Exercises, Project, Exam

Welcome, from the TA team!

data analytics lab

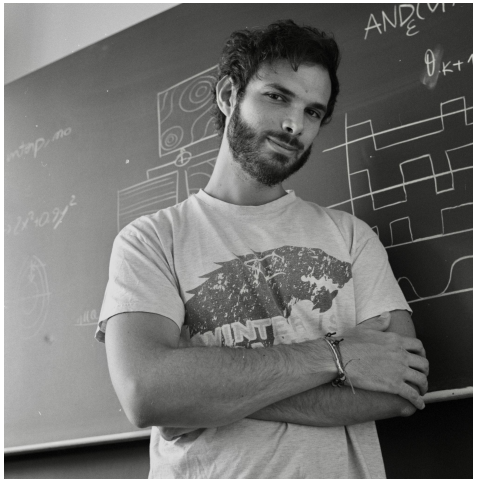
- We are **here to help** you get the most out CIL
  - During exercise session, lecture recap and exercise solutions
  - Presentation of literature/baselines/tips for the projects
  - Support on Moodle

- Most of the TAs took CIL themselves ;)
  - We know which topics are harder to understand 🥵 and will focus on them!
  - We will make sure everyone has the same tools to write a great project and successfully pass the exam 🚀

Head TAs

please contact us for high-level questions/problems

Leonard Adolphs
leonard.adolphs@inf.ethz.ch



Antonio Orvieto
antonio.orvieto@inf.ethz.ch

You will meet the other TAs later!!

Weekly Work

Links to Zoom rooms and recordings, moodle at:
http://www.da.inf.ethz.ch/teaching/2022/CIL/

1) Follow/watch the lecture each **Friday at 10** (week t)

2) Study the lecture, test yourself on the provided exercise sheet

3) Join the Zoom Exercise session each **Friday at 16** (week $t+1$)
   a) Lecture of the previous week is discussed
   b) Selected exercises are solved including live coding if it makes sense
   c) Live coding: TAs will show applications of the theory seen in class
   d) Project overview/baselines

No graded assignments

4) Anything unclear?
   a) Write a Moodle post (best option – fast detailed answer by TAs, available to all)
   b) or join the Zoom Q&A on Thursday at 14 (week $t+2$)

No Exercise Week 8!

Week 9

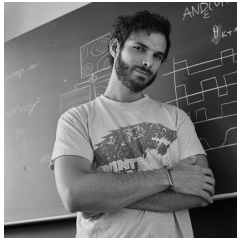
Leonard Adolphs
Intro to PyTorch

Week 13

Xiang Li
Project 1 Baseline (April 1)

Week 19

Dario Pavllo
Neural Nets 1

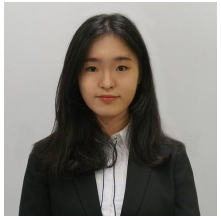


Week 10

Antonio Orvieto
Linear Algebra Recap

Week 14

Xinrui Lyu
Matrix Completion 2



Week 20

Linfei Pan
Project 3 Baseline (May 20)

Week 11

Lorenzo Noci
PCA and SVD

Easter Break

Week 17

Rita Kuznetsova
Latent Variables

Week 21

Sotiris Anagnostidis
Neural Nets 2

Week 12

Luca Biggio
Matrix Completion 1

Week 18

Harish Rajagopal
Project 2 Baseline (May 6)

Week 22

Gregor Bachmann
Generative Models

Exam

- During summer exam session (August)
  - There will be a re-take exam in the winter session
- 120 minutes, 120 points maximum (not all points needed for a 6), no written aids
- No coding, exercises similar to the sheets (samples on website)
- Will not ask you complex formulas by heart, do not worry ! :)
- Project grade usually helps in borderline cases :)

Final Grade = 70% exam + 30% project

Project

- Project work is mandatory. **Groups of three or four students**. No exceptions.
- Project Submission deadline: **August 1st**
- **3 Options (you only have to do 1 project)** → 3 Kaggle competitions.
- You will get a dedicated tutorial session for each project, as soon as relative lecture material is covered.
- Need to write a 4 pages report + submit the code.

Working in groups is a fundamental experience but..

There might be problems due to the fact that CIL is not mandatory for everyone!

- As soon as you have decided to take CIL, look for a group. Post on Moodle!
- Finalize groups by end of May, so that you are sure your colleagues are not dropping out.
- Rule of thumb: construct 4-people groups, just in case ;)
- If you are 4 and 2 of you drop out, you will likely be merged to another group. This is not ideal, so please inform your team members well in advance!
- If you have any other problem or feel workload is unbalanced in the team, contact the head TAs.

Project grading

- **Project Grade** = 30% Kaggle score (private test set) + 70% report and ideas quality
- The Kaggle score will be translated to a grade based on the scores of a baseline solution and the competition winner
  - To get a 4 on the project, you need to do a bit better than the baseline provided in class.
  - Kaggle grade = 4 + 2* (x- baseline_score)/(max_score - baseline_score).
- You can
  - Base your implementation/idea on techniques not seen in class.
  - We like comparisons, please test your novel technique with standard approaches. Be scientific.
  - You can also train your model on additional datasets

Final Grade Composition

Re-submission

Been here before?

- You can take over the grade from the previous year
- But you have to communicate this to one head TA
- You also have the choice to redo the project and join a group

Option 1: Collaborative Filtering

discussed April 1st

	Ben	Tom	John	Fred	Jack
Star Wars	?	?	1	?	4
WallE	5	?	3	4	?
Avatar	3	4	?	4	4
Trainspotting	?	1	5	?	?
Shrek	5	?	?	5	?
Ice Age	5	?	4	?	1

- **Data**: ratings of users to a bunch of movies. Not all users rated all movies.
- **Goal**: predict unrated user-movie pairs (matrix completion).
- **Tools**: matrix completion, SVD, alternating least squares, etc..

Option 2: Sentiment Analysis

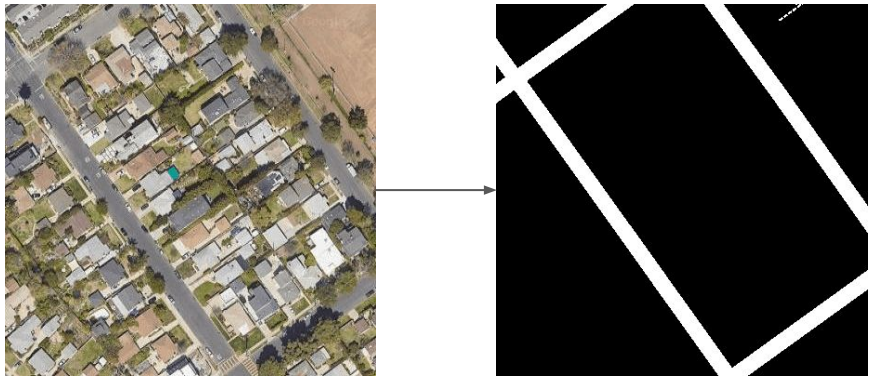
discussed May 6th

Sentiment	Tweet mention
Positive	Maybe I'm mad but I'm now the proud owner of a potentially #bendy #iPhone6, it's so much bigger than the #4s
	Finally got to see an iPhone 6 today. Not revolutionary at all but it's absolutely gorgeous. (And I want one). #iPhone6
Negative	I'm not sure I want it. It's too big to fit in my back pocket! lol #iphone6
	I'm really disappointed with the #iPhone6. It took them 2 years to change the screen & size. Let down.

- **Data**: we provide a large set of training tweets.
- **Ground-truth**: each tweet is labeled as {negative,positive}.
- **Goal**: train classifier to predict polarity
- **Tools**: Transformers, bidirectional LSTMs, word embeddings, etc..

Option 3: Road Segmentation

discussed May 20th



- **Data**: set of satellite/aerial images
- **Ground-truth**: images with pixels labeled as {road,background}.
- **Goal**: train a classifier to segment roads: assign a label {road=1, bg=0} to each pixel.
- **Tools**: Convolutional Neural Networks, Distributed LSTMs, Pyramid Networks, etc..