Introduction to Machine Learning and Big Data

Exercise

Volker G. Göhler

TU Bergakademie Freiberg

April 11, 2024

Organizational Matters

Exercise

- Thursday each week
- 18:00 until 19:30
- KKB-2097

Dramatis Personæ

-  M. Sc. Volker Göhler
-  Humboldt-Bau - Zimmer 228
-  Volker.Goehler@informatik.tu-freiberg.de
-  03731 39 3113

Learning

- code along
- bring your device (don't connect it to the wire, eduroam is okay)
- flipped classroom

Share

- via Opal

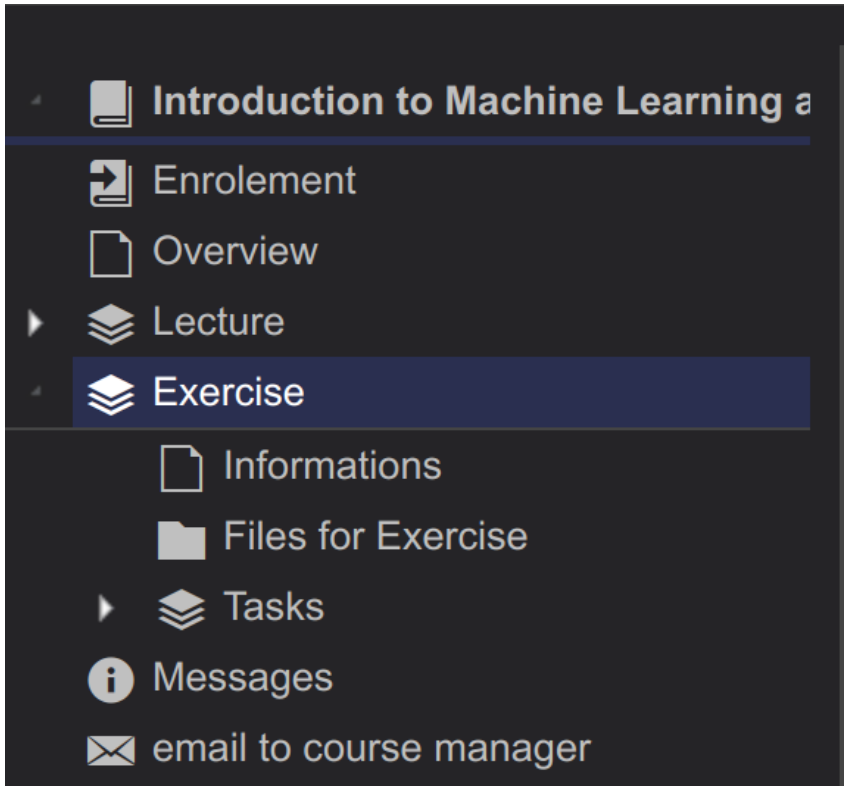

Figure 1: Opal Course Structure

Machine Learning

Software

- Python python.org

Machine Learning

Software

- Python 🔗python.org
- Scikit Library 🔗scikit-learn.org

Machine Learning

Software

- Python python.org
- Scikit Library scikit-learn.org
- jupyter/jupyterlab jupyter.org

Machine Learning?

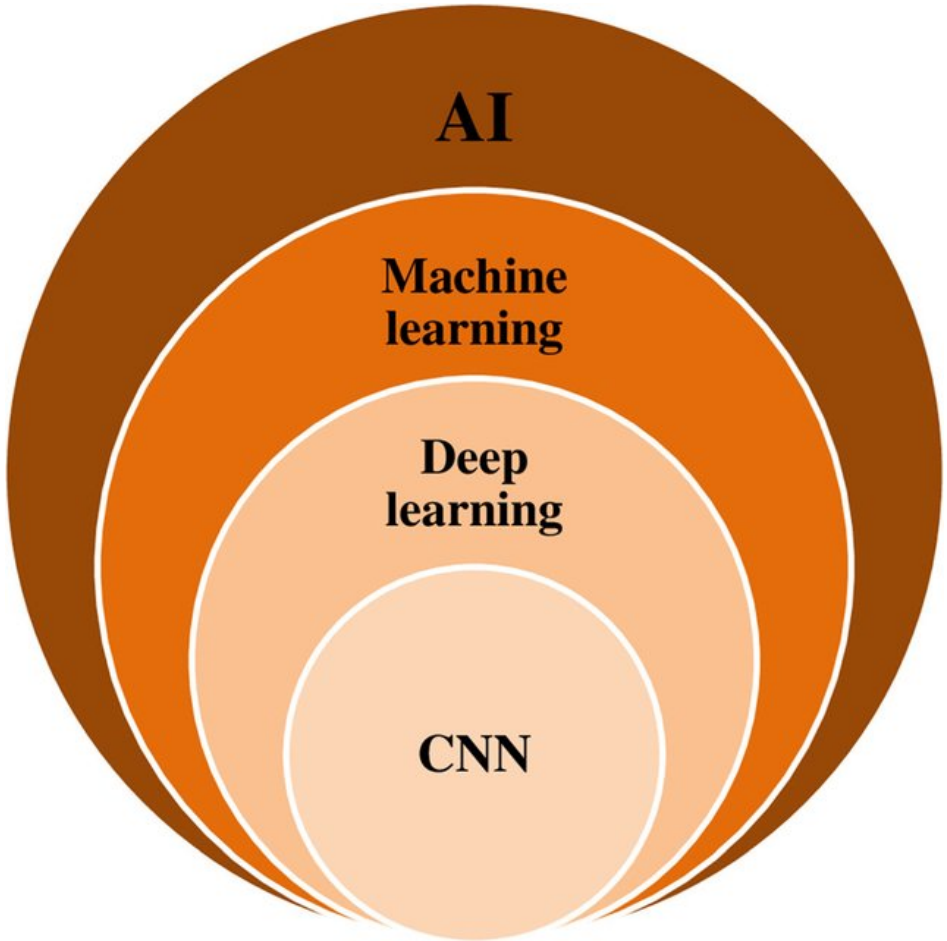

Figure 2: What is AI and ML [@Trusculescu20]

- Which kinds of learning do you know?

- Which kinds of learning do you know?
- Supervised Learning

- Which kinds of learning do you know?
- Supervised Learning
- Unsupervised Learning

- Which kinds of learning do you know?
- Supervised Learning
- Unsupervised Learning
- Semi Supervised

- Which kinds of learning do you know?
- Supervised Learning
- Unsupervised Learning
- Semi Supervised
- Self Supervised

- Which kinds of learning do you know?
- Supervised Learning
- Unsupervised Learning
- Semi Supervised
- Self Supervised
- Reinforcement Learning

No Free Lunch Theorem

- There is no free lunch. [@Wolpert97]

No Free Lunch Theorem

- There is no free lunch. [@Wolpert97]
- All optimization algorithms perform equally well when their performance is averaged over all possible objective functions.

Explanation

Roughly speaking we show that for both static and time dependent optimization problems the average performance of any pair of algorithms across all possible problems is exactly identical. [@Wolpert97]

ML flow

- How do we design a Machine Learning Model?
- What is necessary?

ML flow

- study data
- select model
- train it
- predict
- rinse and repeat

Running Example

jupyter lab example codes

What it is?

- work sheet environment

Where to get the examples?

- work sheets are in the opal share (in the notebook folder)

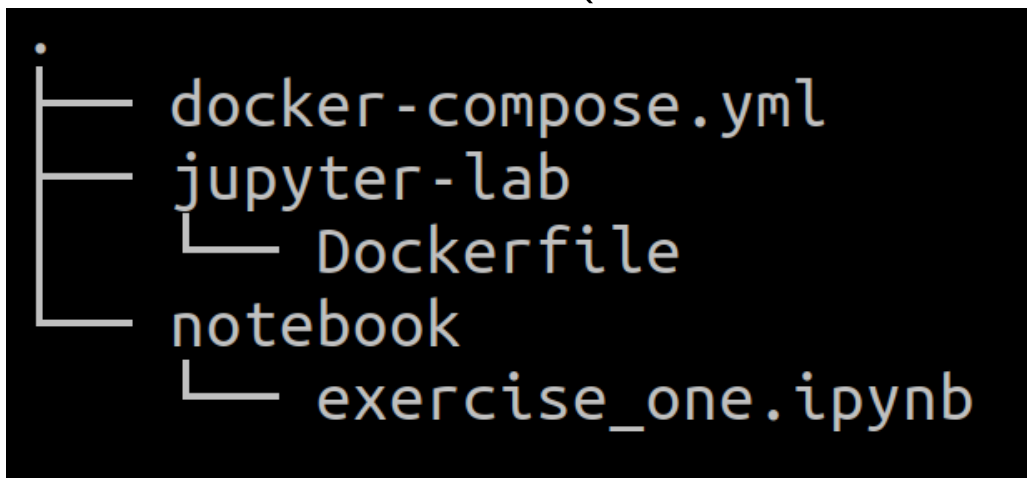

Figure 3: overview of the example structure

Docker

?

How To Run?

at home:

- in folder run: docker-compose up
- in browser: 127.0.0.1:8895

Password: imldb

or install jupyterlab directly

in pool:

- use anaconda

Task!

Your model performs great on the training data, but generalizes poorly to new instances of data.

- What is happening?
- Name three strategies to mitigate this.

- Opal: Exercise -> Tasks -> Task One
- answer the task and upload the answer in Opal

Thank you!

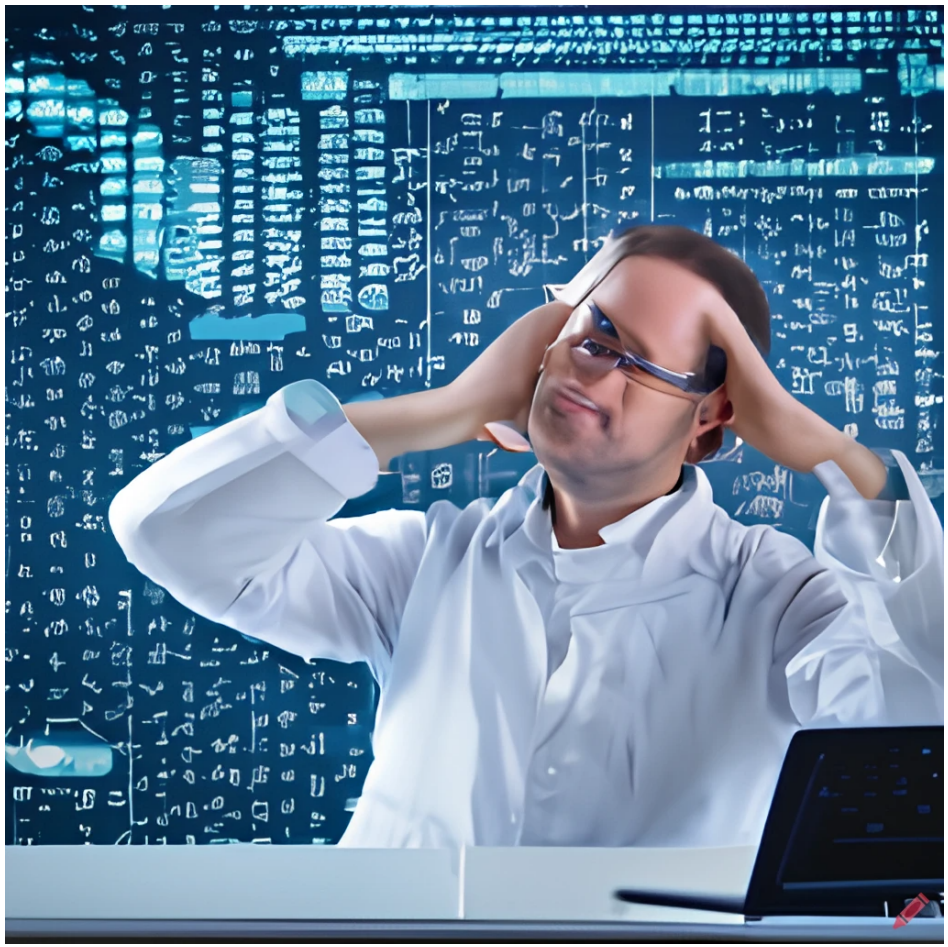

Figure 4: craiyon.com

Reference