Classification and Representation Learning

Course 1 : Introduction

Hoel Le Capitaine

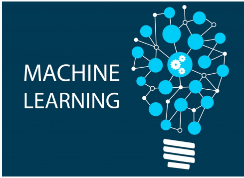
Academic year 2017-2018

Course organisation

Organisation

- mix of classes and autonomous work
- class with prof : 1h15min
- during autonomous time : exercises (left from tutorials), project (see later), read textbooks/papers (`go further` section)
- the distinction between prof class and autonomy should be clearly indicated on celcat (and the same applies for the other courses of the master, tell me if not)
- dot not hesitate to stop me and ask questions whenever something is not clear

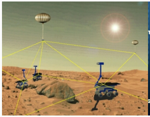
This is about Machine Learning

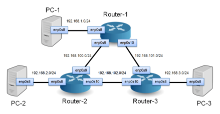


Some definitions of Machine Learning

- **Arthur Samuel (1959)**
  Machine learning: field of study that gives computers the ability to learn without being explicitely programmed.
- **Herbert Simon (1983)**
  Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more efficiently and more effectively the next time.
- **Tom Mitchell (1997)**
  A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

When do we need machine learning?

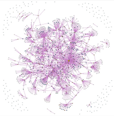


Automatic learning required when

- Human expertise does not exist (navigating on Mars, genome sequencing).
- Humans are unable to explain their expertise (speech recognition).
- The environment constantly changes with time (routing on a computer network, realistic robot).
- The solution needs to be adapted to particular cases (biometrics).
- The data to be analysed is too big (data-mining, web-search).



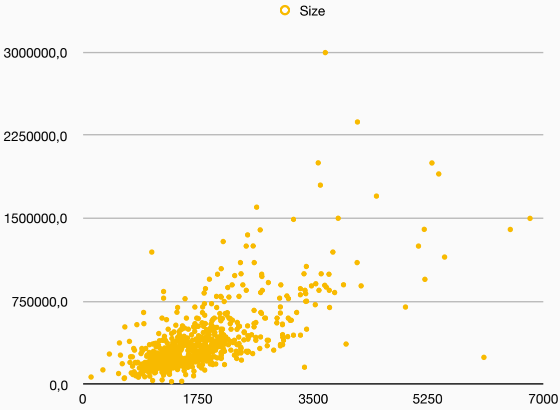
A number of different settings

Classification of main learning techniques

- **supervised learning** : the system is trained on an existing set of examples (inputs/outputs are known)
- **unsupervised learning** : the system must find some patterns and relations (only inputs are known)
- **reinforcement learning** : the system performs some actions and receives rewards (or punishments)

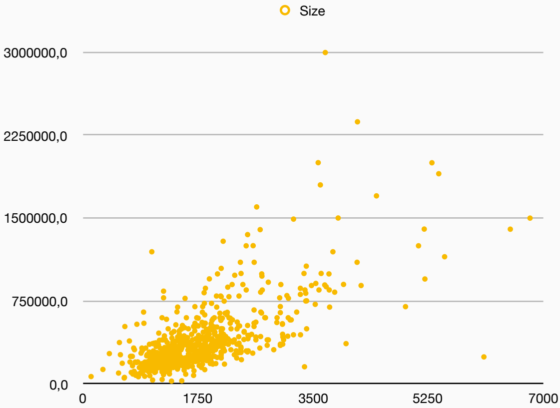
Other approaches : semi-supervised learning, transfer learning, ...

Sell your house :



Supervised learning

Sell your house : **regression** problem



Supervised learning

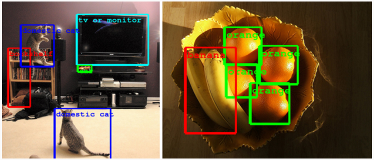
Cancer diagnosis :

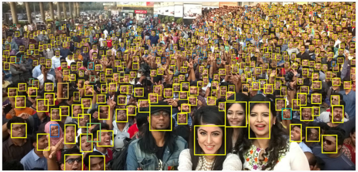
Cancer diagnosis :

Cancer diagnosis : **classification** problem

Prediction and recognition

- prediction
- recognition (object, face, handwriting, speech)

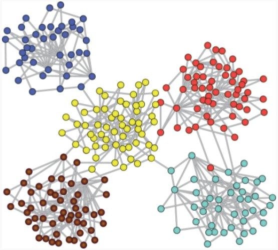


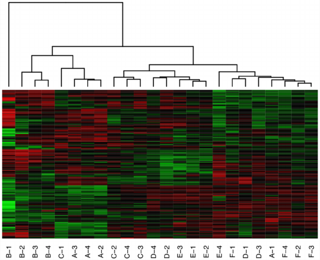


Unsupervised learning (or Knowledge Discovery)

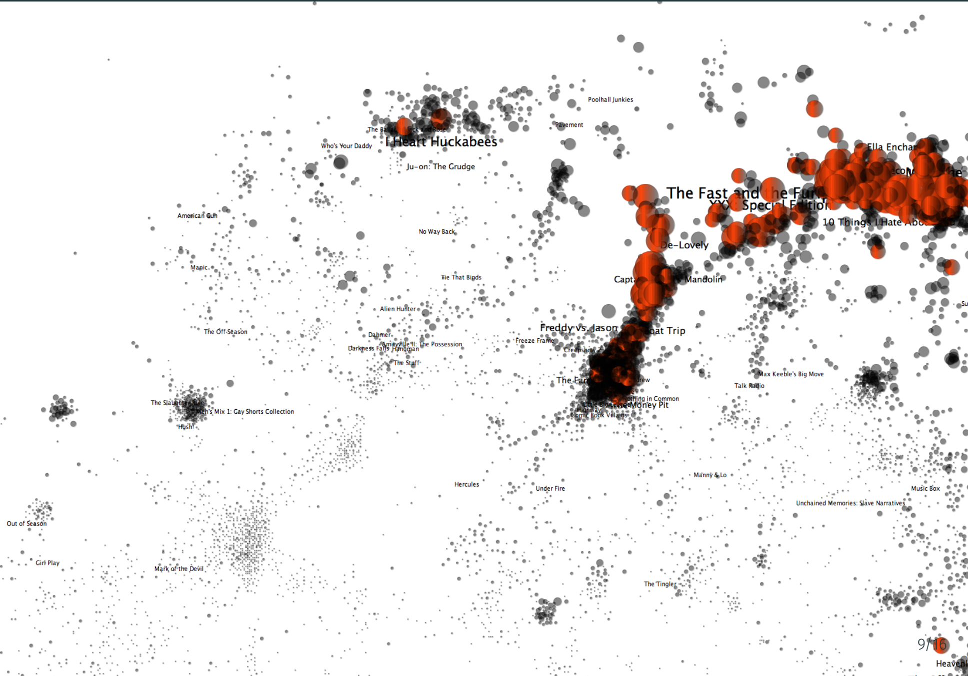
Clustering

- In unsupervised learning, only input data is provided to the algorithm, which has to analyse the statistical properties of the data.
- The goal of unsupervised learning is to build a model or find useful representations of the data, for example:
  - finding clusters of similar data and model their density dimensionality reduction
  - finding good explanations (hidden causes) of the data

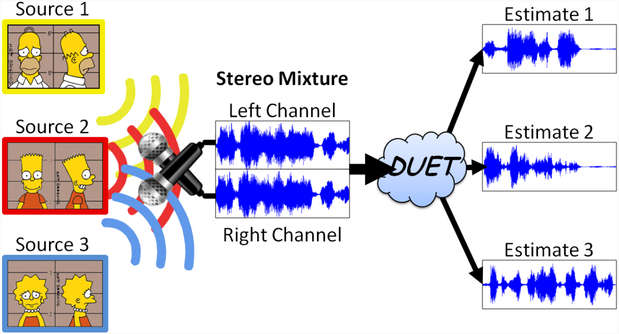




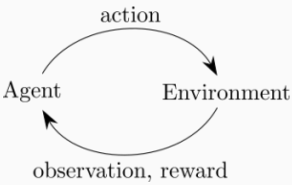
Applications of unsupervised learning



Applications of unsupervised learning



- Cocktail party problem: isolating a source from other ones.
- The separation is possible thanks to the statistical properties of the signals and to an hypothesis on their original distribution.
- This hypothesis (tone, distance) is called "belief" and allows to apply bayesian inference to the problem.
- Applications in signal processing, neural recordings, etc.



A more recent approach

- The agent interacts with its environment and learns a policy to map a state into an action in order to maximize the reward obtained from the environment.
- The correct action is never given, only if it correct or not. The system has therefore to try out all possibilities (exploration).
- Optimization of performance and approximation of functions. Drawback: the exploratory space can become huge.

Applications of reinforcement learning

Games

- Simple RL techniques back in 1995 have found the optimal strategy to play the backgammon.
- No previous knowledge of Backgammon: the algorithm learned the optimal strategy by playing against itself.
- RL algorithms are also applied to Poker, Chess or Go, but with less success until now.
- The main problem for RL is that reward (i.e. win or loss) should not be delayed too much, as in a game that never ends.
- Playing video games, e.g. Mario `https://www.youtube.com/watch?v=L4KBBAwF_bE`

Main other application: robotics (`https://www.youtube.com/watch?v=ZBFwe1gF0FU`)

Overview of this course

What will we do

- (re)calls on linear algebra and optimization
- linear learning machines
- learning theory and evaluation
- metric learning
- neural networks
- deep learning

Assessment

About your mark ...

- a project (30%)
- a 1.5h exam (70%), around week 45/46

The project

Principle

- Students form groups of 2-4 members.
- A list of candidate papers will be posted, and each group should pick one from the list.
- Each group is required to give an oral presentation (10 minutes) to the class about the content of the paper in the last two weeks, and submit a report at the end.
- The report should include at the minimum a summary of the method/framework, and experimental results obtained by playing the code published along with the paper.
- refer to extradoc for details/templates/papers (not yet)

References

Textbooks (not mandatory)

- Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, pp. 241-249). New York: Springer series in statistics.
- Murphy, K. (2012). Machine Learning : A probabilistic perpestive. MIT Press.
- Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.