Machine Learning in Healthcare

#C01 Introduction to ML in Healthcare

Technion-IIT, Haifa, Israel

Assist. Prof. Joachim Behar
Biomedical Engineering Faculty
Technion-IIT

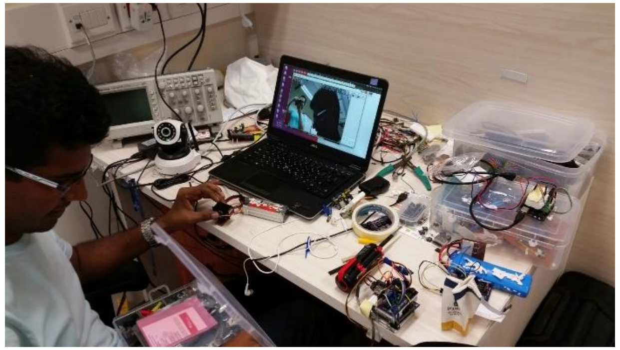
AIMLab.

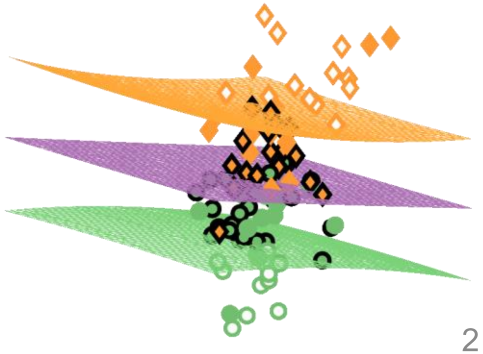
https://aim-lab.github.io/

<u>Scope:</u> Researches innovative **pattern recognition algorithms** to exploit the information encrypted within large datasets of physiological time series. AIMLab. leverages these new data-driven algorithms toward the creation of novel intelligent **remote patient monitoring systems**.

<u>Lab Research:</u>

- Sleep Medicine
- Fetal Medicine.



Course Structure

Course Structure - Overview

- Lectures – **Joachim Behar** (jbehar@technion.ac.il)
  - Intuitions and mathematical background.
- Laboratory sessions – **Alon Begin** (alonbegin24@gmail.com)
  - Hands on computer.
  - Open desk: 18-18:30 and emails.
- Workshops, assignments and Challenge – **Moran Davoodi** (morandavoodi@gmail.com).
  - Open desk: before/after the tutorial sessions.

Anne Weill, PhD

Technion, BME

Doron Shaked, PhD

GE Healthcare

Danny Eytan, MD-PhD

Rambam Hospital

Uri Shalit, PhD

Technion, Industrial Eng.

Kamal Masarweh, MD

Rambam Hospital

Instructor

Joachim Behar

Teaching Assistants

Alon Begin

Moran Davoodi

Time and Location

Tuesday 12:30 - 14:30
Room 202

Tuesday 18:30 - 20:30
Lab 121

Course Aims

- You will learn:
  - Python for data science.
  - Structuring machine learning projects.
  - ML main concepts and familiarity with main classifiers and theory.
  - Deep Learning.
  - ML in healthcare.
- Prerequisites:
  - Algebra.
  - Introduction to probability.
  - Coding skills (although you’re not expected to know Python.)
  - Signals and Systems.

Course Structure

- Three modules, 4-weeks each.

ML basis
Weeks 1-4

Popular Classifiers
Weeks 5-8

Deep Learning
Weeks 9-13

- One assignment for each part (3 in totals).
- Technion-MLH Challenge #1.

Course Evaluation

- Grading
  - Three assignments (A1 20%, A2 30%, A3 30%),
  - Challenge (20%),
  - Course (lectures and TA sessions) attendance > 90% (bonus of 5%).
- Important deadline
  - A1: W06
  - A2: W10
  - Challenge: first submission on test set allowed in W8.
  - Challenge: second submission on the test set and final report allowed **before** W12.
  - A3: W16 (i.e. up to three weeks after the official end of the course.)
- Penalties
  - Plagiarism ➔ 0%.
  - Late submission ➔ -25% per week passed the deadline (cumulative every week.)

Important: *assignments and Challenge source codes are submitted through GitHub. Reports are submitted in Word documents to Moran.*

Assignments

- Three assignments:

Assignments	Dataset type	Number of instances (n) and attributes (p)	Learning objectives	Deadlines
#A1 Cardiotography Data Set	The dataset consists of measurements of fetal heart rate (FHR) and uterine contraction (UC) features on cardiotocograms labelled by expert obstetricians. The aim is to automate the analysis of the FHR and approach the obstetricians' labels.	n=2126 p=23	Data exploration. Features standardization. Linear Classification.	**W06**
#A2 ICU Mortality prediction	Features derived from physiological time series and demographics for the purpose of predicting what patients will die in the intensive care unit (ICU).	n=3000 p=35	Non-linear classifiers.	**W10**
#A3 X-ray	Medical images classification.	n=400	Deep Learning	**“W16”**

AIMLab.

Technion-MLH Challenge

The consequences of air pollution in Haifa on the severity of admissions at the Rambam Children's hospital respiratory unit. Retrospective analysis 2001-to date.

The challenge will be presented in W2 by KM from Rambam and we will explain the rules of the competition. Submission by W12.

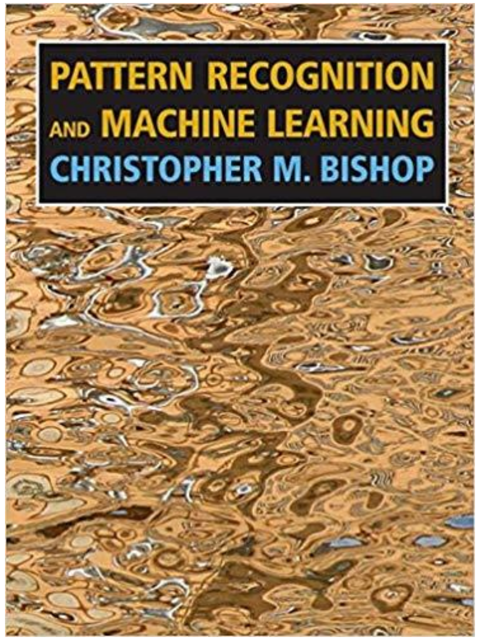
Workshops

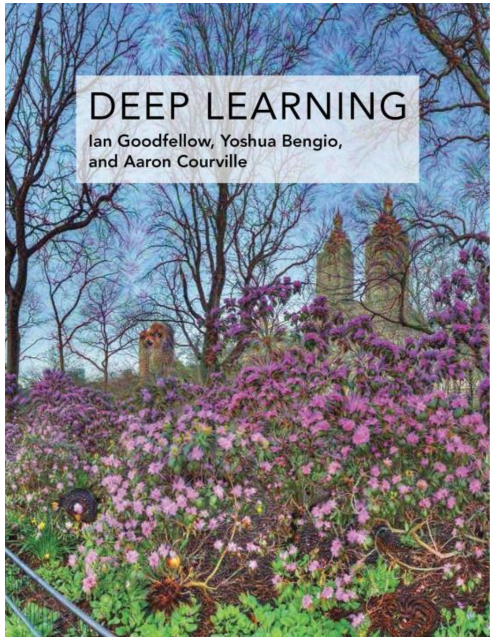
- Three workshops during the semesters. These are optional but we highly advise you to attend.

Workshops	Subjects covered	Date
#W1	Environment, Git, GitHub and numpy.	**W01**
#W2	Supervised Learning.	**W06**
#W3	Deep Learning.	**W10**

Textbooks

- Recommended but not required for completing the course:
  - Bishop, Christopher M. Pattern recognition and machine learning. springer, 2006.
  - Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.





Other resources

- List of other recommended resources:
  - Machine Learning Mastery: https://machinelearningmastery.com/
  - Toward Data Science: https://towardsdatascience.com/
  - Kaggle: https://www.kaggle.com/
  - Google: https://developers.google.com/machine-learning/crash-course/ml-intro

Introduction to ML

“The fourth industrial revolution”, but why?

$$\Psi(1s) = 2a_0^{-1.5}e^{\frac{-r}{a_0}}$$

$$\Psi(2s) = \frac{1}{\sqrt{8}}a_0^{-1.5}(2-\frac{r}{a_0})e^{\frac{-r}{2a_0}}$$

$$\Psi(2p) = \frac{1}{\sqrt{24}}a_0^{-1.5}(\frac{r}{a_0})e^{\frac{-r}{2a_0}}$$

Computer → “easy one!”

Human → “Give me a couple of hours…”

Computer → “not sure…”

Human → “It’s obviously a cat”

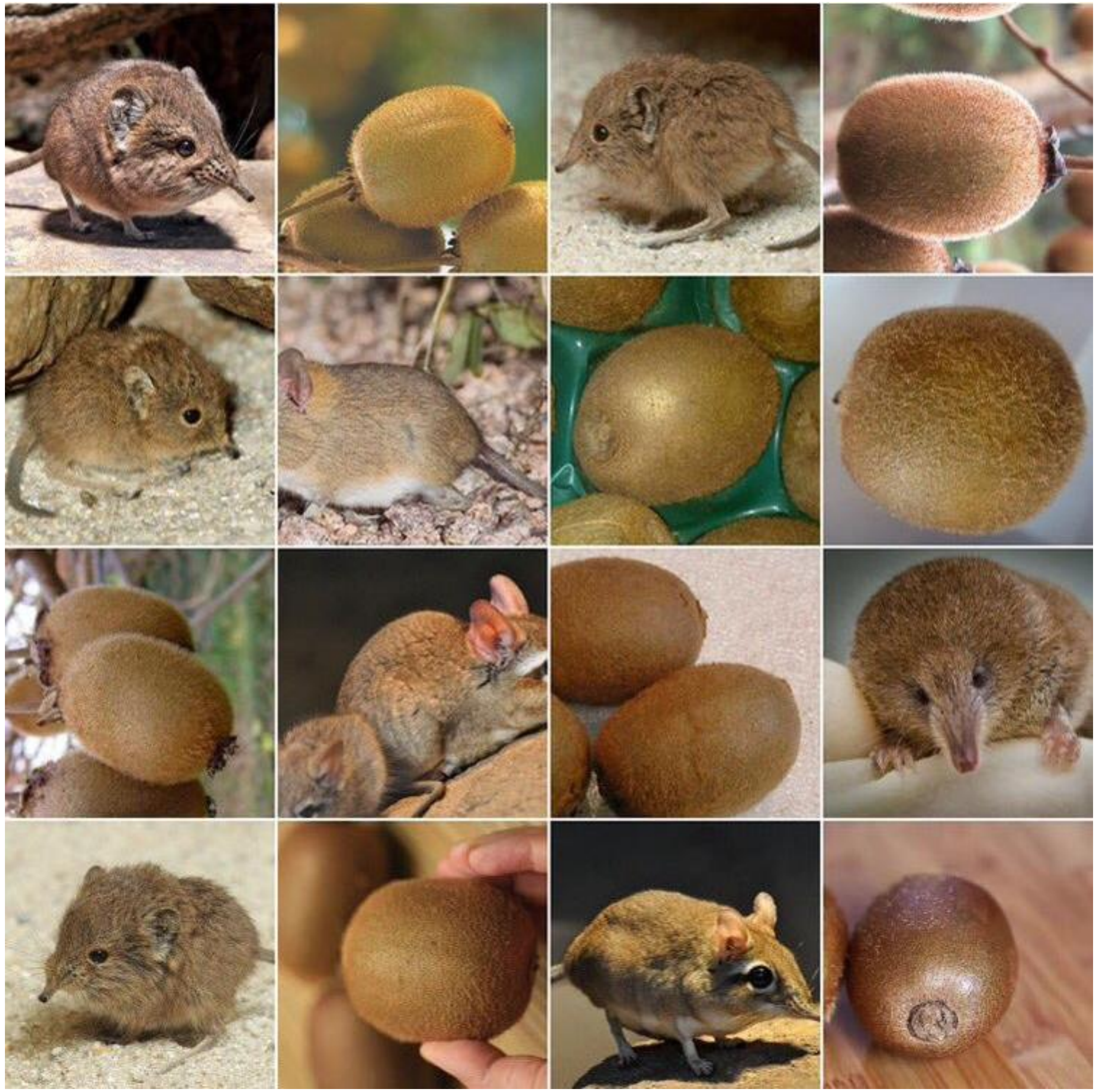
Puppy or bagel?

Credits Karen Zack

Sheepdog or mop?

Credits Karen Zack

[Shrew or kiwi?]

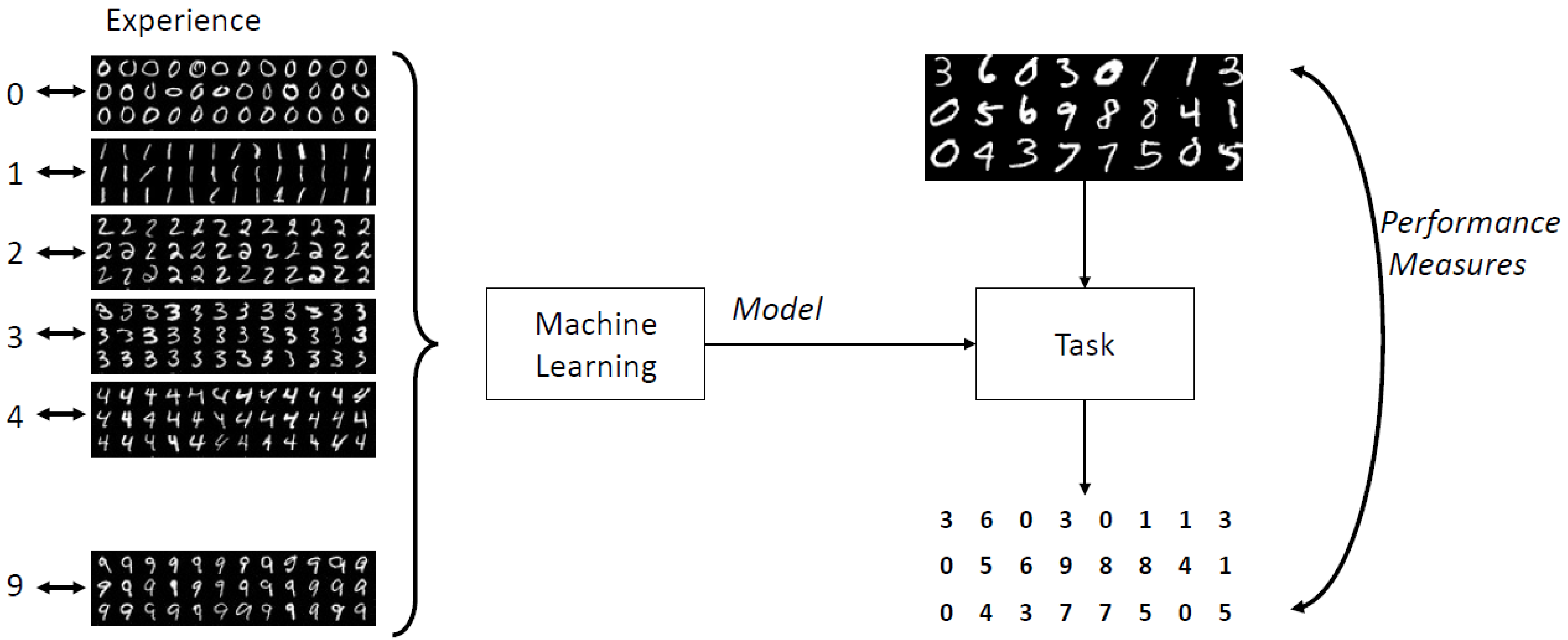


Credits [Karen Zack]

What is machine “learning”?

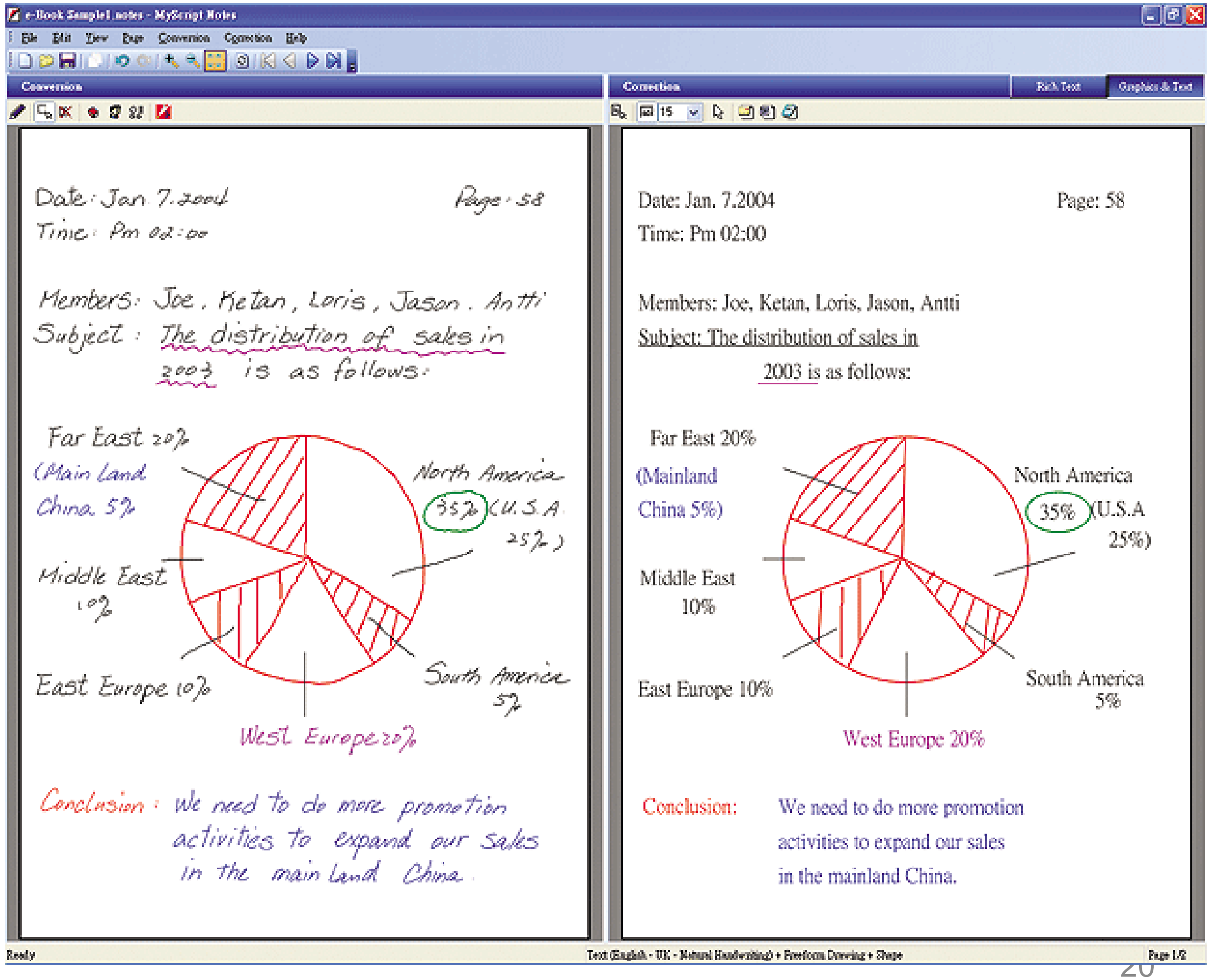
- Machines have been good (better than humans!) at doing complex tasks for a long time.
- The challenge is in extracting <u>semantic information</u> from the signal.
  - Tasks that are easy for people to perform but hard to describe formally.
  - The “meaning” of an image for example, recognizing spoken words.
- How can we program a computer to extract this information?
  - This is what we will cover in this course!
  - With a particular emphasis on dealing with medical data.
- A definition of ML: “Field of study that gives computers the ability to learn without being explicitly programmed” **Arthur Samuel** (1959)

Numbers recognition: Supervised learning on the MNIST dat



AIMLab.

Optical character recognition



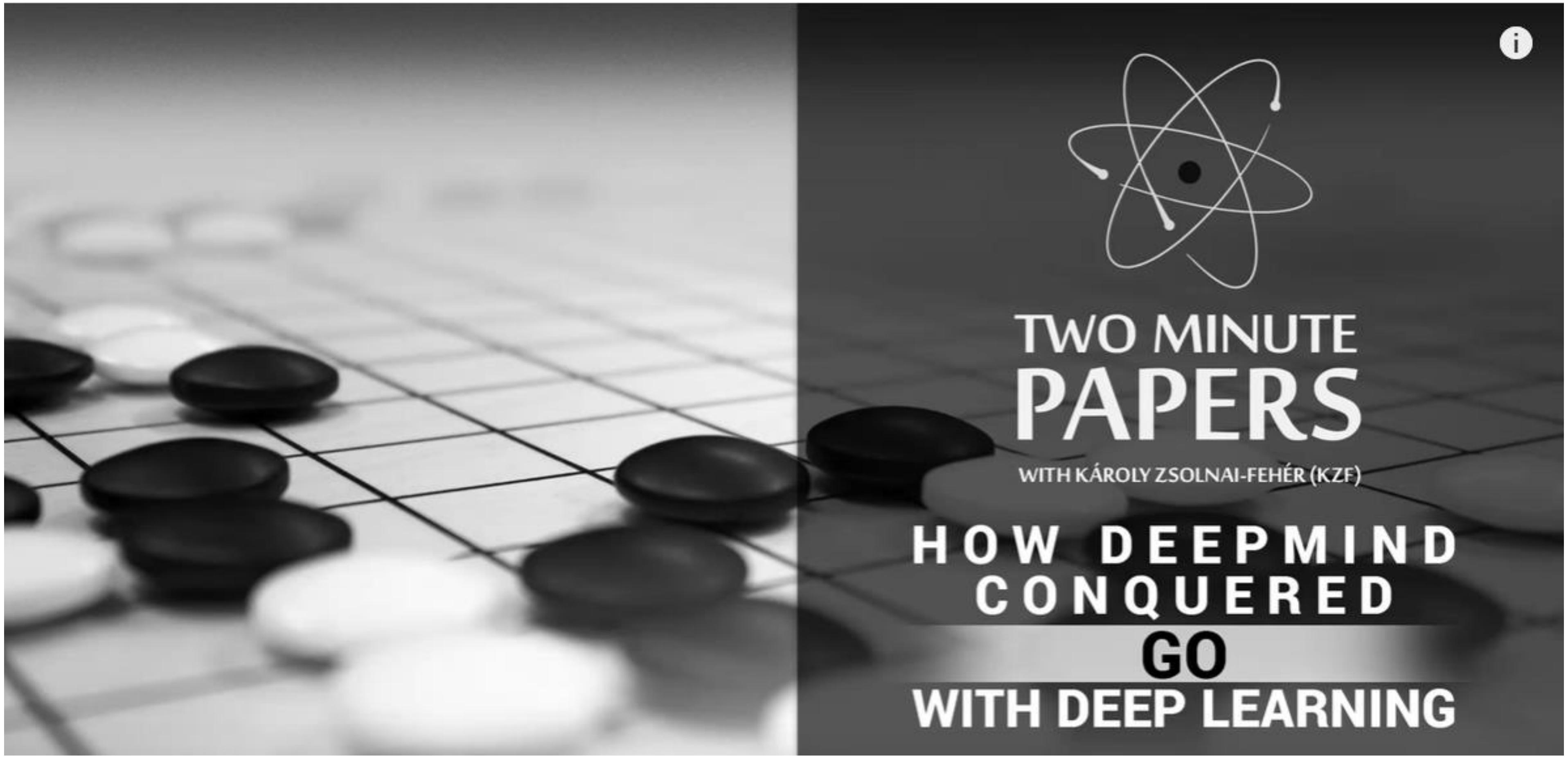
Objects Recognition: Supervised Learning

Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

Playing Atari: Reinforcement Learning

Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).

AlphaGo



[For more inspiration, movie on the AlphaGo](https://www.youtube.com/watch?v=jGyCsVhtW0M): https://www.youtube.com/watch?v=jGyCsVhtW0M
Latest news, AlphaGo Zero: https://deepmind.com/blog/alphago-zero-learning-scratch/

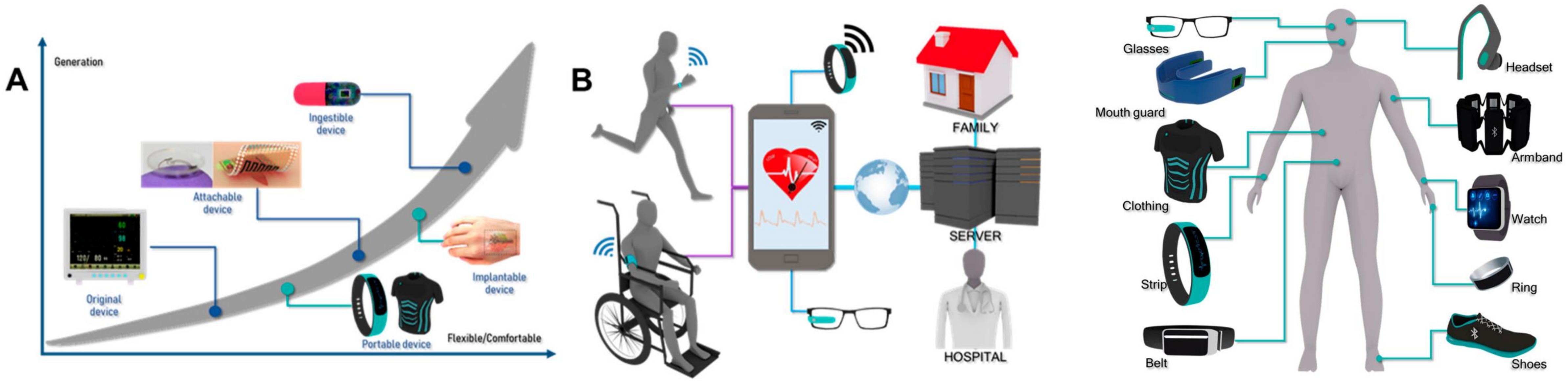
ML in Healthcare

ML in Biomedical Engineering

- Computer Assisted Diagnostics and monitoring:
  - From vital signs and waveforms such as ECG, EEG.
  - From medical imaging such as CT, MRI, ULS.
- Genetics and genomics research.
- Drug discovery.

21th Century: Big Data in Healthcare

- Easier to collect medical data in general.
- New portable sensors: open to long term monitoring and phenotyping of patients.
  - Many are low cost and widely available.



Guk, Kyeonghye, et al. "Evolution of Wearable Devices with Real-Time Disease Monitoring for Personalized Healthcare." Nanomaterials 9.6 (2019): 813.

21th Century: Electronic Medical Record (EMR)

- Consists of the systematized collection of patient and population electronically-stored health information in a digital format (Gunter et al. 2005).
- Remove the need to document history in paper and is more robust, complete, standardized and accessible.
- EMR may include different types of data: demographics, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, etc.
- Healthcare providers use them to improve their care management.
- In term of research, they significantly ease population based study by enable to look at long term trends and changes in patients.

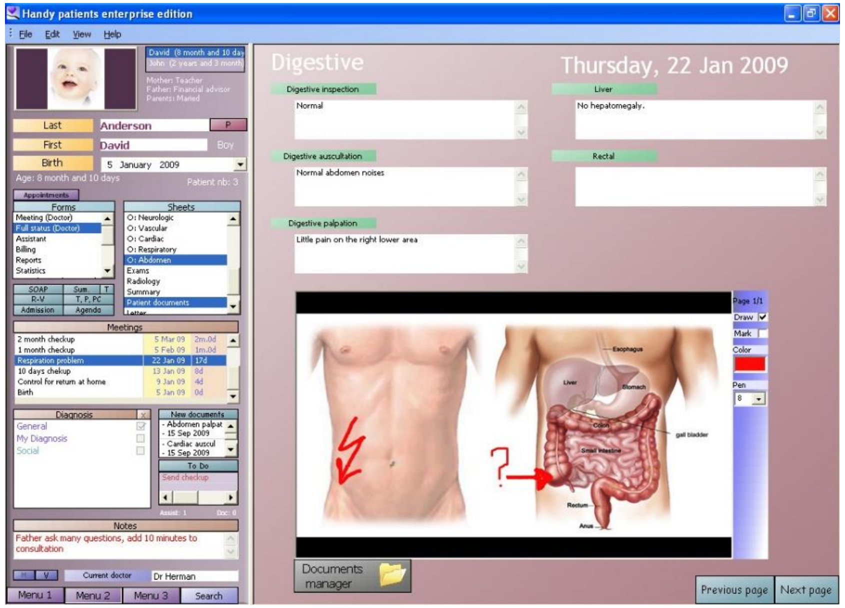
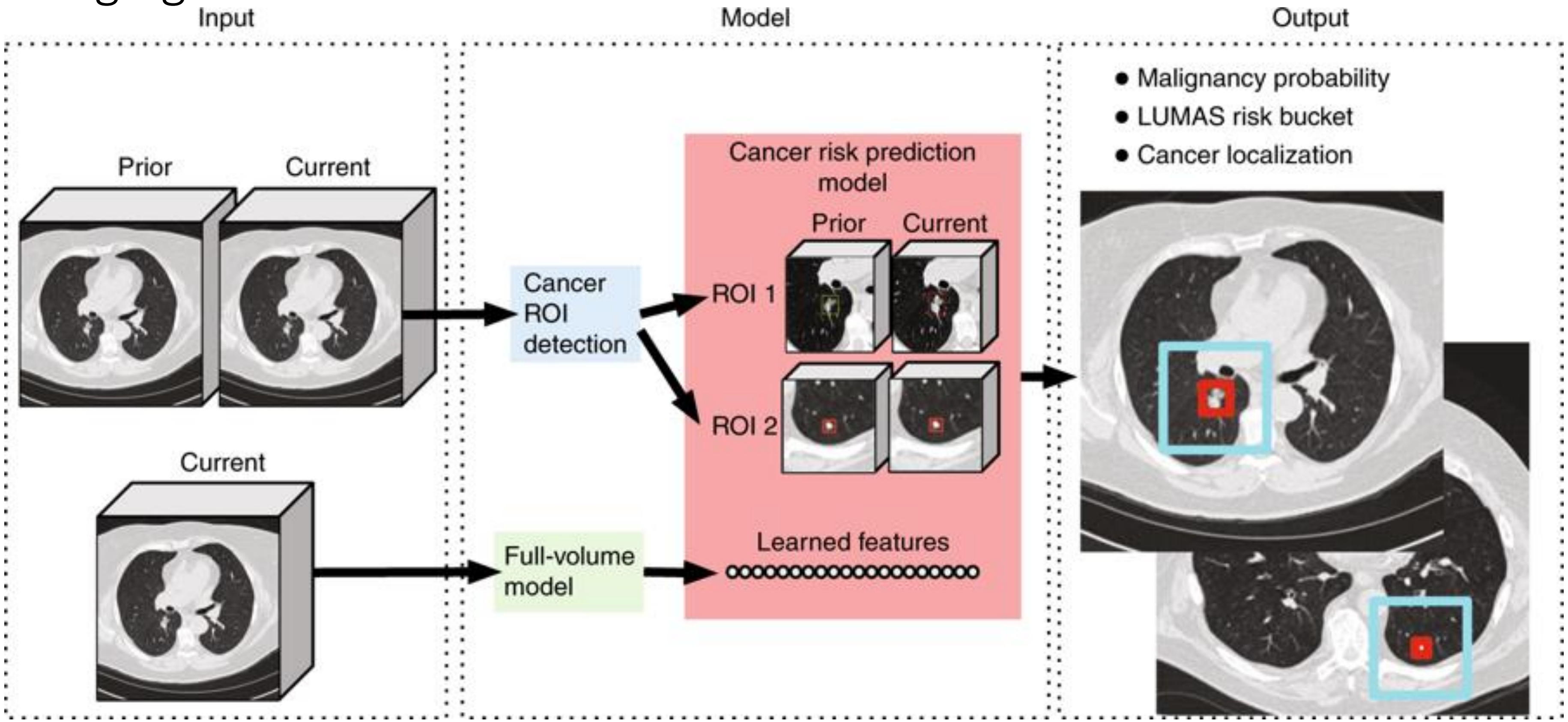


Image: By Oguntoye patients electronic medical record (free open source version), GPL, https://commons.wikimedia.org/w/index.php?curid=8894074

Examples: medical images

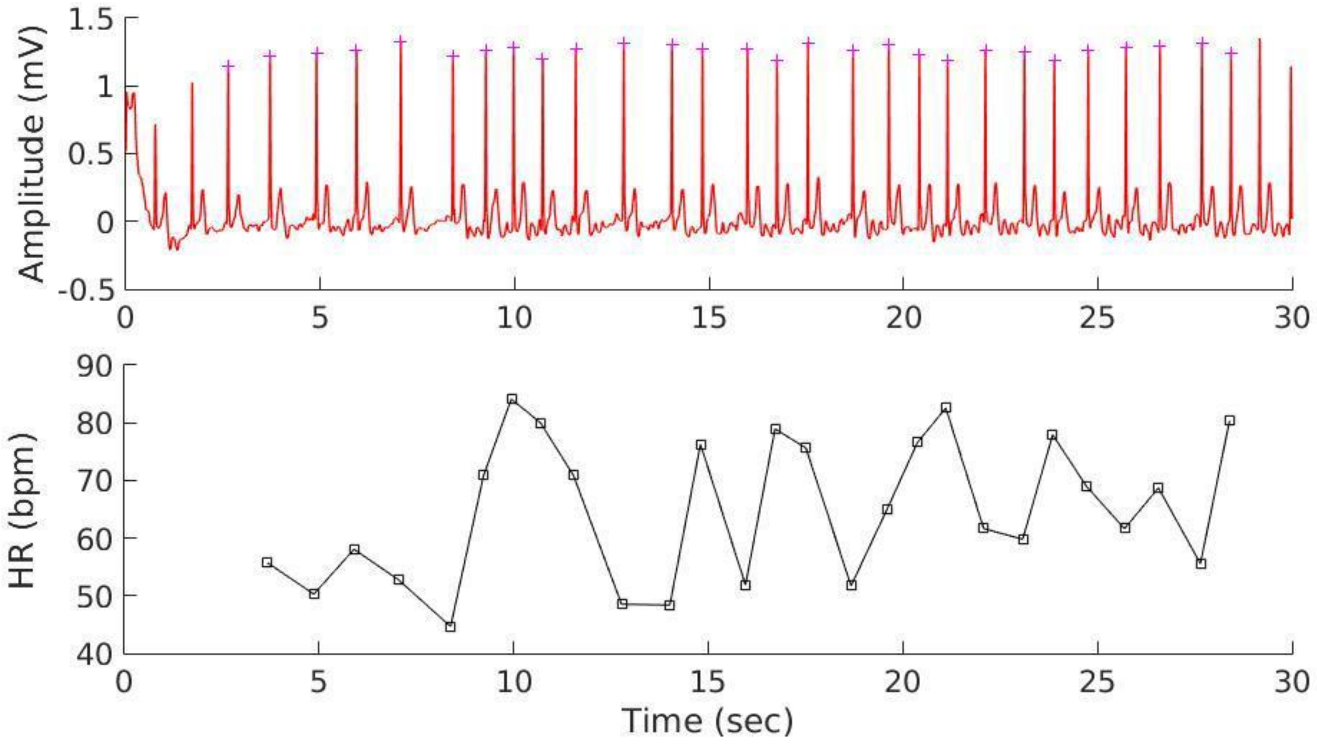
- Medical imaging:



Ardila, Diego, et al. "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography." Nature medicine 25.6 (2019): 954.

Examples: pattern recognition in physiological time-series

- Recognition of cardiac arrhythmias.

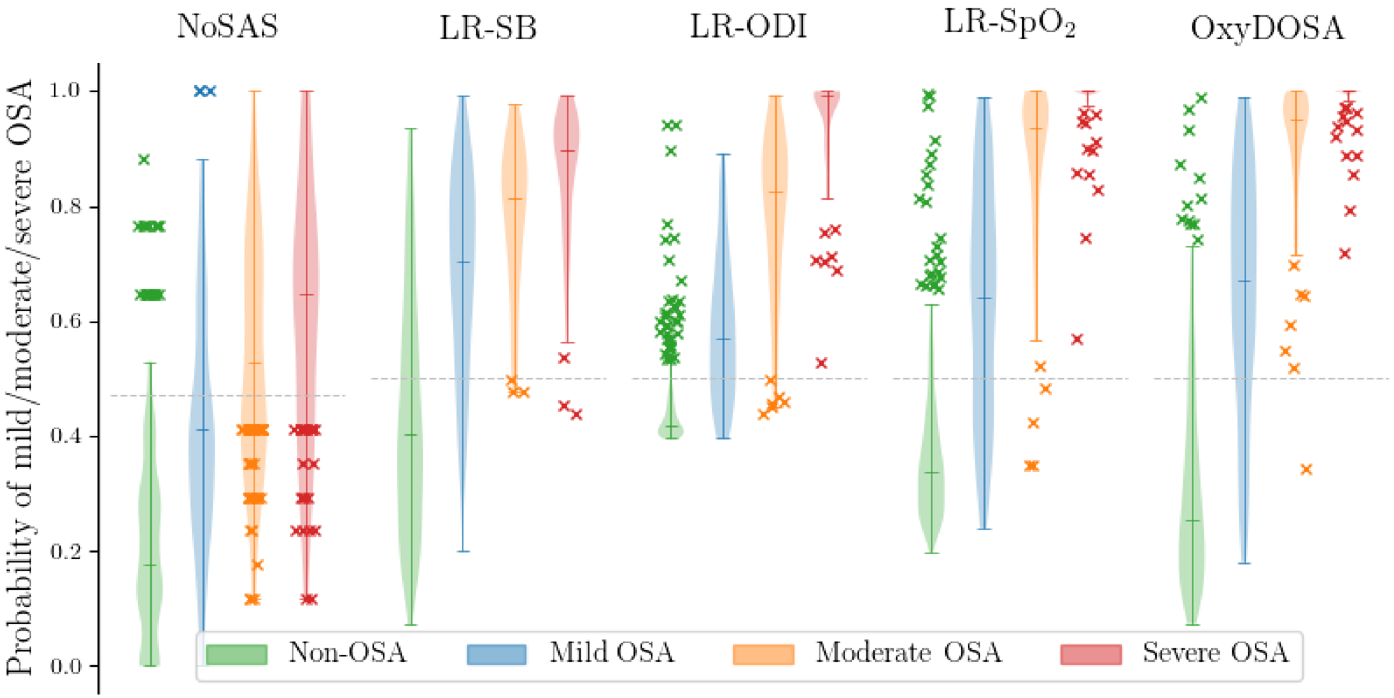


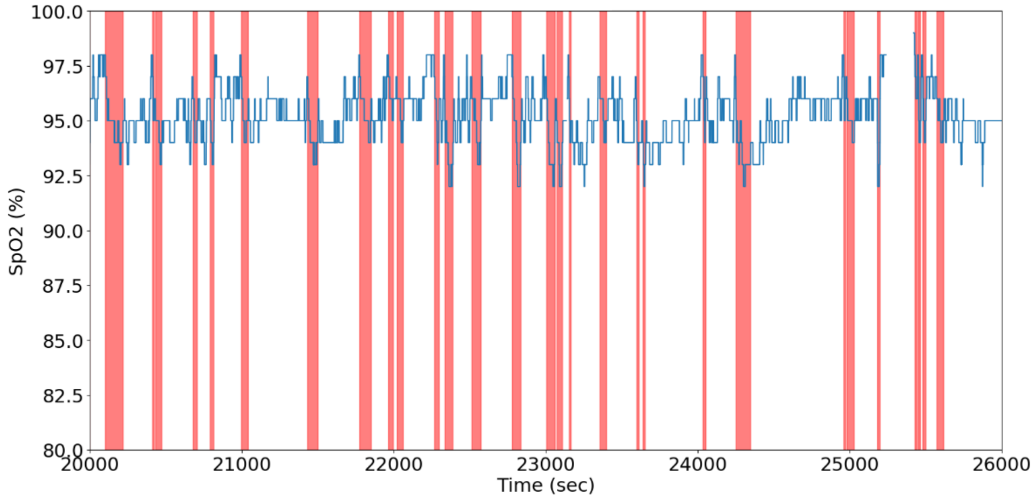
	n	a	o	~
N	4618	23	376	33
A	20	605	108	5
O	533	93	1792	39
~	50	7	52	175

Behar, Joachim A., et al. "Rhythm and quality classification from short ECGs recorded using a mobile device." 2017 Computing in Cardiology (CinC). IEEE, 2017.

Examples: pattern recognition in physiological time series

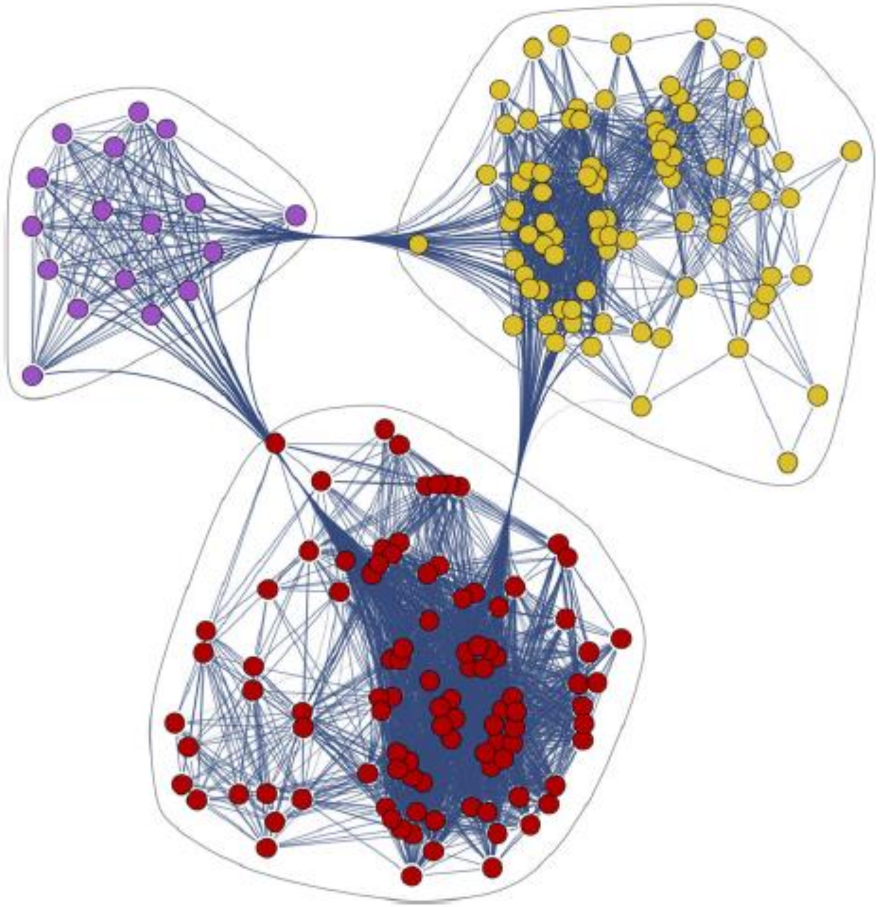
- Detecting obstructive sleep apnea from biomarkers, demographics and anthropometric information.





Behar, Joachim A., et al. "Feasibility of Single Channel Oximetry for Mass Screening of Obstructive Sleep Apnea." EClinicalMedicine (2019).

Examples: unsupervised learning for disease phenotyping

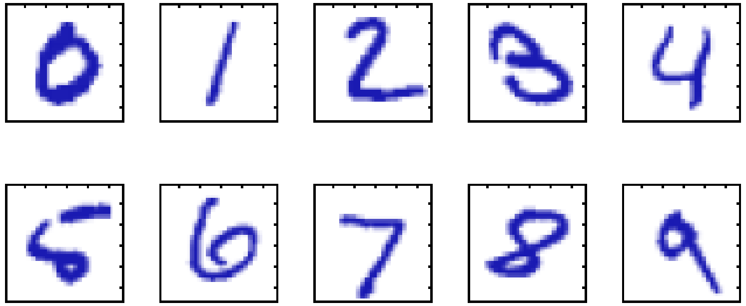


Lacedonia, Donato, et al. "Characterization of obstructive sleep apnea–hypopnea syndrome (OSA) population by means of cluster analysis." Journal of sleep research 25.6 (2016): 724-730.

Important Concepts in ML

AIMLab.

Vocabulary

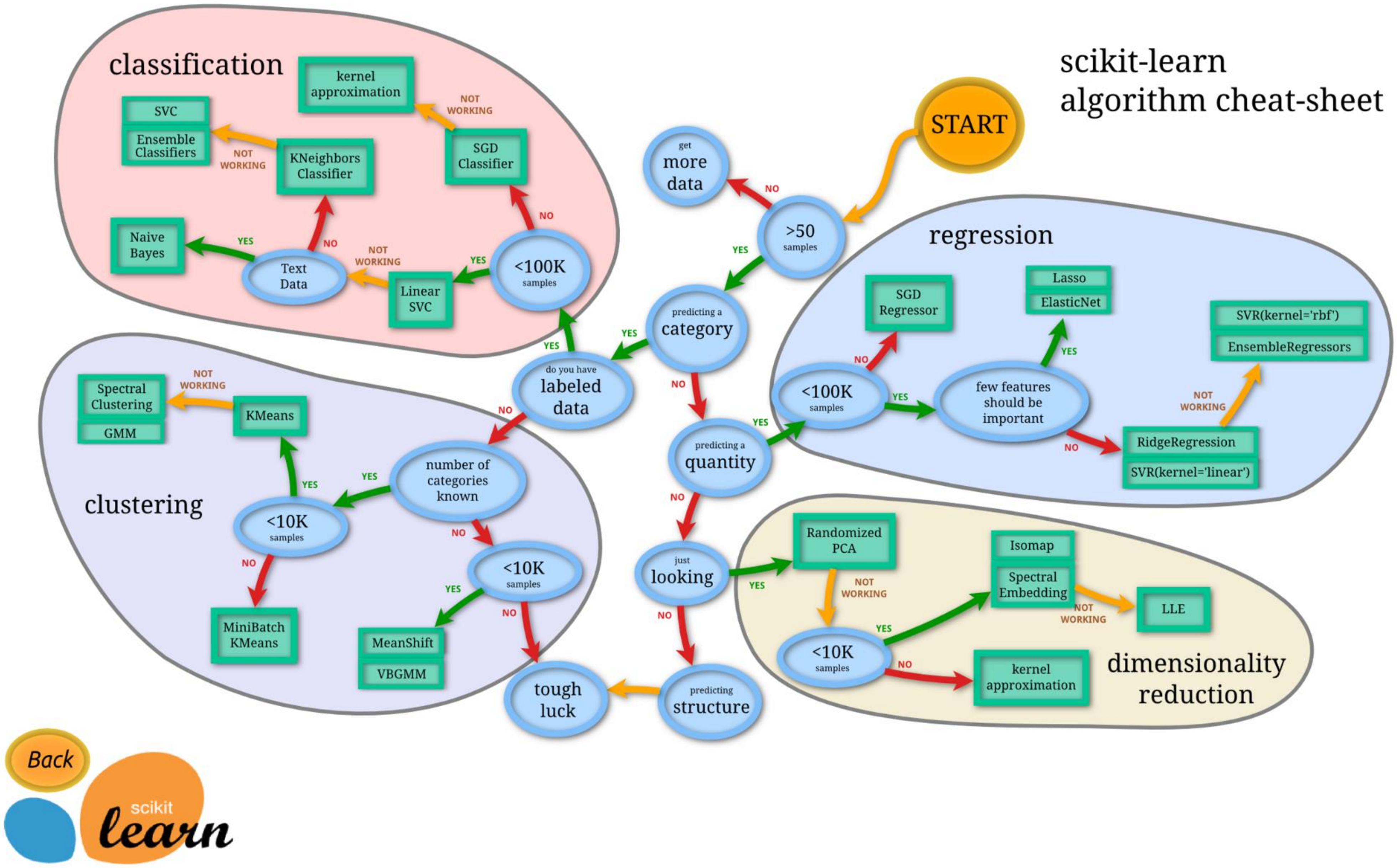


- Consider a set of m **training examples:**
  - $\left\{x_{train}^{(1)}, x_{train}^{(2)}, \dots, x_{train}^{(m)}\right\},$
- Consider the corresponding **target labels**:
  - $\left\{y_{train}^{(1)}, y_{train}^{(2)}, \dots, y_{train}^{(m)}\right\}$
- We seek the **target function** f so that:
  - $f\colon\ x^{(k)} \to y^{(k)}$
- This function is learned during the learning phase by using the **training set**. In a second step, we seek to evaluate it on an unseen set of examples that we will call the **test set**:
  - $\left\{x_{test}^{(1)}, x_{test}^{(2)}, \dots, x_{test}^{(p)}\right\}, \left\{y_{test}^{(1)}, y_{test}^{(2)}, \dots, y_{test}^{(p)}\right\}$
  - The ability to accurately categorize unseen examples is known as **generalization**.

Categories of ML

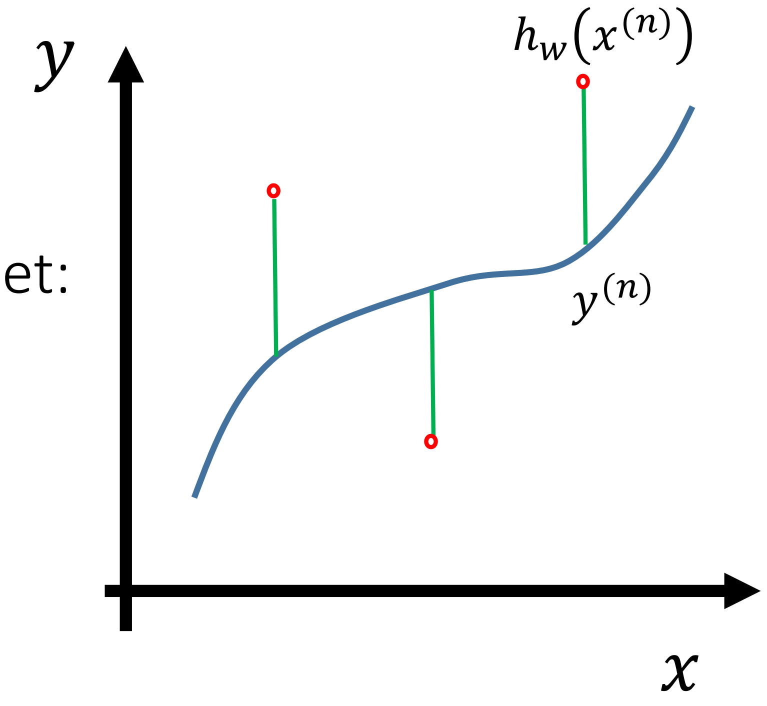
- **Supervised Learning**: the training data comprises examples of the input vectors along with their corresponding target vector.
  - Classification: output is one or finite number of discrete categories.
  - Regression: output is one or more continuous variables.
- **Unsupervised Learning**: input vectors but no target values.
  - Clustering: to discover groups of similar examples within the dataset.
  - Density estimation: to estimate the distribution of the data within the input space.
  - Visualization: project data from high-dimensional space down to two or three dim.
- **Reinforcement Learning**: finding suitable actions to take in a given situation in order to maximize a reward.

AIMLab.


scikit-learn
algorithm cheat-sheet
START
>50 samples
get more data
predicting a category
do you have labeled data
<100K samples
SGD Classifier
kernel approximation
Linear SVC
Text Data
Naive Bayes
KNeighbors Classifier
SVC
Ensemble Classifiers
classification
number of categories known
<10K samples
KMeans
Spectral Clustering
GMM
MiniBatch KMeans
MeanShift
VBGMM
clustering
predicting a quantity
regression
SGD Regressor
few features should be important
Lasso
ElasticNet
RidgeRegression
SVR(kernel='linear')
SVR(kernel='rbf')
EnsembleRegressors
just looking
Randomized PCA
Isomap
Spectral Embedding
LLE
kernel approximation
dimensionality reduction
predicting structure
tough luck
YES
NO
NOT WORKING
Back
scikit learn

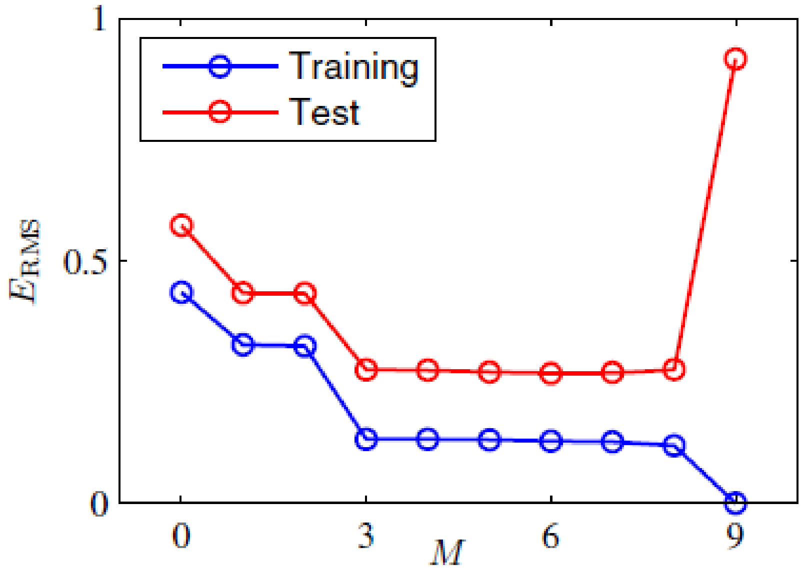
Other important concepts

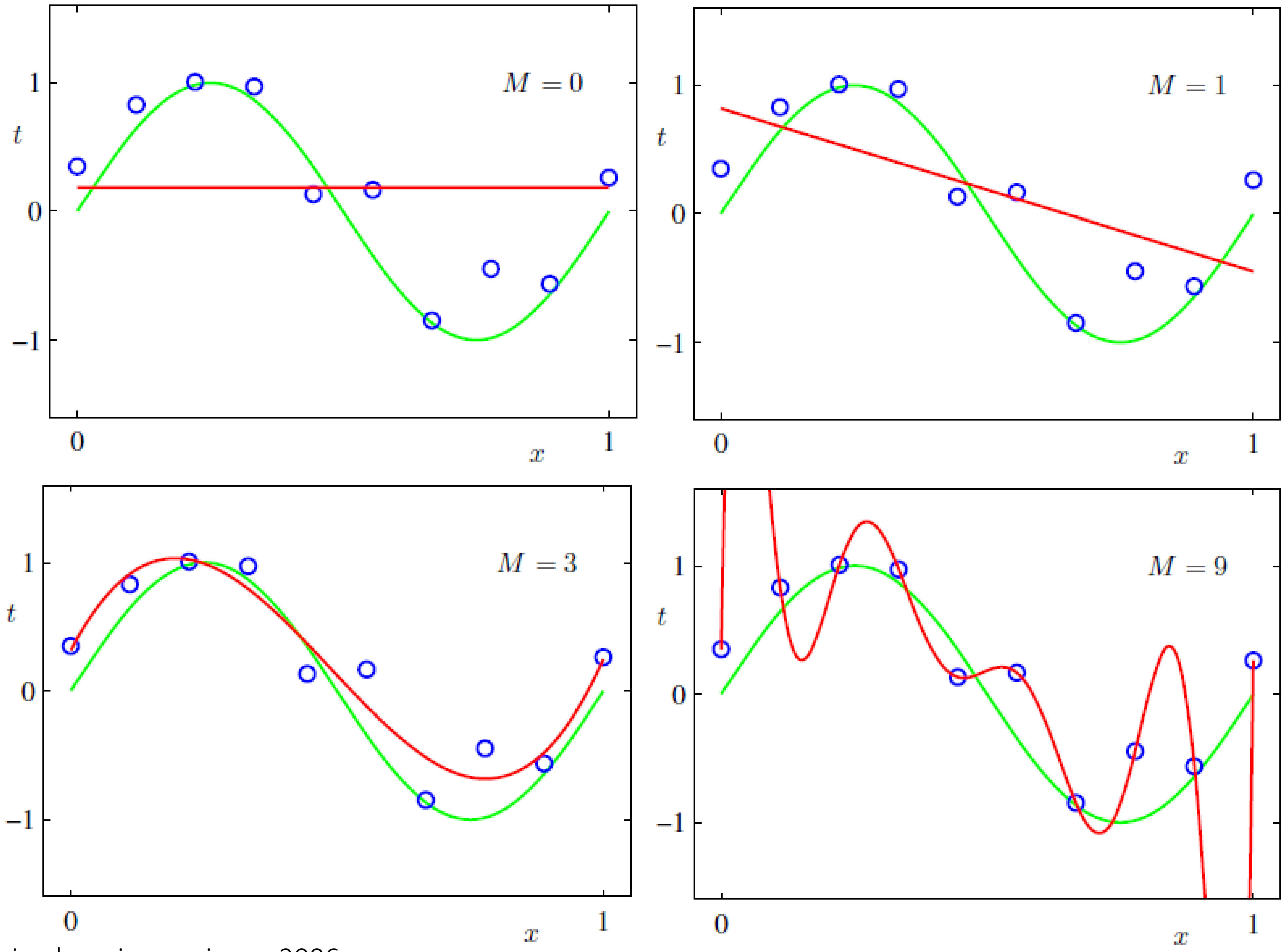
- We will use the curve fitting toy example.
- Consider a set of training examples and their associated target:
  - $\left\{x_{train}^{(1)}, x_{train}^{(2)}, \dots, x_{train}^{(m)}\right\}$,
  - $\left\{y_{train}^{(1)}, y_{train}^{(2)}, \dots, y_{train}^{(m)}\right\}$
- We will consider a polynomial **hypothesis function**:
  - $h_w(x) = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M$
  - M is the order of the polynomial.
- The value of the coefficients will be determined by fitting the polynomial to the training data. This will be achieved by minimizing an **error function** such as the mean square error:
  - $J(w) = \frac{1}{2}\sum_{n=1}^{N}\{h_w\left(x^{(n)}\right) - y^{(n)}\}^2$



Model selection

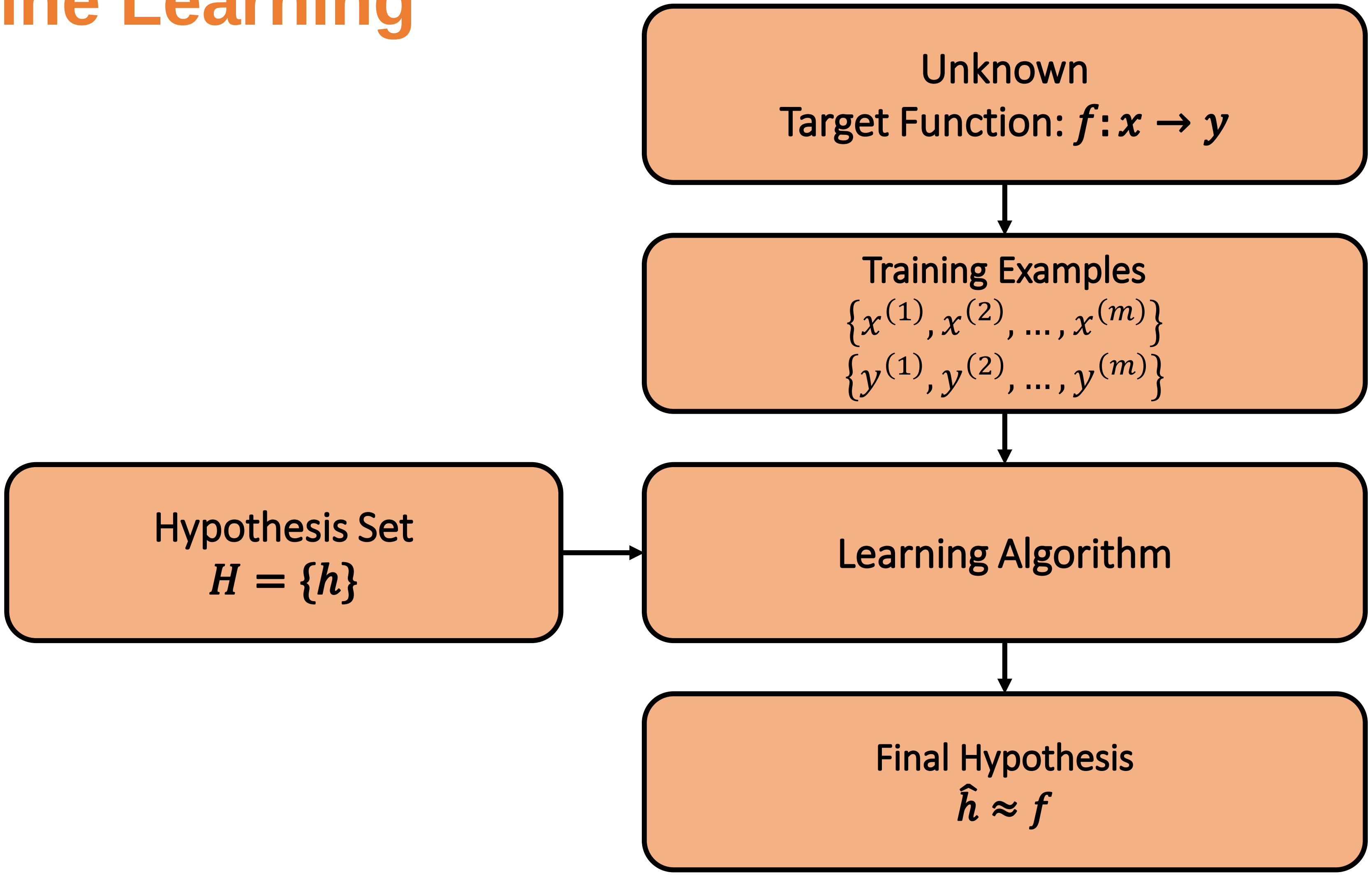
- Under fitting (high bias),
- Overfitting (high variance).





Images: Bishop, Christopher M. Pattern recognition and machine learning. springer, 2006.

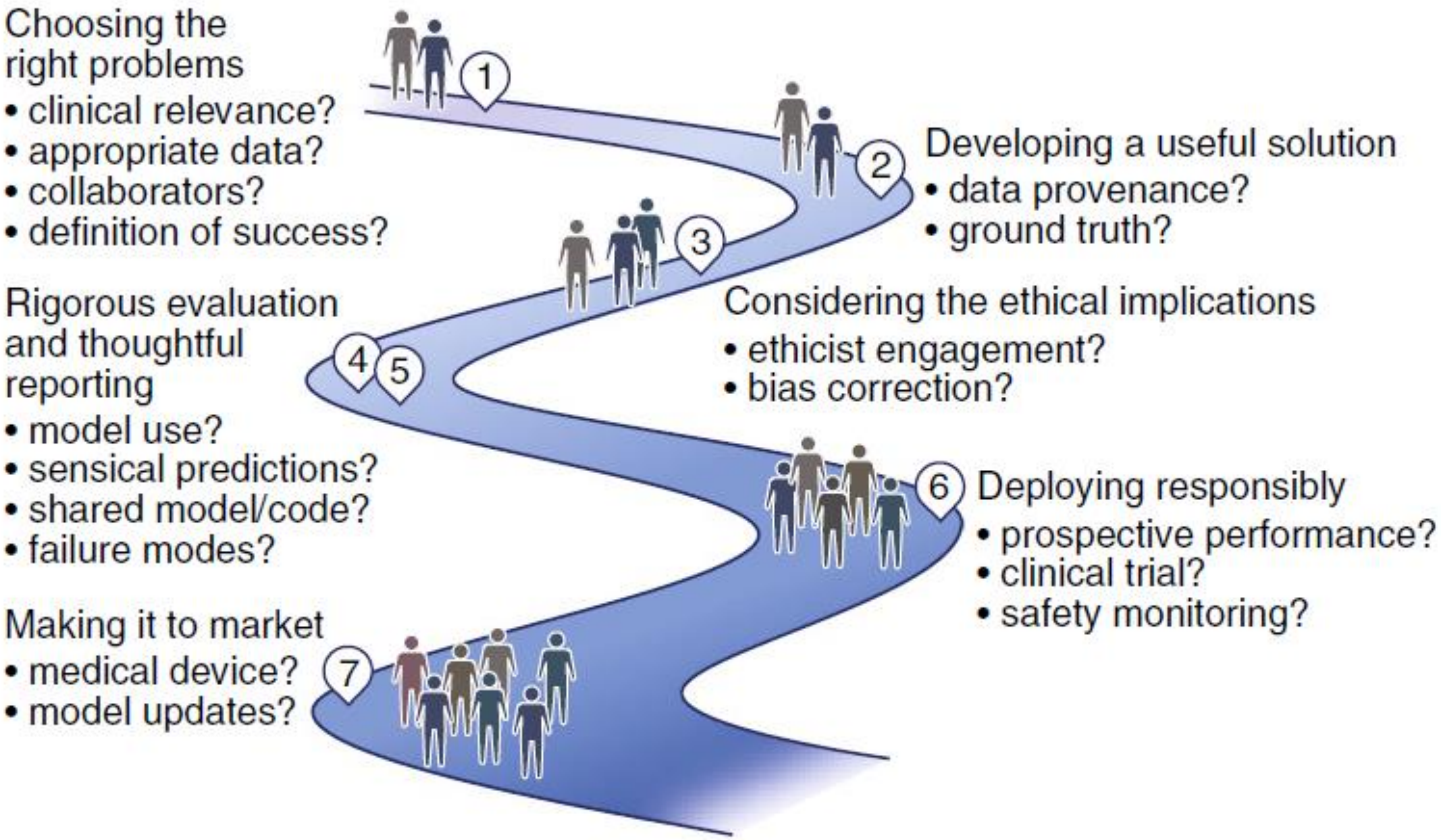
Machine Learning



References

[1] Deep Learning Course notes. Prof. Gilles Louppe. Universite de Liege. Spring 2019.
[2] Introduction to Machine Learning. Doron Shaked, Sagi Schein, Omer Barkol. BME course notes 2018.
[3] Pattern Recognition and Machine Learning. Springer 2006. Christopher M. Bishop.

Examples: unsupervised learning for disease phenotyping



Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... & Ossorio, P. N. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature medicine*, 1-4.

AIMLab.

History

1952: Program for playing Checkers improves itself
1957: Perceptron
70-80's: Expert systems
1986: Back Propagation
80-90's: Signal and image processing
1995: Support Vector Machine
90-2K's: Classification
2006: Netflix prize
2K's - : Big Data
2010: ImageNet Large Scale Visual Recognition Challenge (ILSVRC)
10's - : Deep Neural Networks

Rule based systems imitating experts in specific domains

Communication, Compression, Medical imaging, Medical signals, Digital audio, Image enhancements…

Spam filters, Face detection, Face recognition, OCR, Early medical diagnostics, Video Tracking…

Collaborative filtering, FinTeach, Search