Computational Intelligence Laboratory

Overview & Organization

Thomas Hofmann

ETH Zurich – cil.inf.ethz.ch

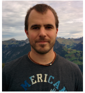
21 February 2020

Section 1

Organization

CIL Team

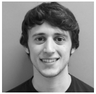
Week 8-10

Yannic Kilcher  Jonas Kohler 

Week 11-13

Antonio Orvieto  Calin Cruceru 

Week 14-15, 18

Dario Pavllo  Giambattista Parascandolo 

Week 19-21

Leonard Adolphs  Gregor Bachmann 

Weekly Schedule: 2 + 2 + 1

- **Lecture**: Fri 8-10, HG E 7
- **Exercises**: Thu 14-16, CHN C 14 **or** Fri 15-17, CAB G61
  - all three exercise sessions are "identical"
  - first hour: pen-and-paper exercise, immediate discussions
  - second hour: group work on programming assignment
  - may need to "load balance", if distribution is skewed
- Voluntary presence time: Mo 11-12, CAB H53
  - TAs help completing programming assignments
- **Webpage:** `http://cil.inf.ethz.ch`
  - only accessible from the ETH network or VPN

Recording, Slides & Forum

- **Lecture slides** are posted before class
- Lectures are automatically **recorded** (voice + screen)
- Q&A platform – `piazza.com`
  - please sign up for an account

Programming Project

- **Project work is mandatory**
- Joint work in groups of **three or four** students
- Solving problems by applying techniques learned in class
- Submitting solution via kaggle.

Been here before? :-/

- You can re-submit previous year's project.
- You have to let us know that you are a "one person group"
- You have the choice to redo the project and join a group

First Week

Introduction to Numpy

- Exercise session, this week

Reading material

- Linear algebra background – on the course website

Grading Criteria

Written Exam

- 120 minutes written exam, closed book **NO WRITTEN AIDS**
- Problems in the spirit of the pen-and-paper exercises

Semester Work

- Develop a **novel solution** for one of the application problems
  - compare with two baseline techniques already implemented in the weekly programming assignments
  - competitive criteria: run-time, accuracy, ..
- Write up in the form of a short **paper**
  - non-competitive criteria: paper review, creativity of solution, quality of implementation

Grading Formula

- **Final examination** – during examination period
- Grading
  - project $< 4.0 \Longrightarrow$ fail
  - project ≥ 4.0 :

$$\text{final grade} = \frac{7}{10}\,\text{exam grade} + \frac{3}{10}\text{project grade}$$

- Again: **NO WRITTEN AIDS** allowed during the exam

Section 2

Goals & Content

Learning Goals

- Acquire and deepen **fundamental concepts** of machine learning
- Implement and compare **techniques** and **models**
- Goal:
  - understand and analyse methods
  - **extend** and **special-case** general methods
  - find best **solutions** to practical problems
- Expected background:
  - introductory class on machine learning
  - mathematics: linear algebra, vector analysis, prob & stats
  - programming: Python (easy to learn)

Contents

- Largely the focus of the class is on **data modeling**, **learning representations**, **unsupervised learning**, and on solving **complex prediction problems**
  - large parts of introductory ML classes: simple classification and regression problems (models, algorithms, theory)
- Specifically (see syllabus)
  - Data reduction (dimension reduction and data clustering)
  - Matrix completion and factorization
  - Latent variable models (mixture & topic models)
  - Sparse coding and dictionary learning
  - Deep neural networks and deep generative models
  - Reinforcement learning (*if time permits*)

Section 3

Projects

Project Overview

- Real world data sets and challenges – you pick one!

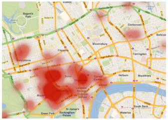
1

preferences recommender

2

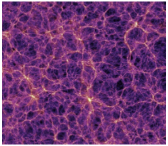
texts & tweets sentiment

3



aerial imagery segmentation

4



cosmology galaxies

- Combine and extend techniques ⇒ novel solution
  - compare to baselines developed during the course
- Produce a write up of your findings ⇒ scientific short paper
  - emphasize experimental protocol, metrics, and reproducability

Project 1: Collaborative Filtering

Recommender system: present items of likely interest to a user

- Products: Amazon, ...
- Movies: Netflix, IMDB, ...
- Music: LastFM, Spotify, ...
- Social Media: Facebook, ...

Collaborative filtering (CF) makes recommendations based on:

- (known) preference of a user towards other items
- collective preferences of other users

Project 1: Collaborative Filtering

Viewers were asked to rate some movies (items):

	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

- Not all viewers rated all movies.
- Goal: predict unrated user-movie pairs (**matrix completion**)
- Should we recommend Fred to watch "Ice Age"?

Project 2: Sentiment Classification

Automatic sentiment analysis to give a machine the ability to understand text and its **polarity**.

- **Data**: we provide a large set of training tweets.
- **Ground-truth**: each tweet is labeled as {negative, positive}.

Positive: "i have the worlds best dad"

Negative: "pouring rain outside . wish i could go out"

- **Goal**: train classifier using **word vectors** to predict polarity

Project 3: Semantic Segmentation

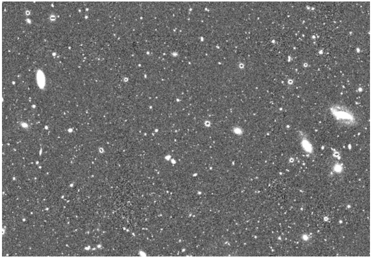
Extract roads from satellite images

- **Data**: set of satellite/aerial images acquired from GoogleMaps
- **Ground-truth**: images with pixels labeled as {road, background}.

- **Goal**: train a classifier to segment roads in these images, i.e. assign a label {road=1, background=0} to each pixel.

Project 4: Galaxy Image Generation

- **Data**: astronomical images acquired from wide field imaging surveys.
- **Ground-truth**: pixels labeled as $\{$background, star, galaxy$\}$



- **Goal**: train a generative model that can generate galaxy images.