

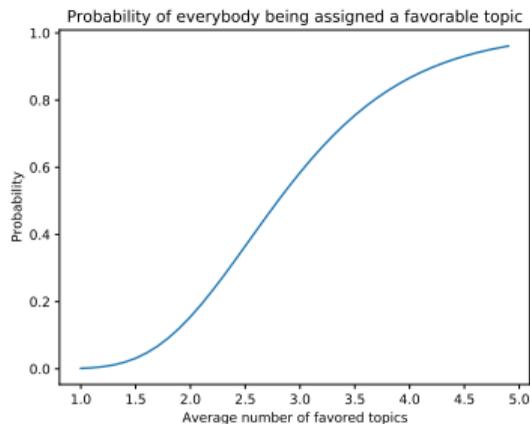


## Seminar Unsupervised Machine Learning - Anomaly Detection



## Today

- Kick-Off Meeting
- Some Formalities
- Short Overview of the Topics



- Choose a couple topics
  - Since we are only a few, you can make these requests quite complicated (I prefer topic 1, but I would also take 3 or 7, except when I can do it in german, then I would prefer topic 12)
- Send your choice to [Simon.Kluettermann@cs.tu-dortmund.de](mailto:Simon.Kluettermann@cs.tu-dortmund.de) (till tomorrow 13.10.2021 23:59)
- You will be assigned one in the next days
- Earlier submission will not be preferred, so take your time

- This seminar is overbooked and there are dozens of students that we were not able to assign a seminar this semester
- Many of you should also have time to hear a Seminar next year, while staying in regular study time.
- So if you're not sure that you want to participate in this course, please consider that you are taking the place of a potentially very motivated student!
- If you want to make place for another Student: No hard feelings, just write an email to [Simon.Kluettermann@cs.tu-dortmund.de](mailto:Simon.Kluettermann@cs.tu-dortmund.de).



## Objective of this Seminar

- Introduction to some fundamental research problems
  - Researching current scientific ideas
  - Understanding benefits and drawbacks of state-of-the-art techniques
  - Writing a clear and concise scientific report
  - Presenting and discussing your findings

→Great start for a final thesis.... →maybe just talk to your supervisor about this



## Timeline

- 1 Presentation in Class (Blockseminar, 26-28.01.2022 13:00-17:00)
  - 2 Discussion of your Findings (afterwards)
  - 3 Writing of your Report (till 30.01.2022 23:59)
  - 4 Criticise the Paper of another random Student (till 06.02.2022 23:59)
- All parts required!
  
  - Everything will be done in english. If this is a problem for you, please write [Simon.Kluettermann@cs.tu.dortmund.de](mailto:Simon.Kluettermann@cs.tu.dortmund.de)



## Tasks of this Seminar

- 1 Choose a couple of topics from our list, you will be assigned to one of them
  - 2 Read and understand the paper given to you
  - 3 Find, read and understand related literature. It is probably impossible to get a good picture about your topic from just one paper
  - 4 Critically analyze the suggested ideas and compare them to the literature
- Final Results:
- Presentation (30min +10min discussion)
  - Written Report (at least 6 Pages double column, ACM template equivalent)
  - Critique of another written report



## Research Culture

This course is Research oriented

- Feel free to ask as many Questions as you want
- If you want to discuss your Topic with somebody, make an appointment with your Supervisor
- the same holds for your Presentation/Report
- Any Feedback is always appreciated



## Topic 1: Anomaly Detection for Monitoring

### Anomaly Detection for Monitoring (Schwartz, Jinka, 2015)

**Supervisor:** Jelle Hüntelmann ([jelle.hüntelmann@cs.uni-dortmund.de](mailto:jelle.hüntelmann@cs.uni-dortmund.de))



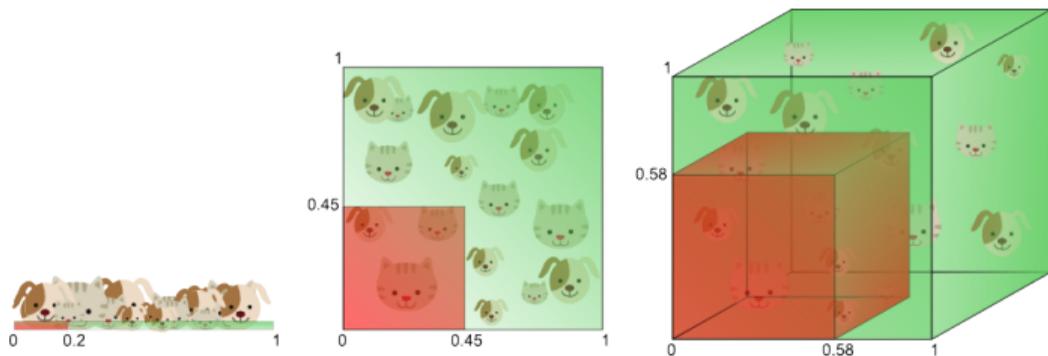
- Time Series analysis
- Less of a Paper, more a Book: great for less experienced People



## Topic 2: High dimensional Data

**A comprehensive survey of anomaly detection techniques for high dimensional big data** (Thudumu et al, 2020)

**Supervisor:** Simon Klüttermann (simon.kluettermann@cs.uni-dortmund.de)

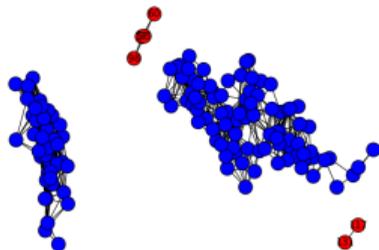


- High dimensional Data becomes hard for AD
- Explain why

## Topic 3: Graph Data

**A Comprehensive Survey on Graph Anomaly Detection with Deep Learning** (Ma et al, 2021)

**Supervisor:** Simon Klüttermann ([simon.kluettermann@cs.uni-dortmund.de](mailto:simon.kluettermann@cs.uni-dortmund.de))



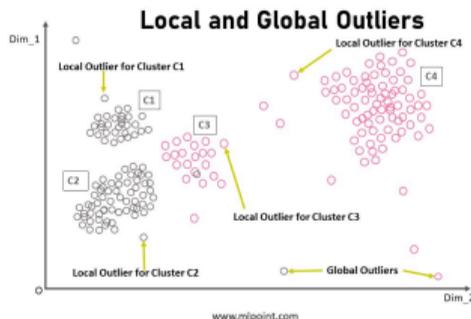
- Graph Data can be super interesting, but also quite complicated
- Either split up into two topics or done by two people



## Topic 4: Density based Outlier Detection

**LOF: Identifying Density-Based Local Outliers** (Breunig et al, 2000)

**Supervisor:** Daniel Wilmes ([daniel.wilmes@cs.uni-dortmund.de](mailto:daniel.wilmes@cs.uni-dortmund.de))



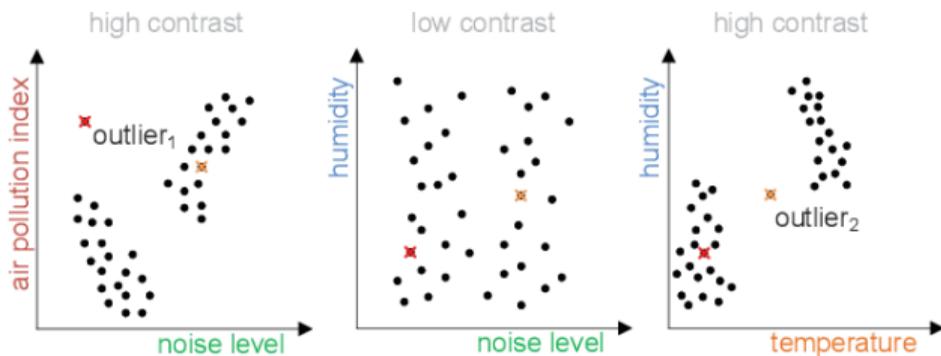
- Classical AD algorithm
- Good for a less experienced student



## Topic 5: Subspace preprocessing

**HiCS: High Contrast Subspaces for Density-Based Outlier Ranking** (Keller, Müller et al, 2012)

**Supervisor:** Daniel Wilmes (daniel.wilmes@cs.uni-dortmund.de)



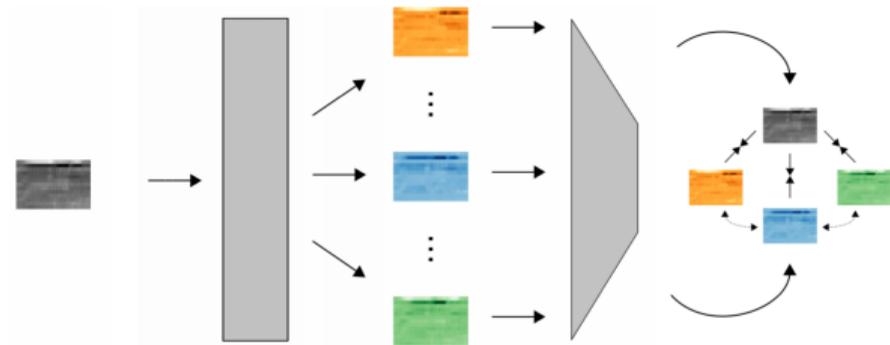
- Filter out useless Features
- Cowritten by Prof. Müller, might be related to a future thesis



## Topic 6: Transformations for Anomaly Detection

### Neural Transformation Learning for Deep Anomaly Detection Beyond Images (Qui et al, 2021)

**Supervisor:** Bin Li (bin.li@tu-dortmund.de)



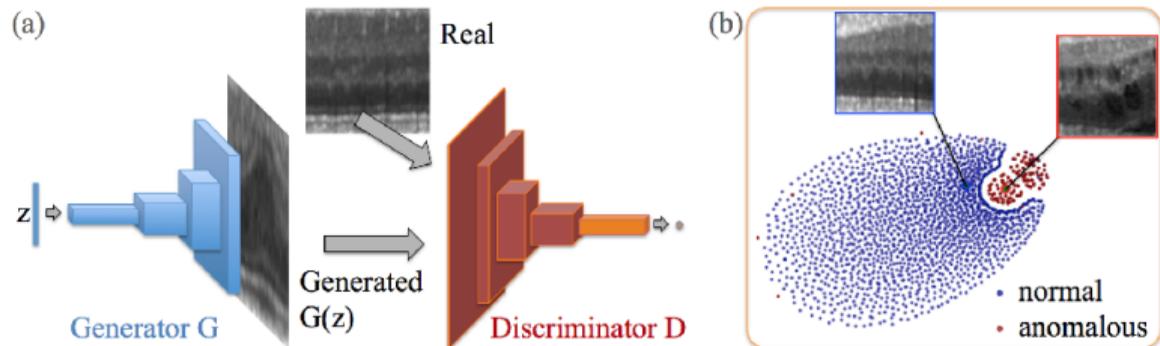
- Pretransformations (like Rotations) can improve ML
- For Time-Series: Make learnable



## Topic 7: GANs for Anomaly Detection

**A Survey on GANs for Anomaly Detection** (Mattia et al, 2021)

**Supervisor:** Simon Klüttermann ([simon.kluettermann@cs.uni-dortmund.de](mailto:simon.kluettermann@cs.uni-dortmund.de))



- GANs generally used for data generation
- Can also be used for AD, how?

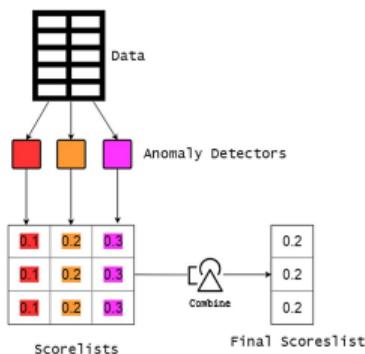


## Topic 8: Anomaly Ensembles

### Unsupervised Anomaly Detection Ensembles using Item Response Theory

(Kandanaarachchi1, 2021)

**Supervisor:** Simon Klüttermann (simon.kluettermann@cs.uni-dortmund.de)



- Combine multiple algorithms
- Could be combined with the following Paper

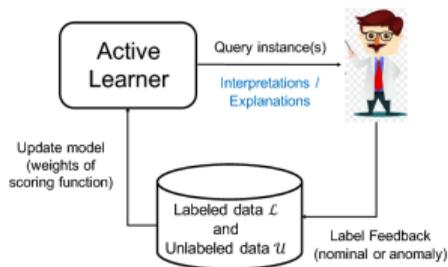


## Topic 9: Active Anomaly Detection

### Active Anomaly Detection via Ensembles: Insights, Algorithms, and Interpretability

(Das et al, 2019)

**Supervisor:** Simon Klüttermann (simon.kluettermann@cs.uni-dortmund.de)



**GOAL:** Maximize the number of true anomalies presented to the user.

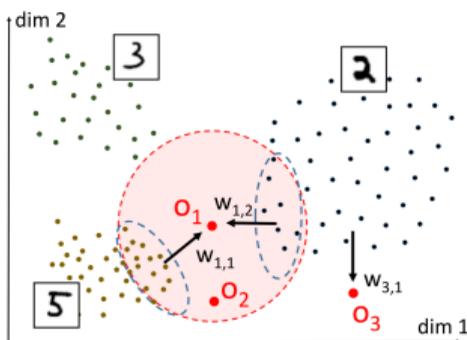
- Active Learning=Learning with humans
- Could be combined with the previous Paper



## Topic 10: Interpretation

### Contextual Outlier Interpretation (Liu et al, 2017)

**Supervisor:** Benedikt Böing (benedikt.boeing@cs.uni-dortmund.de)



- Interpretabile AD: Create Trust
- Understand by surrounding

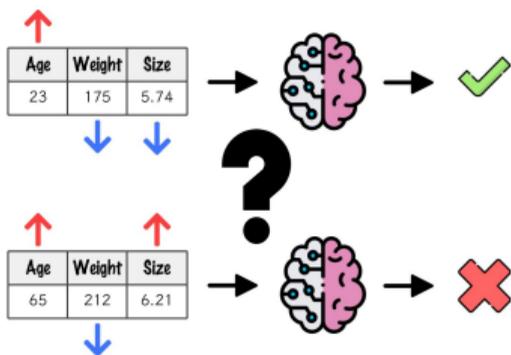


## Topic 11: Explainability

### Additive Explanations for Anomalies Detected from Multivariate Temporal Data

(Giurgiu, Schumann et al, 2019)

**Supervisor:** Chiara Balestra (chiara.balestra@cs.uni-dortmund.de)



- Measure contributions of features
- Your own Literature Research is extra important



## Topic 12: Application: Interpretable Device Failure

**Interpretable, Multidimensional, Multimodal Anomaly Detection with Negative Sampling for Detection of Device Failure** (Sipple et al, 2020)  
**Supervisor:** Bin Li (bin.li@tu-dortmund.de)



- real world data is messy
- detect IOT devices that fail



## Topic 13: Application: Video Data

### **Fast Unsupervised Anomaly Detection in Traffic Videos** (Doshi et al, 2020)

**Supervisor:** Jelle Hüntelmann ([jelle.hüntelmann@cs.uni-dortmund.de](mailto:jelle.hüntelmann@cs.uni-dortmund.de))



- Application Paper
- Complex input data

## Topic 14: Application: Spacecraft monitoring

### **Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding** (Hundman et al, 2018)

**Supervisor:** Daniel Wilmes (daniel.wilmes@cs.uni-dortmund.de)



- Another Application Paper
- This time to supervise Spacecrafts



- 1: AD for Monitoring
- 2: High dimensional Data
- 3: Graph Data (\*2?)
- 4: LOF
- 5: Subspace preprocessing
- 6: Transformations
- 7: GANs for AD
- 8: Ensembles
- 9: Active AD
- 10: Interpretation
- 11: Explainability
- 12: Application: Device Failure
- 13: Application: Video Data
- 14: Application: Spacecrafts
- Any other ideas?

