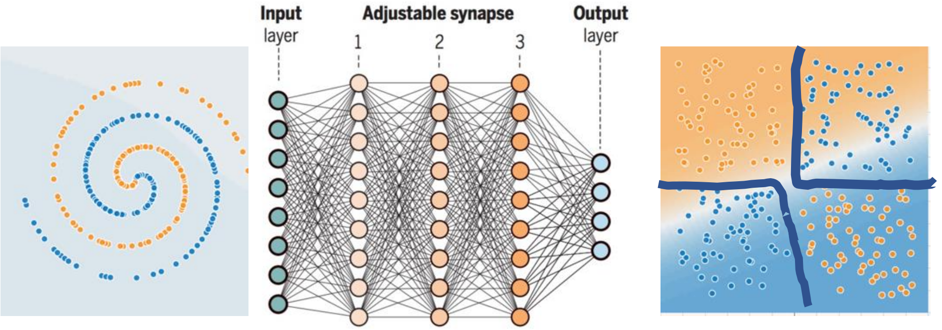
Beyond the Perceptron: **the Multi-Layer Perceptron**

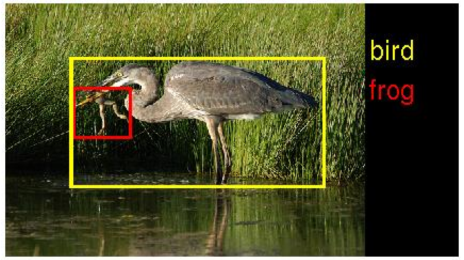


BackProp algo

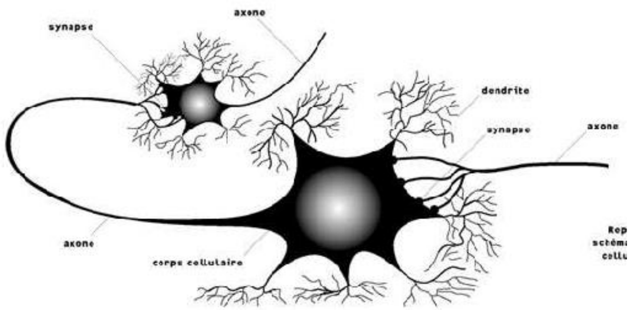
David Everett Rumelhart - Cognitive Psychology
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536.

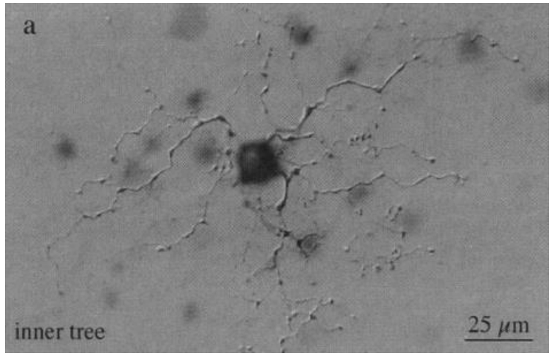
Performance in computer visiobn



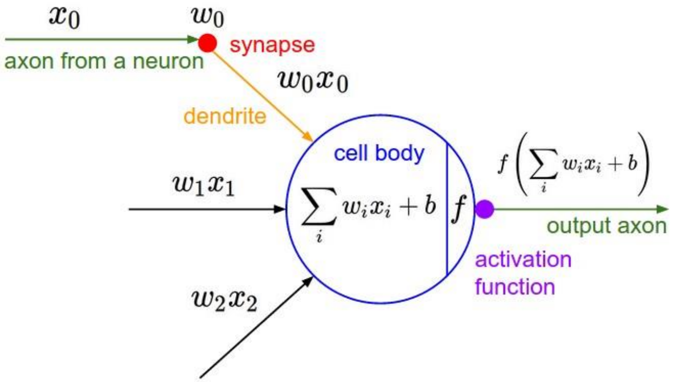


Neural Networks





Artificial neurons (bio-inspired)

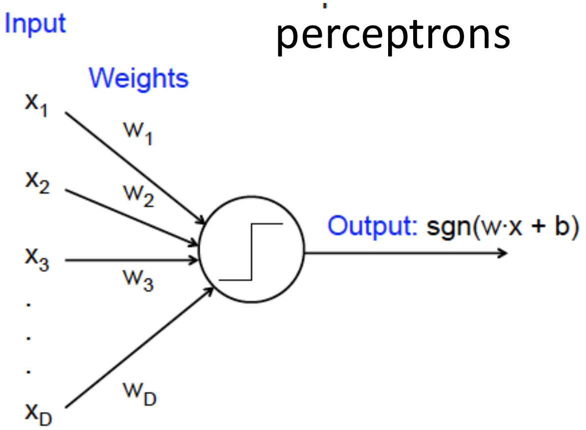


$$y(\mathbf{x}, \mathbf{w}) = f\left(\sum_{j=0}^{M} w_j \phi_j(\mathbf{x})\right)$$

x: obervations, w: weights, y: output
ϕ: models basis or activation functions

$$t(\mathbf{x}) = y(\mathbf{x}, \mathbf{w}) + \epsilon(\mathbf{x})$$

Neural Networks

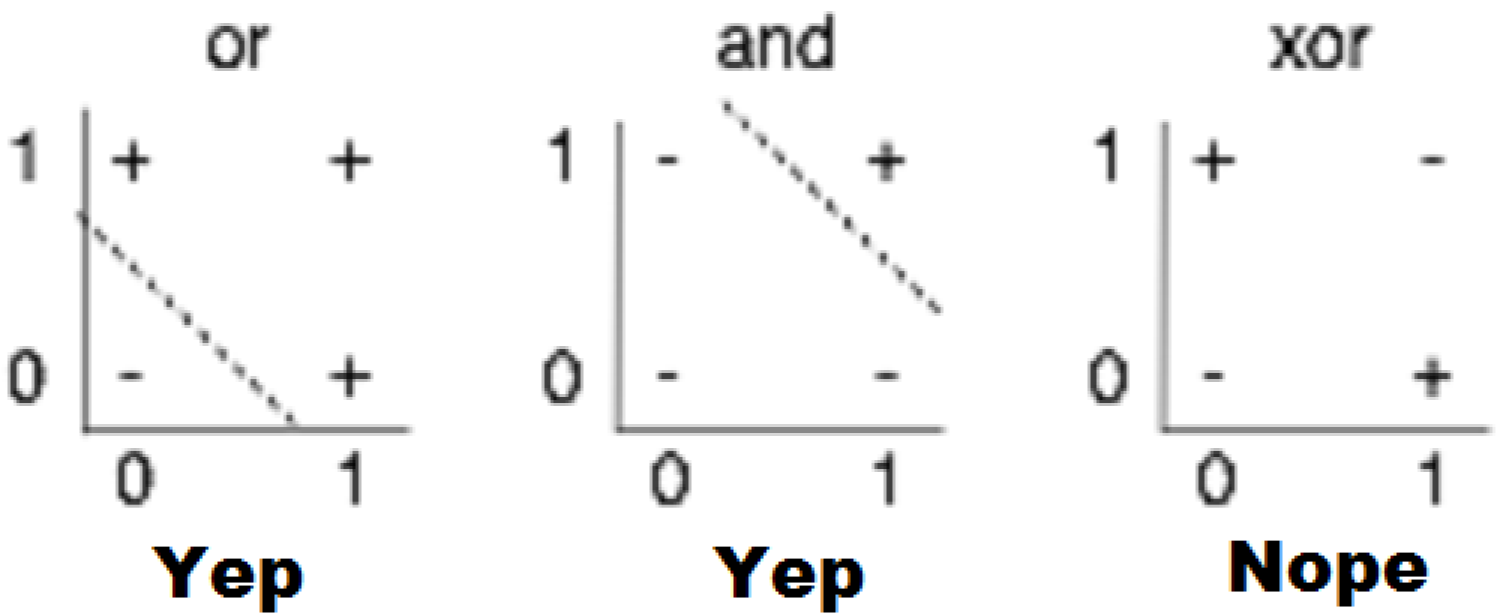


perceptrons

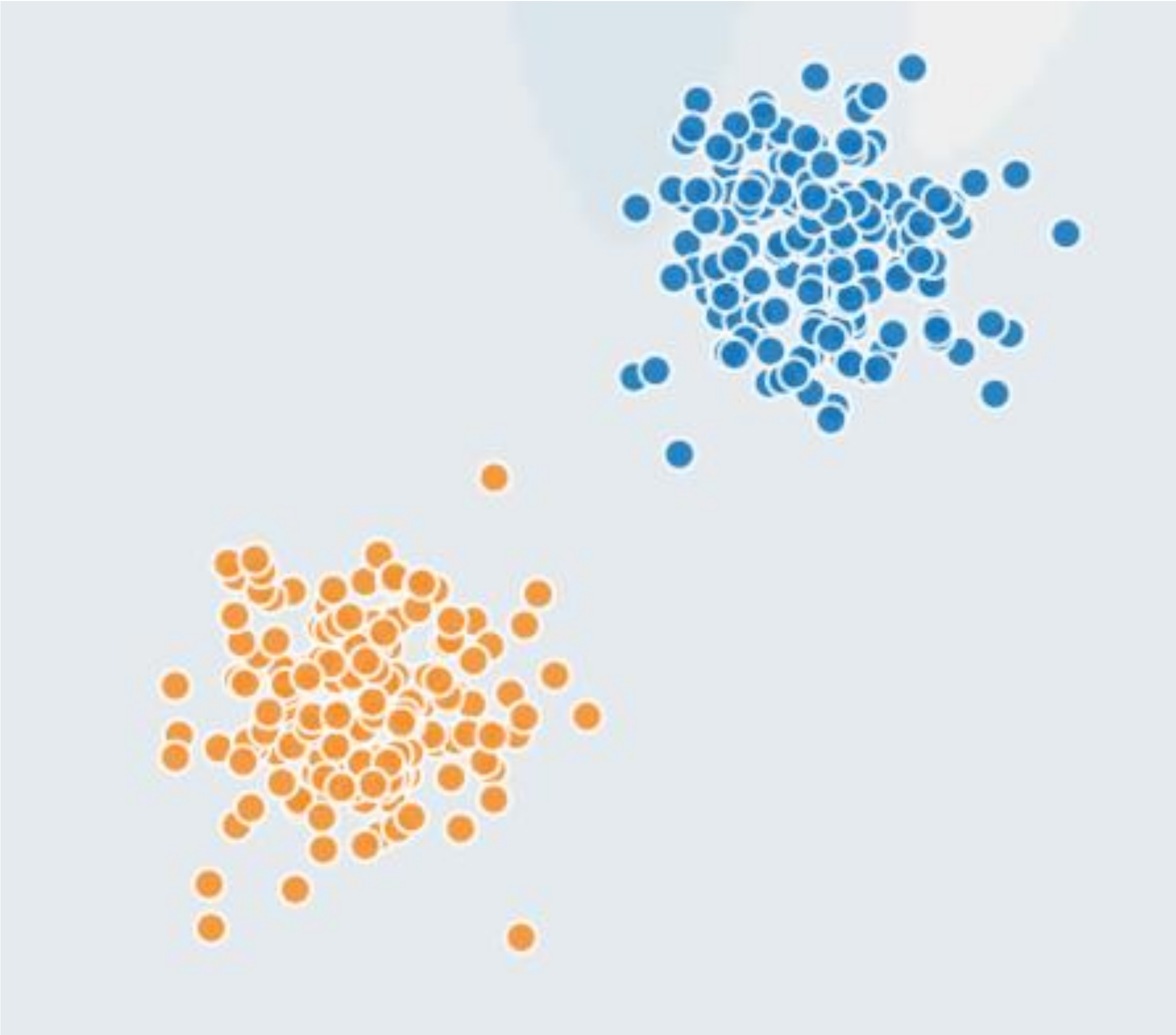
Rosenblatt 1958 Psych Rev.

- Perceptron : no hidden layers
  only linearly separable function.
  Convergence theorem for the learning rule.

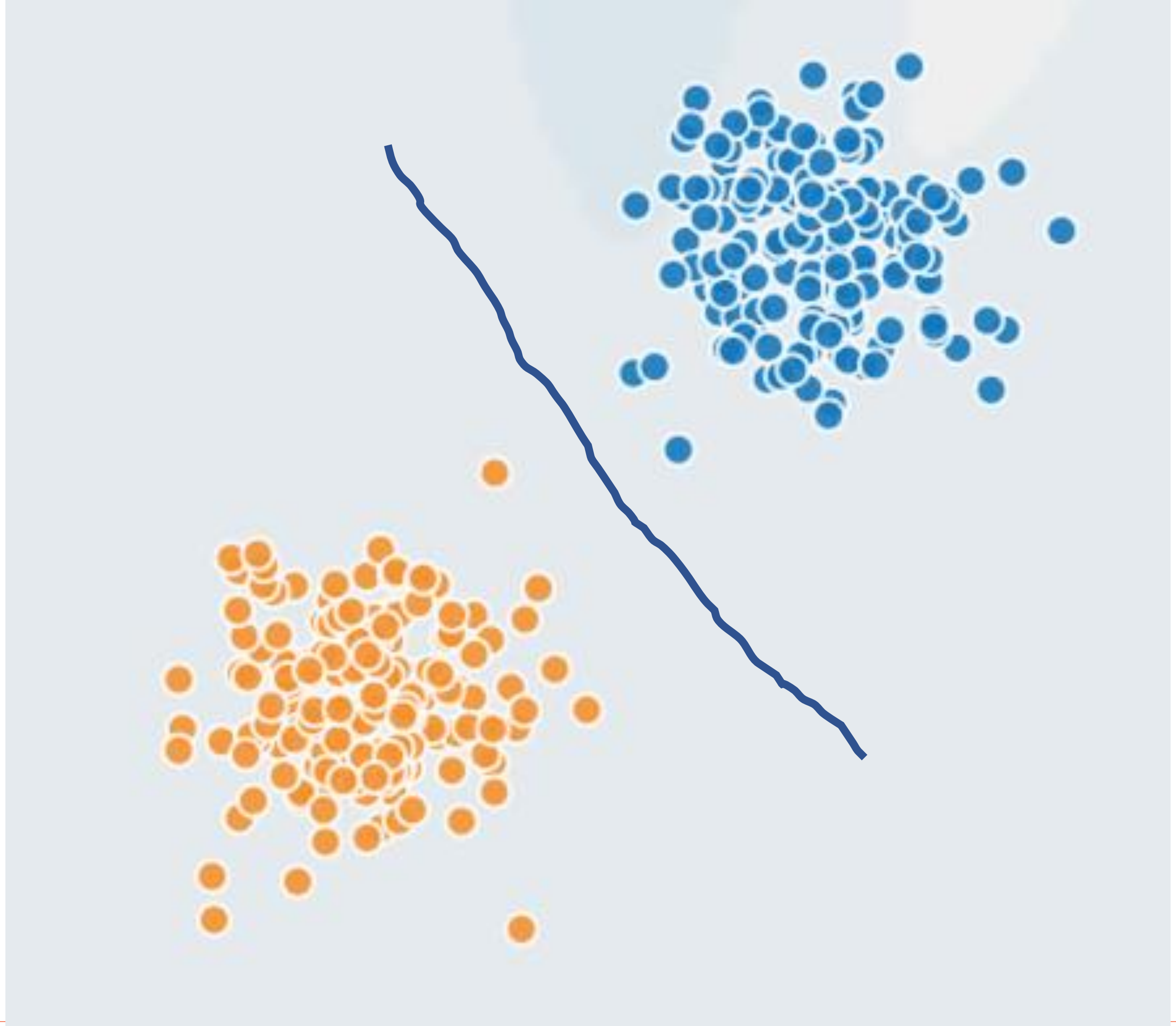
Perceptron limitations



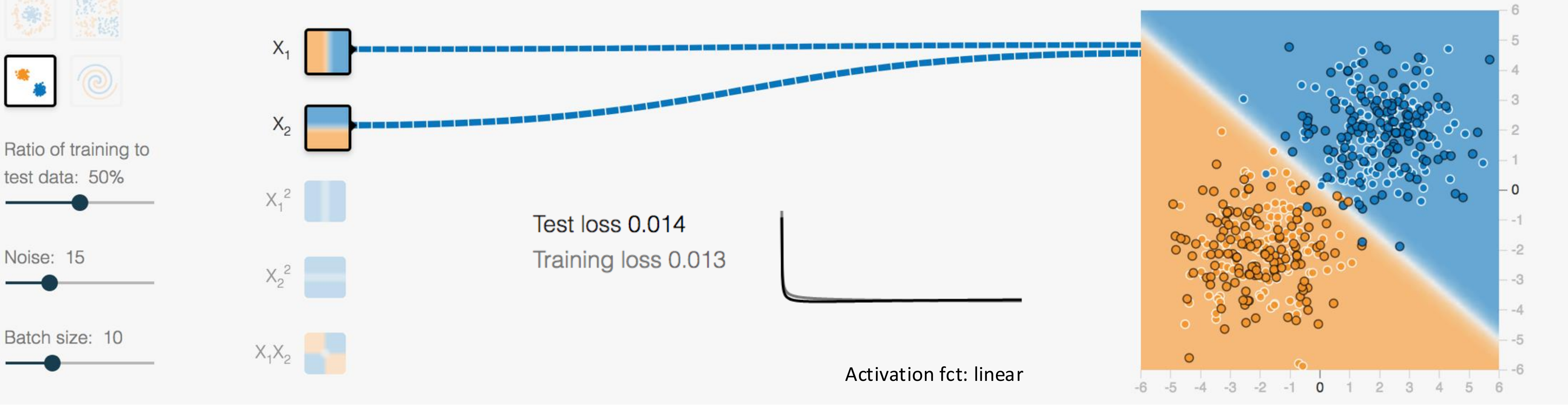
Perceptron



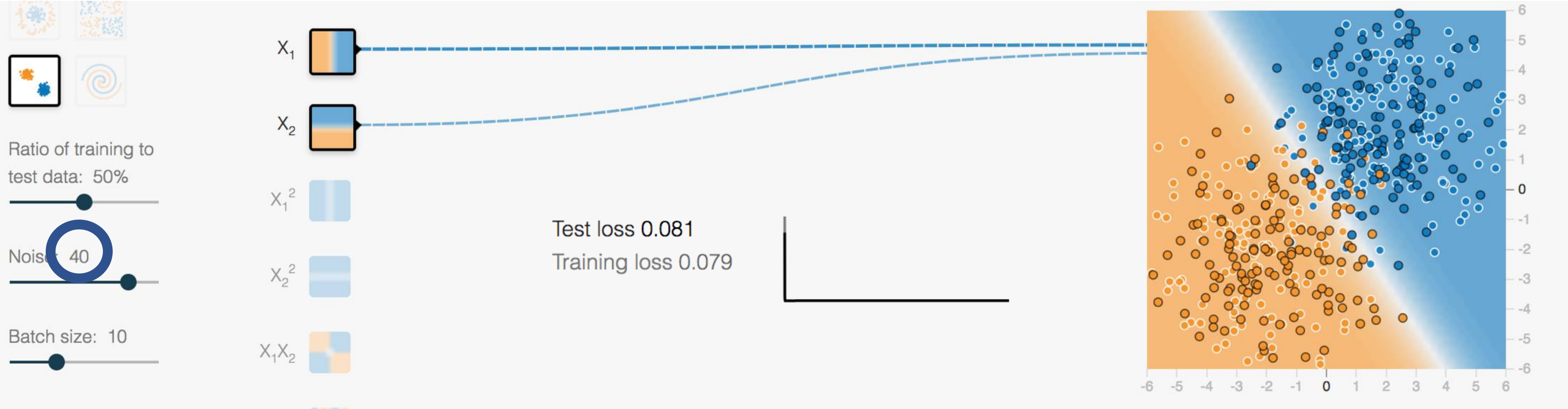
Perceptron



Perceptron

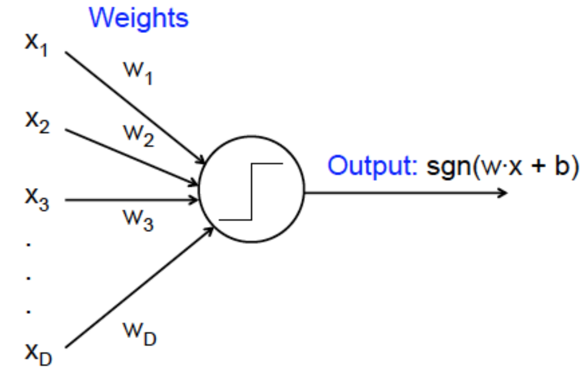


http://playground.tensorflow.org/



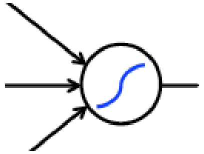
Neural Networks

perceptrons

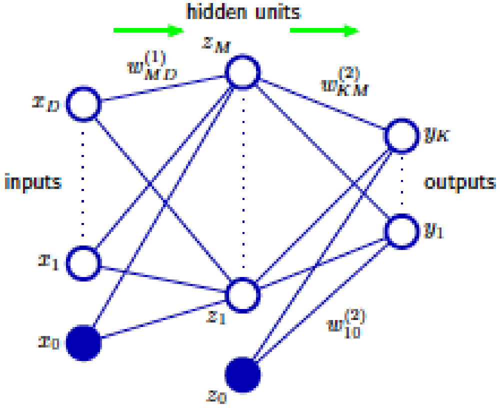


Rosenblatt 1958 Psych Rev.

- Perceptron : no hidden layers
  only linearly separable function.
  Convergence theorem for the learning rule.

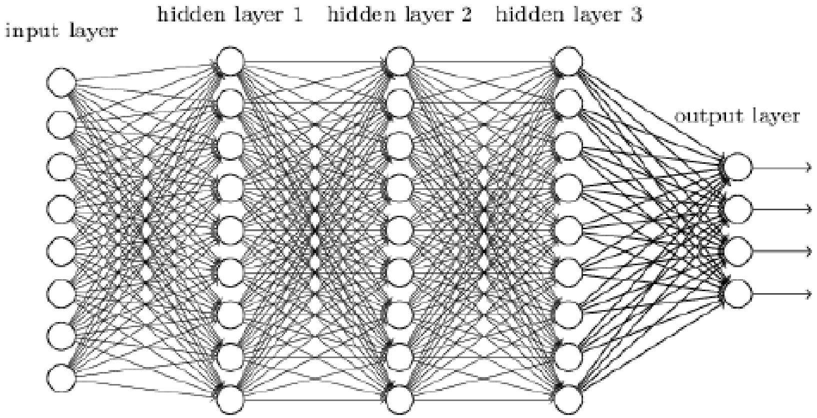


$$g(t) = \frac{1}{1+e^{-t}}$$

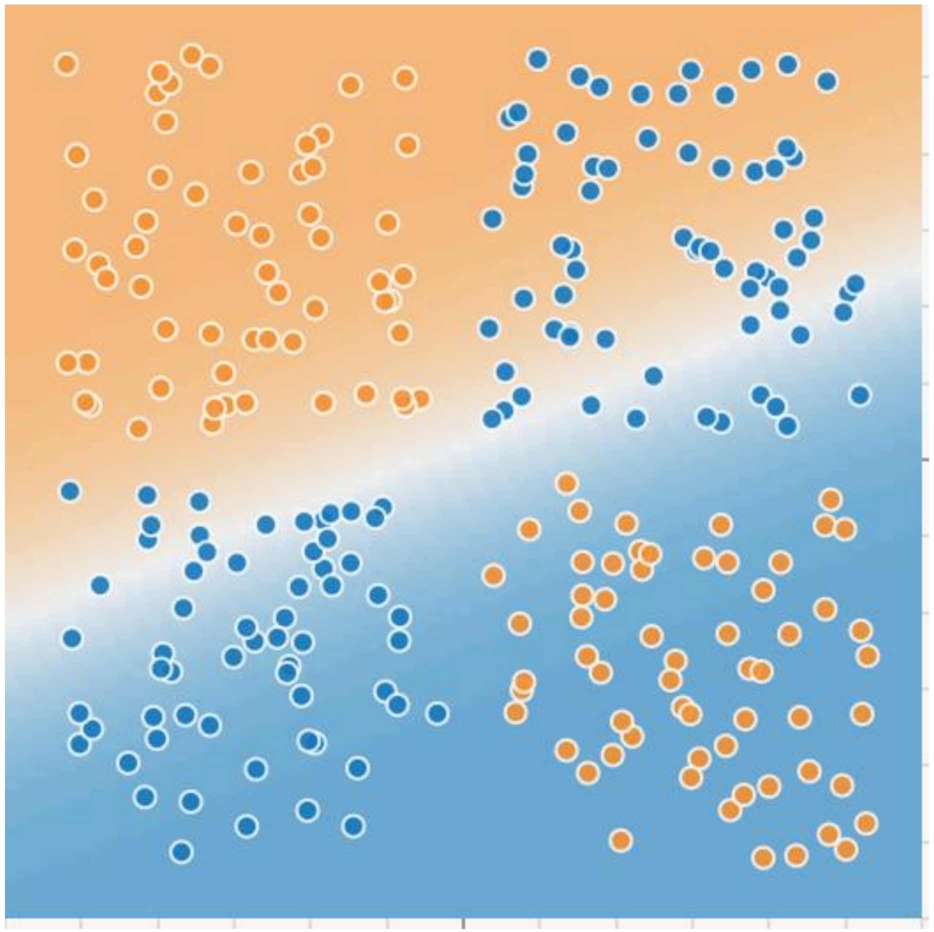


Khalidov & Hansard

Multilayers architectures



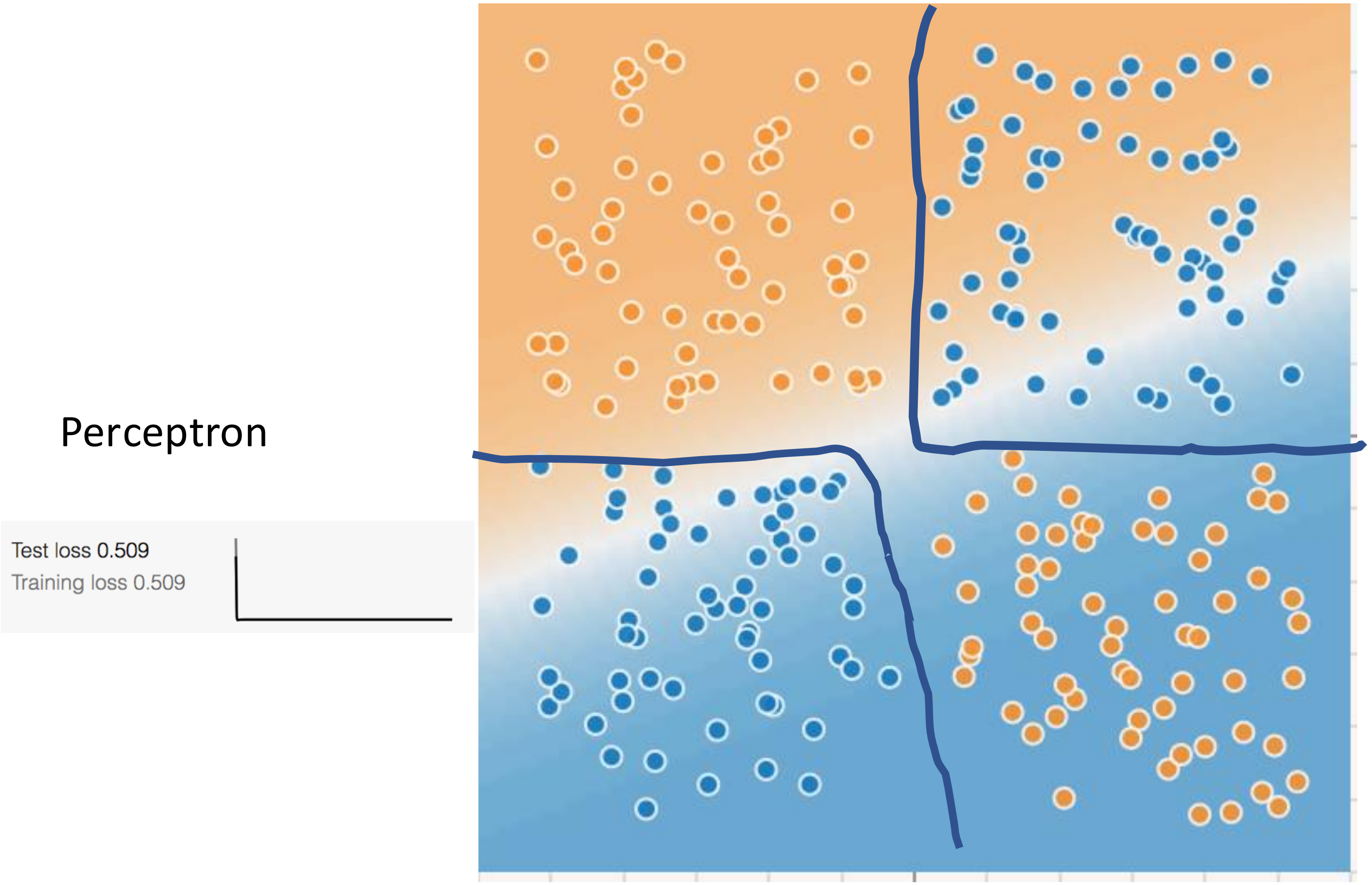
Multi-layer network



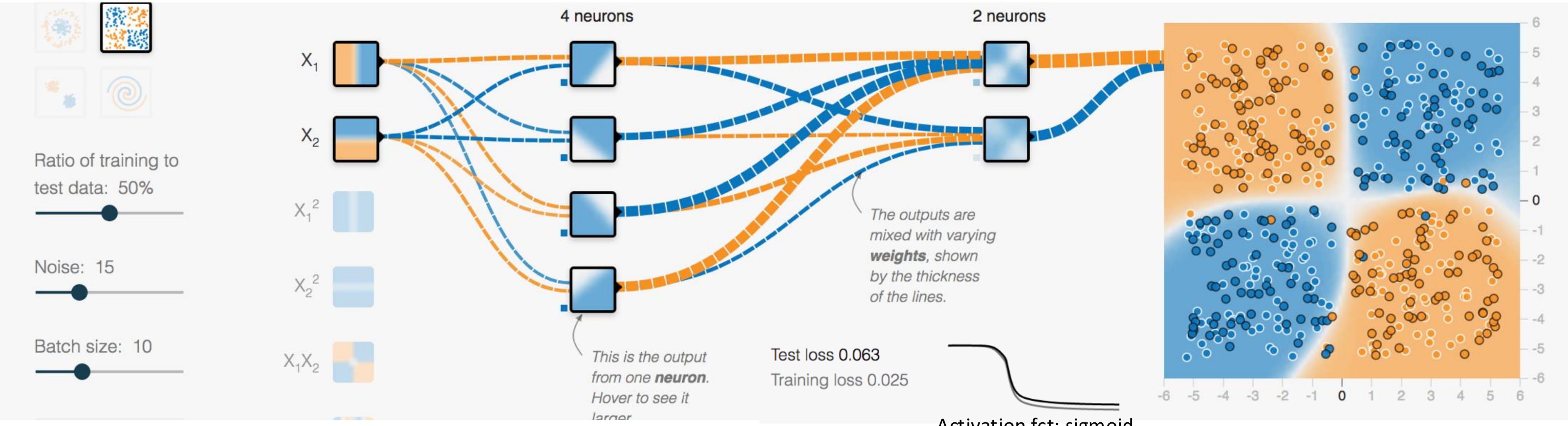
Multi-layer network

- Hidden layers required for non linear separation

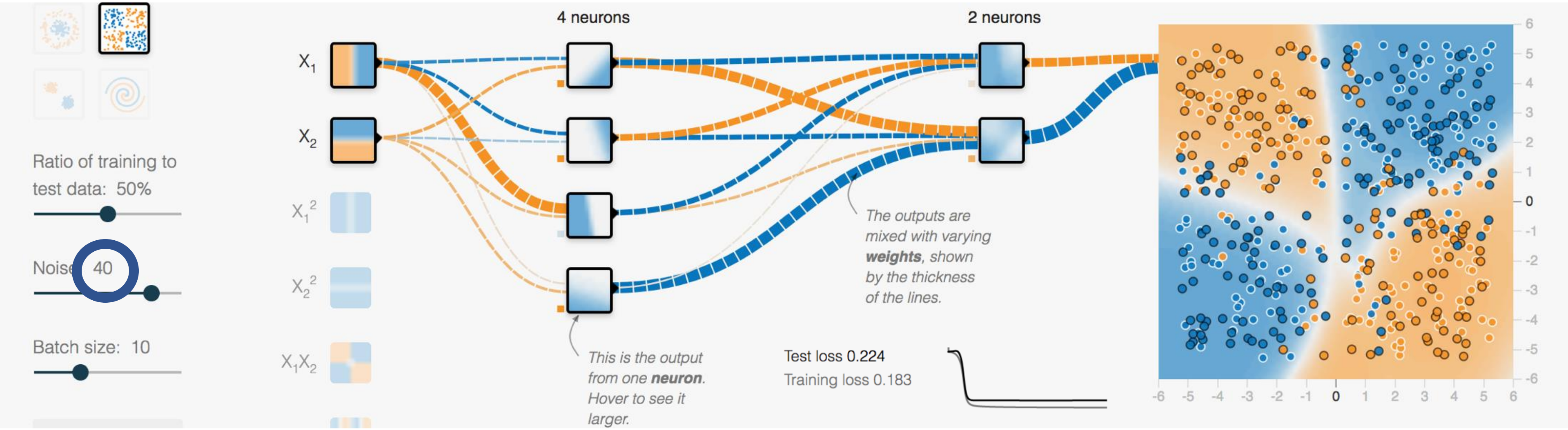
Perceptron



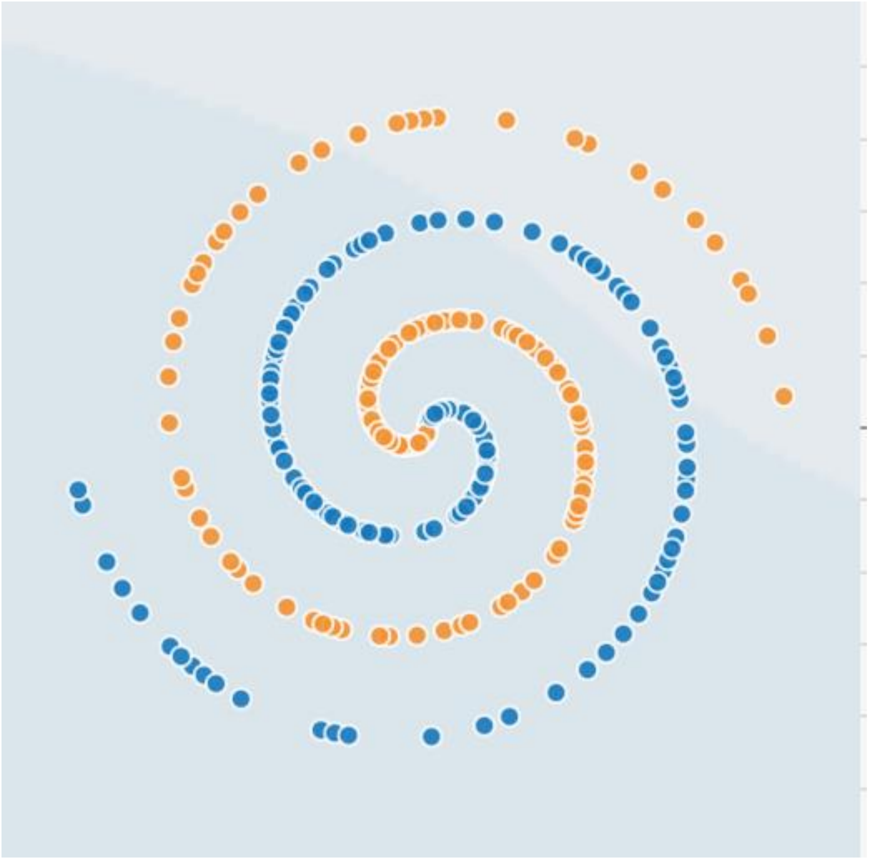
Multi-layer network



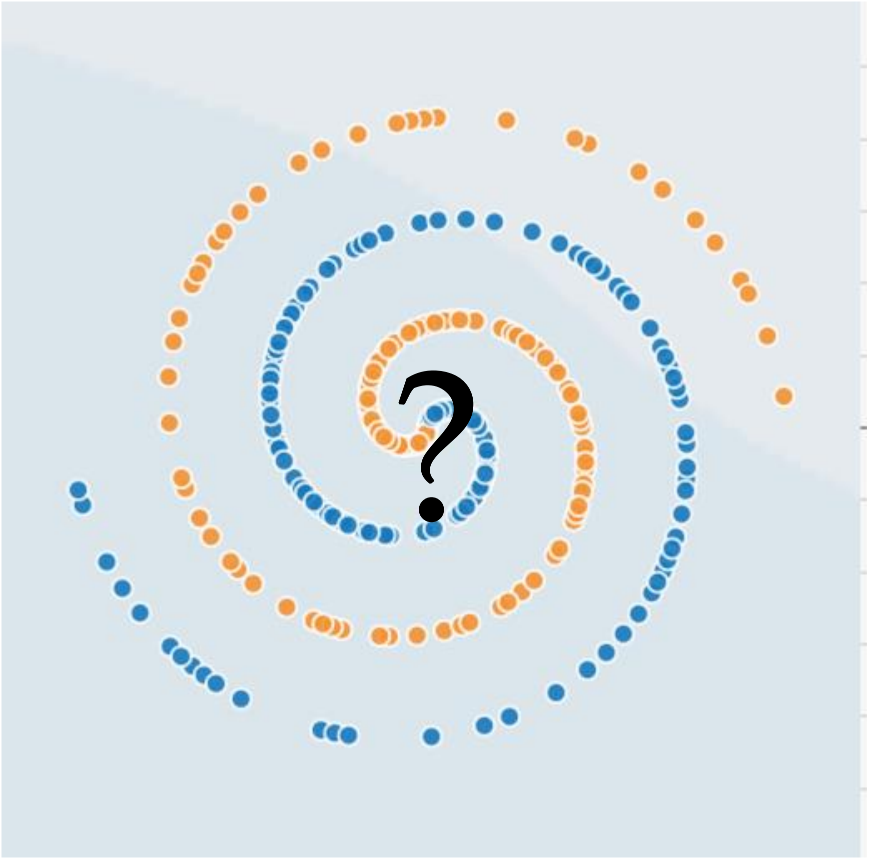
Activation fct: sigmoid



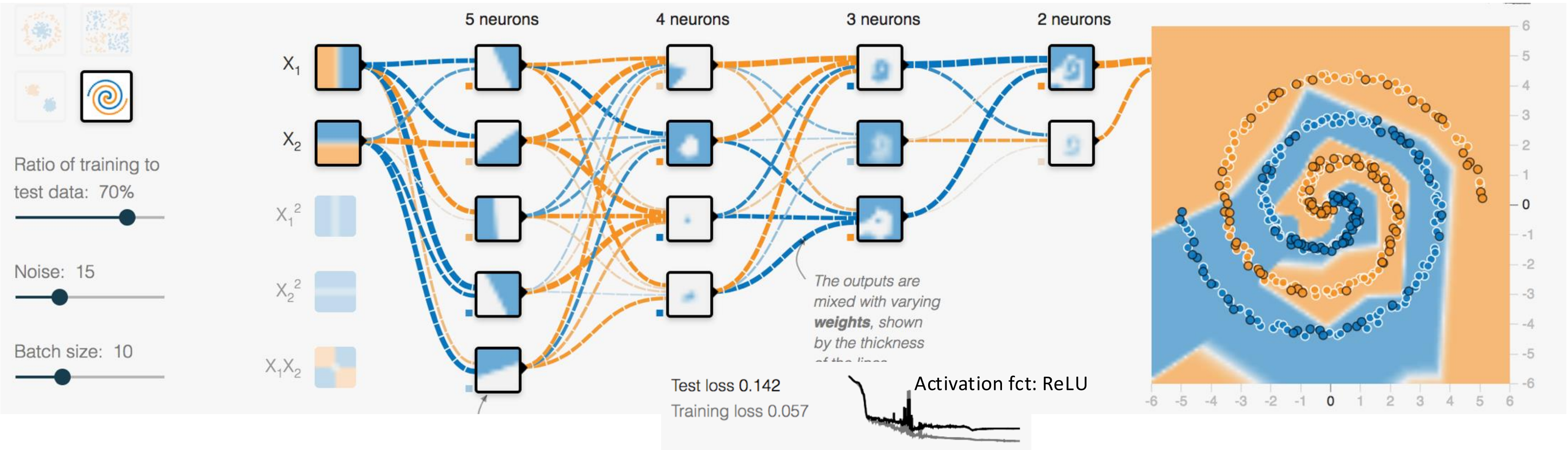
Multi-layer network

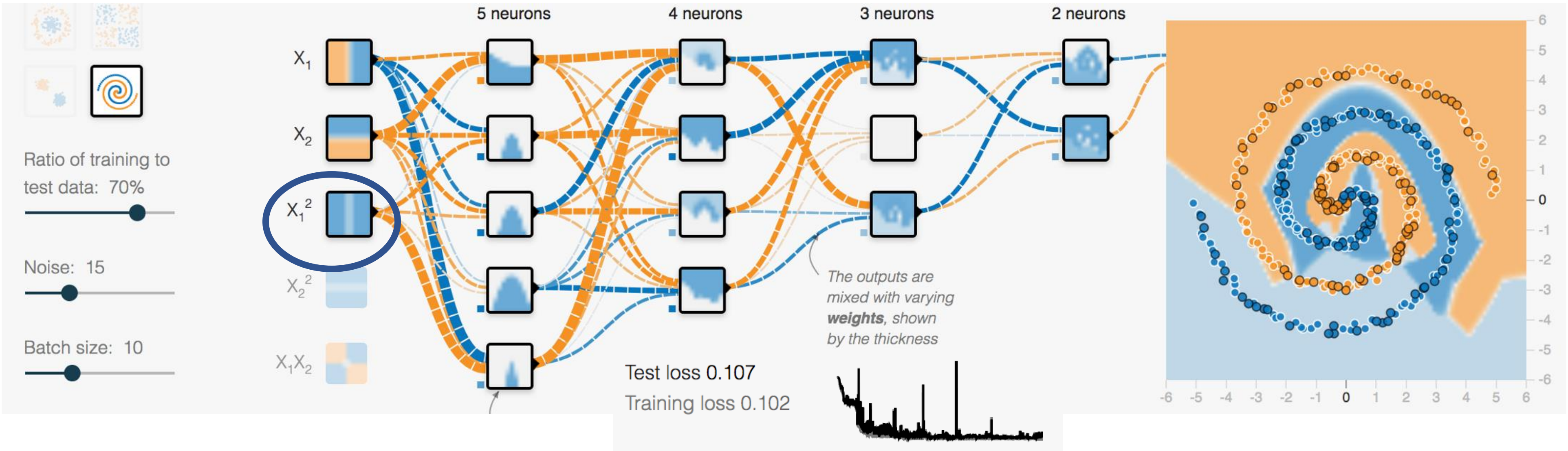


Multi-layer network



Multi-layer network





Multi-layer network