Data Mining

Matteo Francia
DISI — University of Bologna
m.francia@unibo.it

Hi!

Matteo Francia, Ph.D.

- Email: m.francia@unibo.it
- Assistant Professor (junior) @ DISI, UniBO
- www: https://www.unibo.it/sitoweb/m.francia/en

Teaching

- [DTM] Big Data and Cloud Platforms (Module 2)
- [DTM] Machine Learning and Data Mining (Module 2)

Research topics

- Big data / database / machine learning
- Precision agriculture and spatio-temporal analytics

BIG (Business Intelligence Group)

- w/ Prof. Golfarelli, Prof. Gallinucci
- You can check our thesis projects here

Table of Contents and Course Goals

Hands-on *machine learning* (7.5 hours + 7.5 hours with Prof. Gallinucci)

- *Theory*, *practice*, and *case studies* of the main steps of the CRISP-DM methodology
  - We will have coding labs
  - Python libraries for machine learning (mainly `pandas` and `scikit-learn`)

Goals

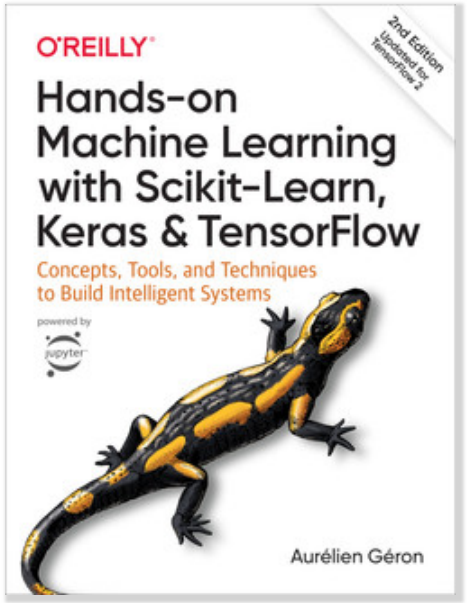
1. *Master the right terms and paradigms* behind AI
2. *Develop (=code) your own (simple) project* based on AI
3. *Make technical considerations* about AI
   - "What happens if I do …?"
4. *Understanding the surrounding AI world*
   - Make decisions with AI
   - Try to make predictions about AI for your company
   - Orient yourself in an extremely dynamic and complex world dominated by AI

Teaching material

Slides and Python notebooks (to be opened on Google Colab) are sufficient to prepare for the final exam

However, I suggest you to read the book:

- *Title*: Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow
- *Author*: Aurélien Géron
- Some notes about the book:
  - The book is intended as a support
  - The book is available for free in the library
  - The book is pretty huge (about 800 pages)
  - Carefully check the summary and select our topics
  - If you want, you can read the other parts to take inspiration (not requested for the final exam)



Hands on ML

Tentative outline

1. (w/ me) From business understanding to data-preprocessing
2. (w/ me) Hands-on machine learning (lab session)
3. (w/ Prof. Gallinucci) Big data pipelines
4. (w/ Prof. Gallinucci) Hands-on big data (lab session)

CRISP-DM