

A Phd Course

Let's start with an observation:

Machine Learning is Optimization

Let's start with an observation:

Machine Learning is Optimization

- Well, maybe a few exceptions apply (looking at you, non-parametric models...)
- ...But broadly speaking, the statement is true

From an optimization point of view:

- Classical ML has focused on large-scale, unconstrained problems
- ...And for good reason!
 - Large input spaces
 - Large parameter spaces
 - Expensive cost functions

Let's start with an observation:

Machine Learning is Optimization

However, that's a bit reductive:

- What if our ML problem has a non-trivial structure?
- What if we have external knowledge?
- What if there are physical laws, or regulations?
- What if we want to use data to help with a decision problem?

Then, we may want to take a broader view...

...Since other optimization techniques may be of help!

This course is about seeing ML and CO as a whole:

- Focus: integration of Machine Learning and Constrained Optimization
 - Emphasis on modeling aspects and knowledge integration
 - No mention of acceleration techniques
- Three parts:
 - Constrained Optimization for Data Mining
 - Handling constraints in Machine Learning
 - Hanlding Machine Learning models in optimization
- It will be far from a complete overview
 - Bias on my own research
 - I'll do my best to provide pointers
 - ...And starting points for related areas

Course material

- Jupyter notebooks for everything
- RISE plugin for the presentation mode
- <u>Docker</u> container for ease of configuration/isolation/reproducibility
- Each part in different github repository

Lectures are meant to be executed

For doing it locally, you need to:

- Install <u>Docker</u> and <u>Docker Compose</u>
- Clone the git repository
- Open a terminal on the cloned repository
- Run docker-compose up and follow the instructions

The first time it will take a while (to download the base image)

Let's check our directory structure:

```
In [6]: !ls -1 ..

total 24
-rw-r--r- 1 1000 1000 888 Nov 13 21:49 Dockerfile
-rw-r--r- 1 1000 1000 2639 Nov 30 09:07 README.md
drwxrwxrwx 2 1000 1000 4096 Nov 29 21:52 data
-rwxrwxrwx 1 1000 1000 734 Sep 19 13:27 docker-compose.yml
drwxrwxrwx 5 1000 1000 4096 Nov 30 09:56 notebooks
drwxrwxrwx 2 1000 1000 4096 Nov 30 10:03 pdfs
```

- The data folder is meant for datasets & co.
- The pdf folder contains PDF exports for the notebooks
- In the Dockerfile you have the container setup instructions
- The docker-compose.yml file specifies how to run the container
- The README.md file contains instructions on how to run the lectures

Let's check our directory structure:

```
In [7]: !ls ../notebooks
        '00. Introduction to the Course.ipynb'
        '01. Constraint Optimization for Data Mining.ipynb'
        '02. The Alternating Direction Method of Multipliers.ipynb'
        '03. Solving the Path Formulation.ipynb'
        '04. Consolidation as Regularization.ipynb'
        '05. Mixed Integer Linear Programming.ipynb'
        '06. Solving the Consolidation Problem.ipynb'
        '07. From Pricing....ipynb'
        '08. ... To Column Generation.ipynb'
        '09. Constraints in the Master.ipynb'
        '10. Constraints in the Subproblem.ipynb'
        '11. Maximum Wait Pricing Model.ipynb'
         assets
         print-pdf.sh
         rise.css
         util
```