Demystifying ML, AI & Automation

Spyros Gkezerlis

24-25.10.2017

TLAs

ML : Machine Learning
AI : Artificial Intelligence
BD : Big Data
NN : Neural Network
DL : Deep Learning
CNN: Convolutional Neural Network
NLP: Natural Language Processing

TLA: Two or Three Letter Acronym

MATH
Mental Abuse To Humans

Intro

Spyros Gkezerlis - @sgez

About
BSc Automation Engineering – Telematics Applications
MSc Information Systems – Reinforcement Learning
Interests – Applied ML, RL especially in time domain

Experience
4 years Software Engineer in Greece
9 years Engineer and Manager in OTE
3 years Program Manager in DTAG

Ariadni Gkezerli
(c) 2017

Why Lectures?

- "Set the record straight" sessions to shed light in the buzzwords of last couple of years so that there is a common understanding of what ML is and then assess potential.
- Pros cons, Tools, Technology etc. should not dictate what we should use! We should start with the problem we want to solve, work to the algorithms needed, then utilize best approach.
  - Inform what ML is and is not
  - Identify potential scope
  - Differentiate between disciplines
  - Awareness on current toolset
- Use a common language

Do you want to become an ML Jedi?

Why ML?

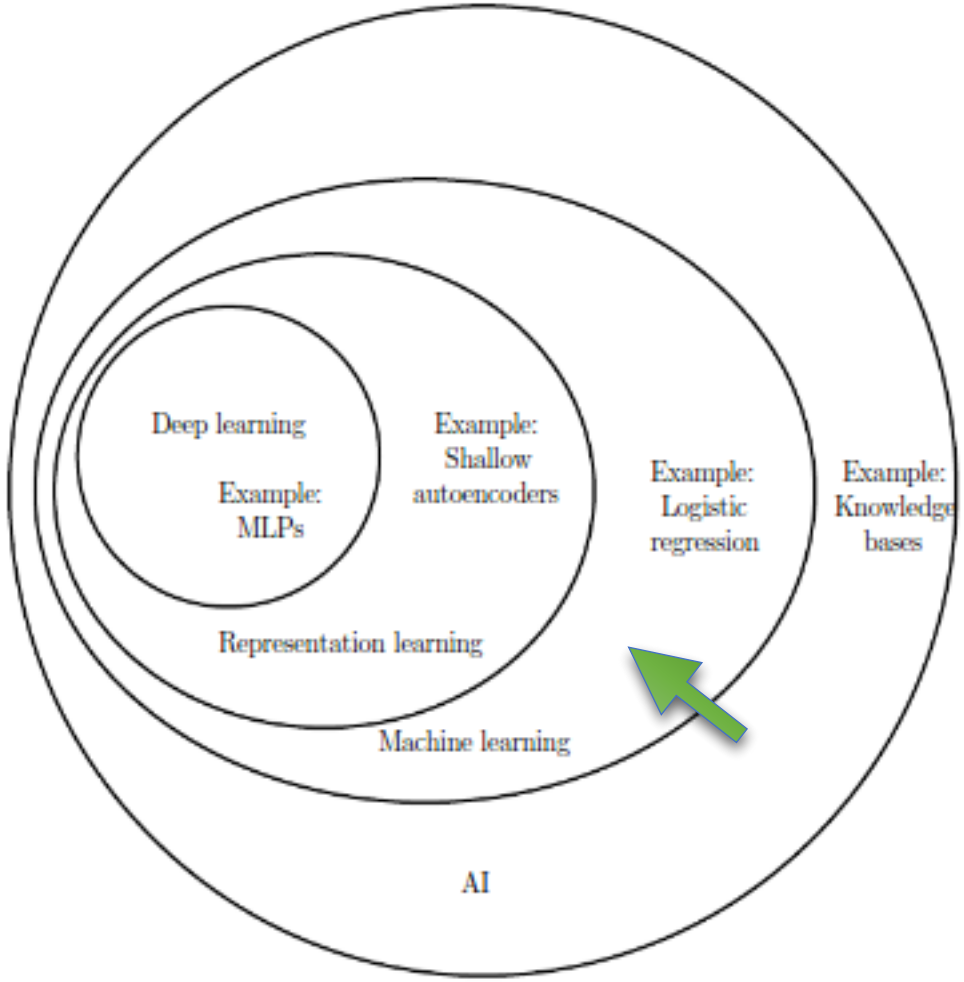
- Why use Automation, Adv.Stats, ML, AI? Because it can help us:
  - Reduce complexity of network
  - Improve experience by Time-to-market, Time-to-repair
  - Repetitive caused costs can be targeted and reduced
- Also… AI is another fun way to solve problems ☺
- But it can solve particular types of problems!!!

Lecture Guidelines

- Stop for questions at any time! Be curious!!!
- Lunch all together for questions and NT
- Workflow to be used?
  Kanban board with TODO, DOING, DONE
- First iteration through everything is a must
  if more items:
  put in TODO
- Set as TODO from the beginning
  Basic understanding of the landscape
- Syllabus will be asked! Learn the new dictionary!

Landscape overview

- Automation
- Statistics & Probability
- Data Mining
- Artificial Intelligence
- Machine Learning
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning
- Big Data Analytics



From Ian Goodfellow's Deep Learning Book: http://www.deeplearningbook.org

Theory

- Probability Theory Bayes Theorem

  P(AlB) ==> Probability of A given that B

  - Example
    Out of 3000 emails received over a certain period, 2000 are spam and 1000 are not. The word “Rolex” appeared in 250 out of the 2000 which are spam and in 5 out of the 1000. So, if an email is received, and contains the word “Rolex”, what is the possibility that it is a spam?

Let S be the event that the message is spam, and E be the event that the message contains the word w. Under our assumption from before, we have that:

$$p(S|E) = \frac{p(E|S)}{p(E|S) + p(E|\overline{S})}$$

https://www.arl.wustl.edu/~todd/cse240/cse240_lecture17_3-24-16.pdf

Theory

- Example – Solution:
  Out of 3000 emails received over a certain period, 2000 are spam and 1000 are not. The word "Rolex" appeared in 250 out of the 2000 which are spam and in 5 out of the 1000.

  So, if an email is received, and contains the word "Rolex", what is the possibility that it is a spam?

Now p(w) and q(w) are empirical estimates of p(E|S) and p(E|S')

$$r(Rolex) = \frac{p(Rolex)}{p(Rolex) + q(Rolex)} =$$

$$\frac{0.125}{0.125 + .005} = \frac{0.125}{0.125 + .005} \approx 0.962$$

https://www.arl.wustl.edu/~todd/cse240/cse240_lecture17_3-24-16.pdf

Theory

- Statistics Markov Chains

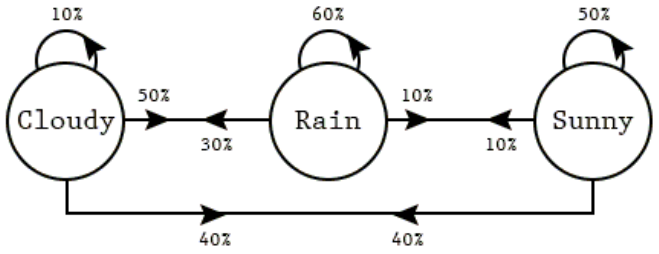
  Markov Chains is a probabilistic process, that relies on the current state to predict the next state. For Markov chains to be effective the current state has to be dependent on the previous state in some way; For instance, from experience we know that if it looks cloudy outside, the next state we expect is rain. We can also say that when the rain starts to subside into cloudiness, the next state will most likely be sunny.

  Not every process has the Markov Property, such as the Lottery, this weeks winning numbers have no dependence to the previous weeks winning numbers.

Markov Table of Probabilities

STATE	NEXT STATE	PROBABILITY	%
CLOUDY	CLOUDY	0.1	10%
CLOUDY	RAIN	0.5	50%
CLOUDY	SUNNY	0.4	40%
RAIN	CLOUDY	0.3	30%
RAIN	RAIN	0.6	60%
RAIN	SUNNY	0.1	10%
SUNNY	CLOUDY	0.4	40%
SUNNY	RAIN	0.1	10%
SUNNY	SUNNY	0.5	50%

Markov State Diagram



Current State Vector

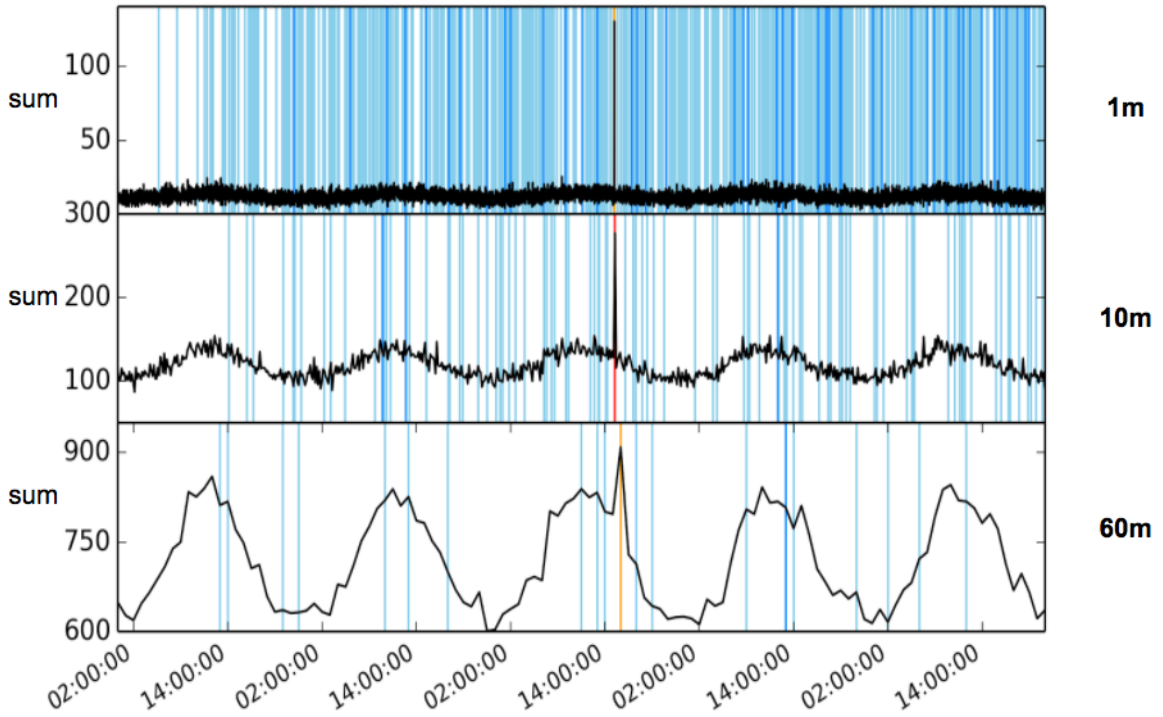
C	R	S
1	0	0

Figure 4

http://techeffigytutorials.blogspot.com/2015/01/markov-chains-explained.html

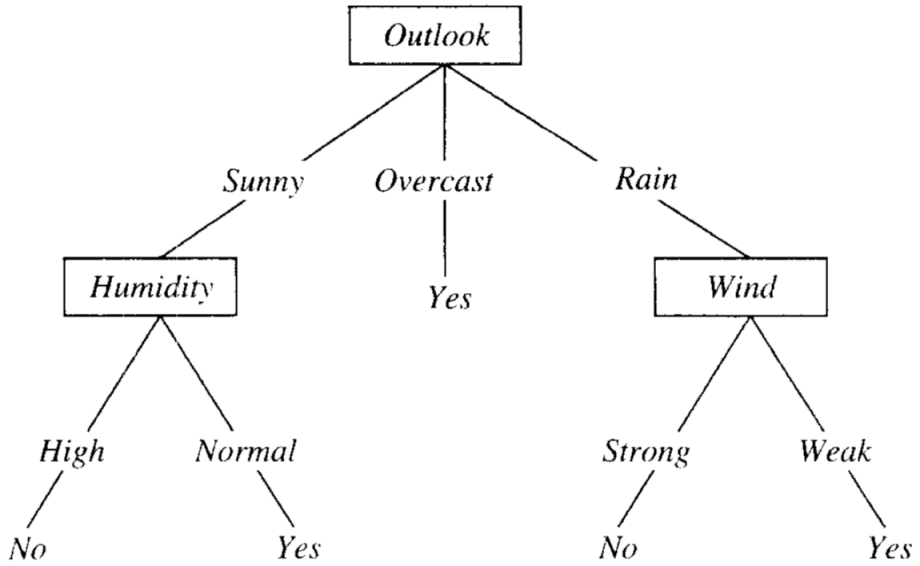
Theory

- Time Domain
- Event & Anomaly Detection
- Pattern recognition
- Learning behaviour



Theory

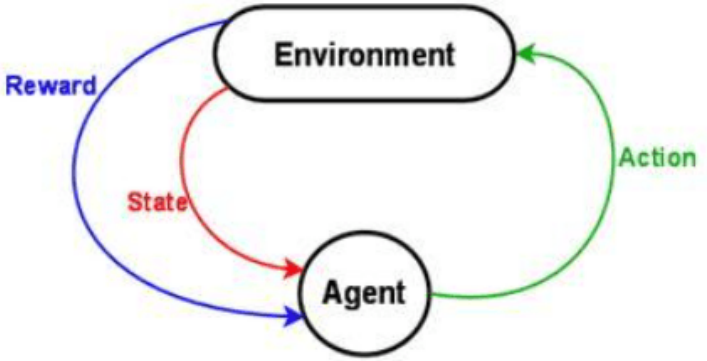
- Supervised Learning - Decision Tree Learning
  DTL is method for approximating discrete valued target functions, in which the learned function is represented by a decision tree.

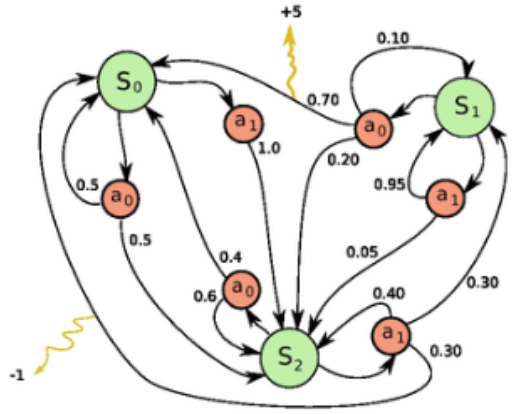


http://www.cs.princeton.edu/courses/archive/spr07/cos424/papers/mitchell-dectrees.pdf

Theory

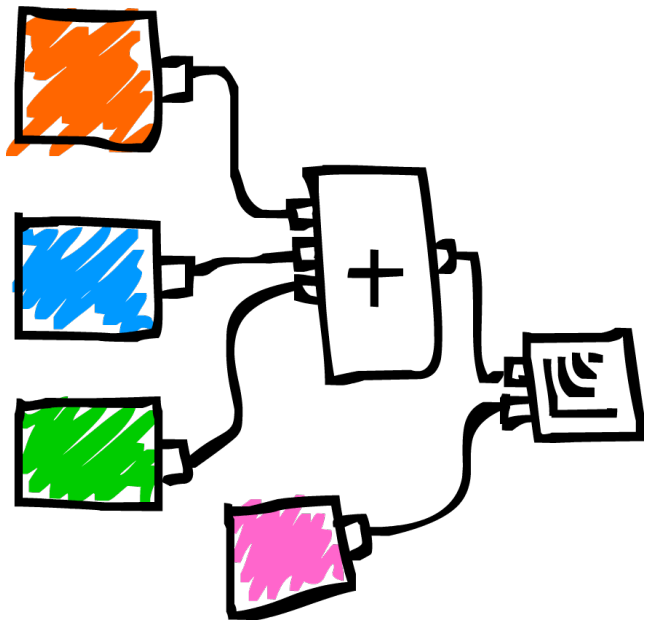
- Reinforcement Learning
  - inspired by behaviourist psychology
  - RL is about how agents ought to take actions in an environment so as to maximise some notion of cumulative reward.
  - A state is the current status
  - An action is decided based on previous reward levels
  - A reward is given based on the next state
  - Usual structure that an agent can use to describe the environment can be a graph / NN which has starting states and end / target states.



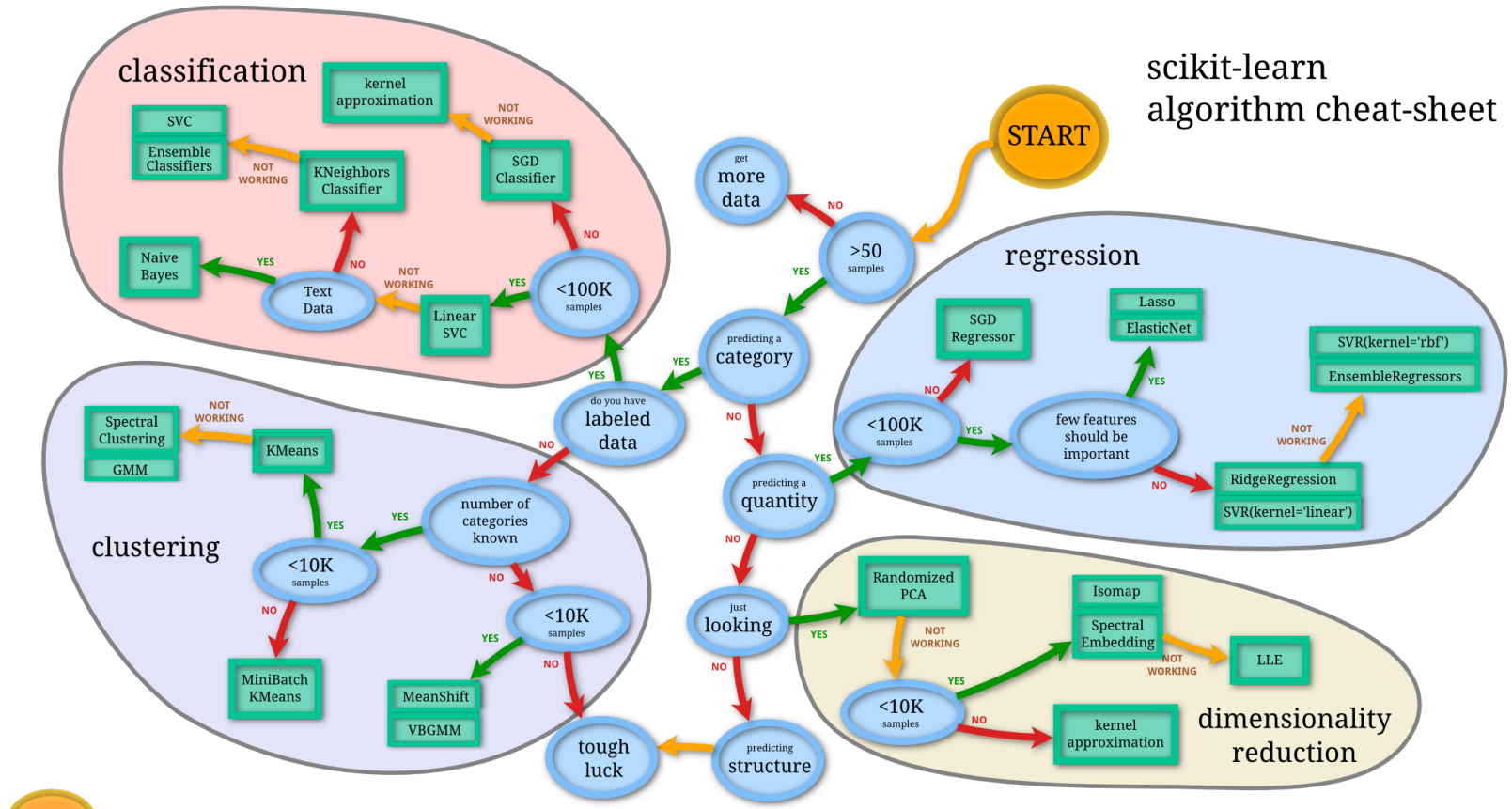


Workflow – The ML Pipeline

- Start with the question
- Find proper data and sources
- Create a data set
- Train system & classify
- Test the system and fine-tune
- Learn & automate
- Apply to workflow



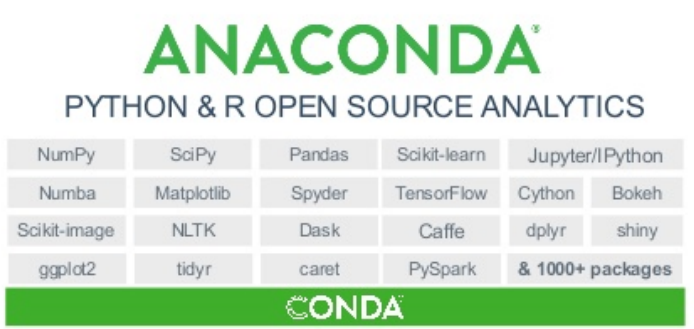
Selecting Proper Algorithms

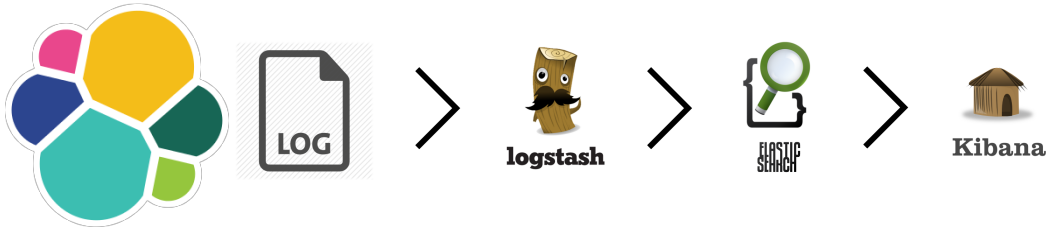


Back

Tools – Comparisons

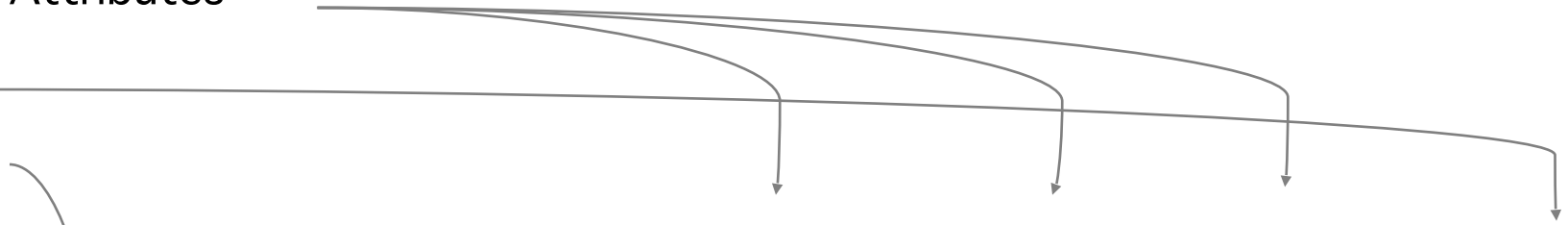
- Weka
  - Data Mining GUI
- Anaconda
    - Python
      - Dask
      - Scikit-Learn for ML/DL
      - Tensorflow for DL
    - R Studio
- Elasticsearch – Logstash - Kibana





Datasets Syllabus

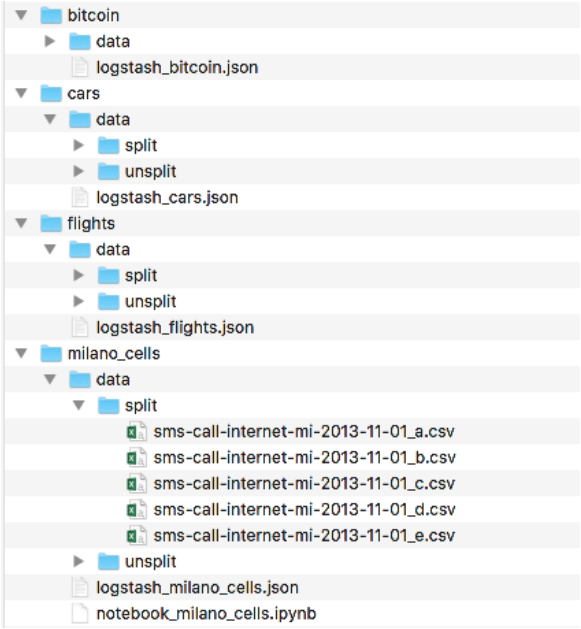
- Dataset
  - Features or Attributes
  - Labels
  - Instances



Age	Nationality	Education	Job
25	Greek	MSc	Engineer
38	American	MSc	Analyst
35	Japanese	PhD	Assistant
…	…	…	…

Example Datasets

- bitcoin
- cars
- flights
- milano_cells



Lectures Page

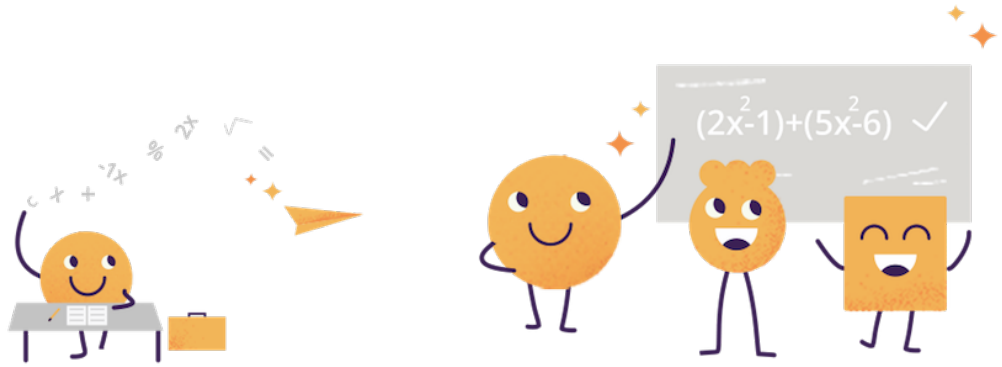
https://github.com/sgez/MLAI

Practice (fun part :-)

- Weka For Data Mining / basic ML
- Anaconda Python & R for ML
- E L K for Visualization & ML

  on

  - bitcoin
  - cars
  - flights
  - milano_cells



Special Thanks!

24.10.2017

- Diogenis Georgiadis
    - Zeta Papadopoulou
        - Christos Kavakopoulos

25.10.2017

- Christos Penitas
    - Antonis Galanakis
        - George Georgitselis