Can a **Machine** Hear If a **Machine** Is Broken ?

Unsupervised Anomaly Detection by Airborne Sound of Industrial Machinery

data science bootcamp final capstone at neuefische GmbH

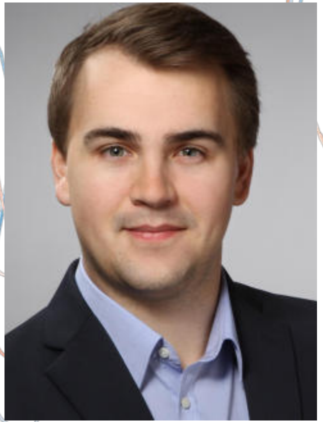
Agenda

1. Scenario / Use-Case / Challenge
2. Dataset and Baseline
3. Audio Feature Extraction
4. Deriving the Detection Model
5. Ensemble & Results
6. Outlook

Introduction Team BA-HanseML

Exploring Machine Learning for Industrial Engineering Challenges

Dipl. Ing. Arne Scheunemann
Automation, Information and
Electrotechnical Engineering
Data Scientist



M.Sc. Bendix Haß
Energy and Process Engineering
Data Scientist

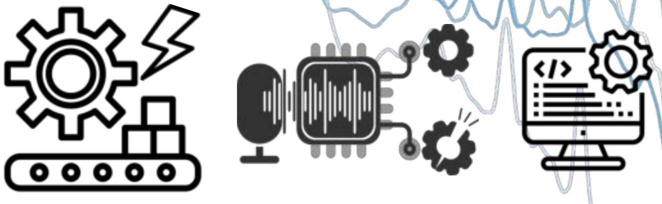
GitHub https://github.com/BA-HanseML/NF_Prj_MIMII_Dataset

Scenario Use Case - Product vision

- Detecting abnormal operation in machinery
  - Part of Condition Monitoring
- Detecting by airborne sound is
  - challenging and normally done by technician
  - benefits for installation - no down time, no glue, no screws, no welding ...
- How would it be used:
  - place the microphone near machine
  - wait for some time (30 min - 2h) to train the AI on normal
  - connect it to the fieldbus of your plant supervision
  - receive a alarm if something abnormal is recorded.
  - yes ... it’s like a smart babyphone for machines parts

Placing and Calibration/Train Time

Auto Detection -> Remote Supervision



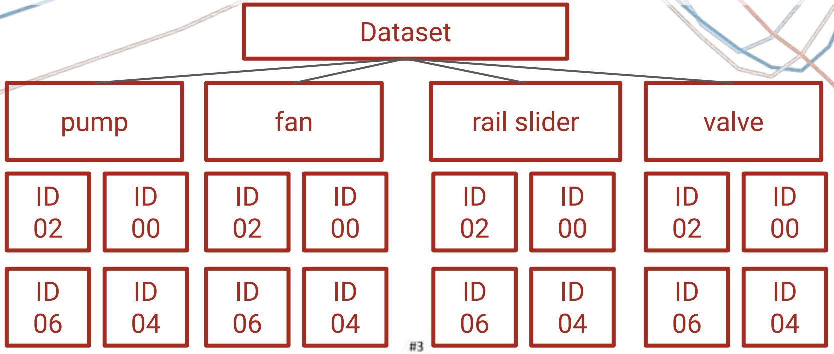
icons: www.flaticon.com

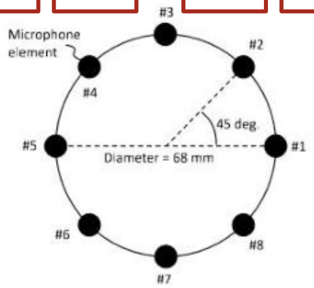
Challenges

- Unsupervised Anomaly detection
  - calibration on normal - only normal is known !
  - there can be many forms of normal
- Failure modes of machines are diverse
  - failure modes are very different from machine to machine
  - much more failure modes than normal modes are possible
  - not all failures are clearly audible
- Airborne Sound - Acoustic
  - Background noise - and it can change (but it is maybe not the machine that changed)
  - Reverb in the room
  - Direction of sound
  - Distance
  - ...

MIMII Dataset

- Malfunctioning Industrial Machine Investigation and Inspection
  - is made available by Hitachi Ltd. under a Creative Commons license (CC BY-SA 4.0)





- 4 machine parts, 4 variants each machine
- 3 levels of noise (SNR 6dB to -6dB)
- ca. 1000x10sec audio file each variant and noise level (ca. 150 abnorm)
- 16kHz sampled
- each audio file has 8 channels
  - microphone is a ring of 8 mics

Machines

examples pictures of similar, but not the actual recorded devices ..

centrifugal pump

solenoid valve

slider rail system

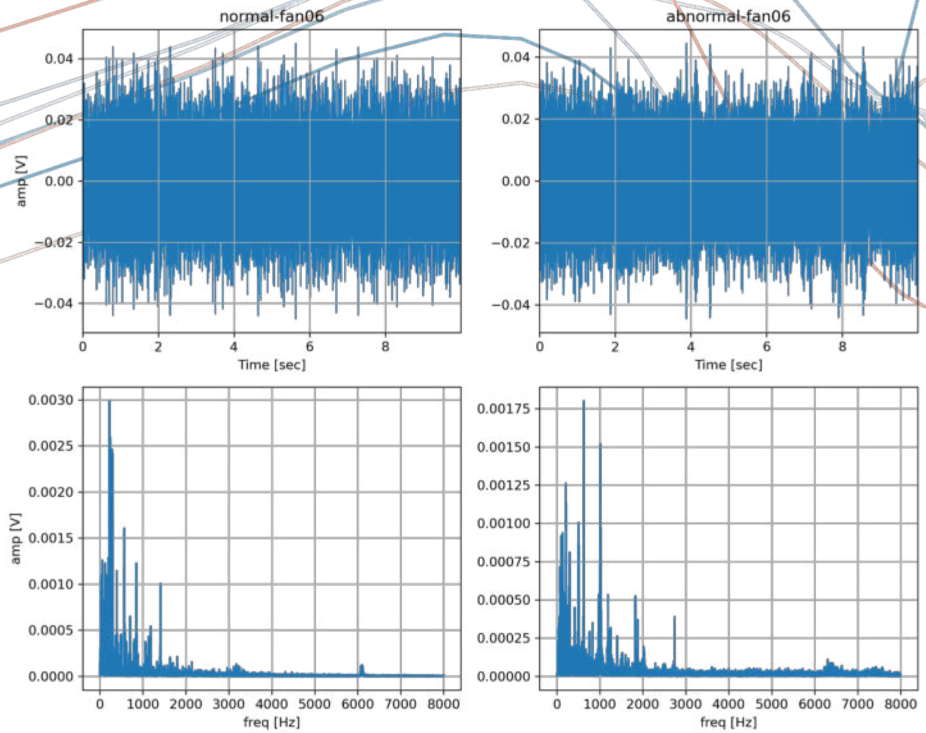
fan

pump: https://www.globustechnomech.com/
fan: https://www.systemair.com/
slider: https://gerald-summers.co.uk/linear-motion/
valve: https://ussolid.com/

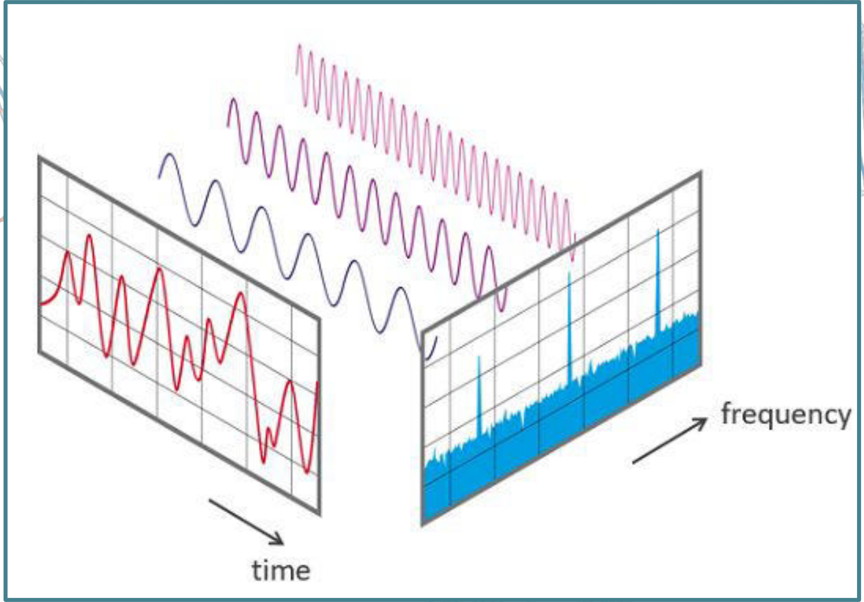
MIMII Baseline

- **Paper:** MIMII DATASET: SOUND DATASET FORMALFUNCTIONING INDUSTRIAL MACHINE INVESTIGATION AND INSPECTION by Harsh Purohit, Ryo Tanabe, Kenji Ichige, Takashi Endo,Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguch
  - describes the data set
- **Source Code:** https://github.com/MIMII-hitachi/mimii_baseline
  - **A solution with a single autoencoder on MEL spectrum frames is given**
  - **And the result of ROC-AUC score for that baseline performing**

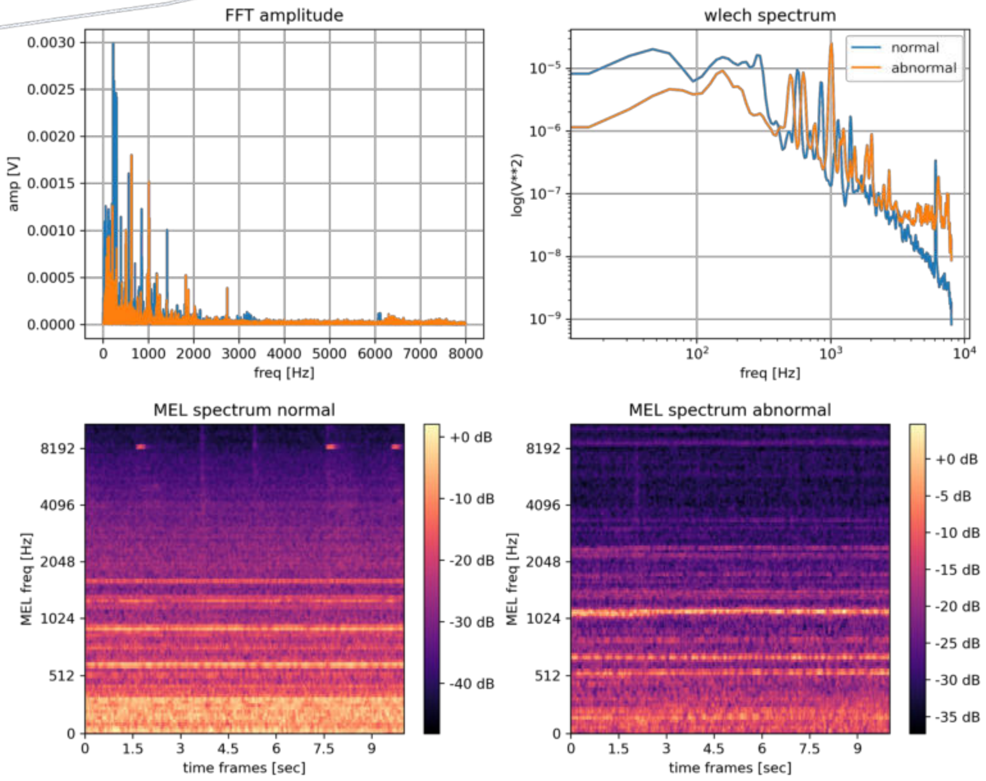
Feature extraction - Time to Frequency



Fourier Transformation

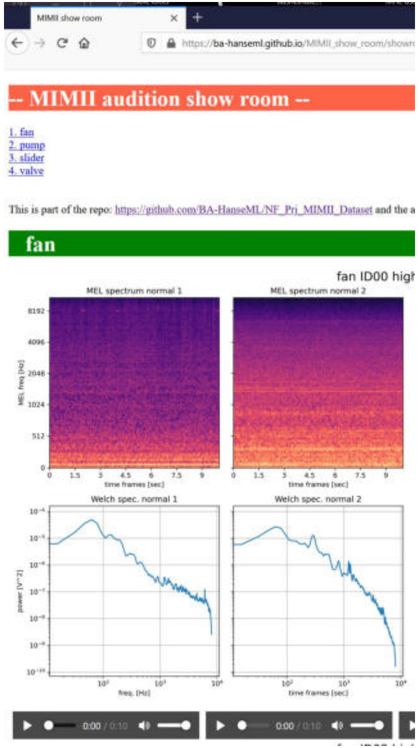


Feature Extraction ML input



- moving to the frequency room helps in making out characteristics!
- Smoothing frequency room with welch spectrum (PSD)
- Time framing and scaling -> 2D MEL spectrum
- other would be possible:
  - plain STFT 2D
  - Cepstrum 1D
  - cepstral cepstrum - >MFCC
  - ...

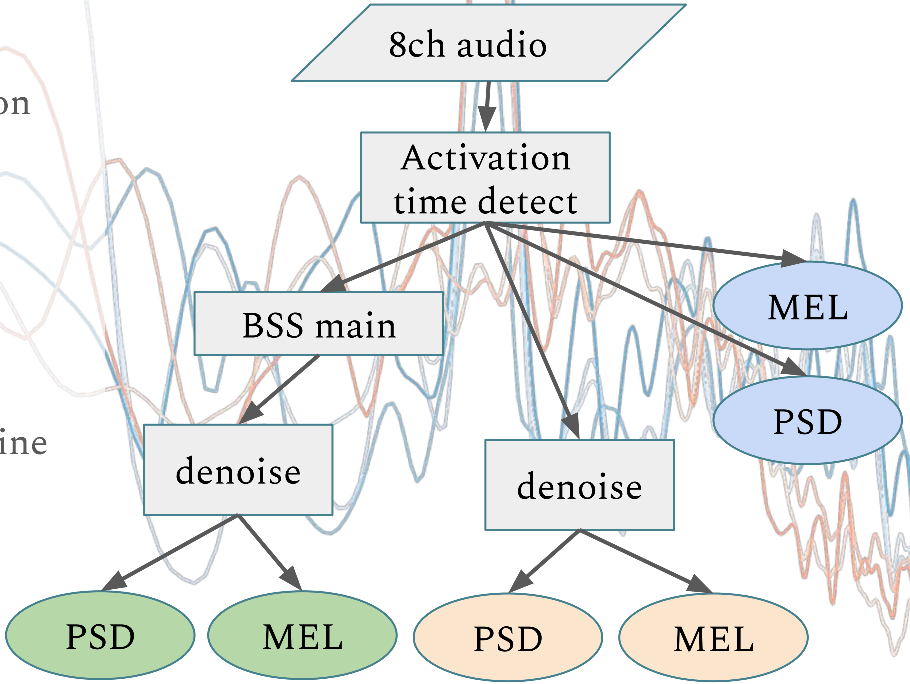
Small Audition in the Showroom



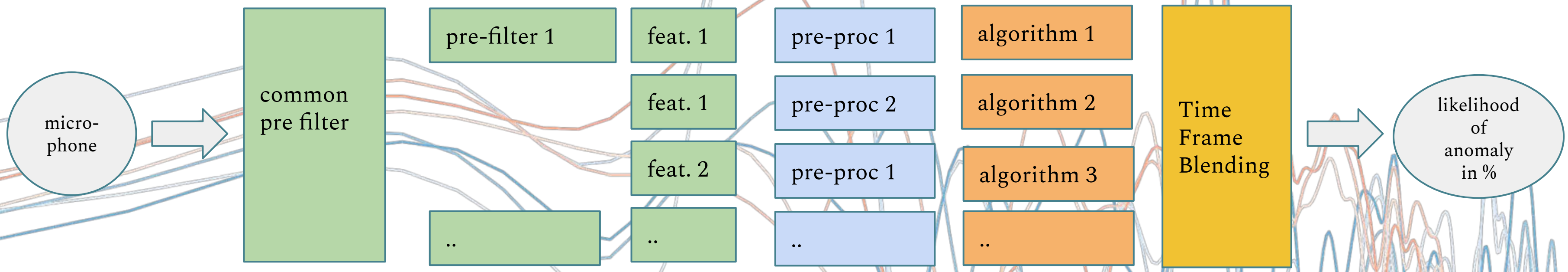
https://ba-hanseml.github.io/MIMII_show_room/showroom.html

Feature Extraction Filter Diagram

- Blind mono Denoise (BD)
  - denoising on the 2D spectrum by convolution
- Activation time detect
  - sporadic machines like the valve
  - deactivate on inactive time frames
- Blind Source Separation (BSS)
  - FastICA individual component analysis
  - naive approach two 2 sources = noise+machine
- Other (not applied but explored)
  - DOA direction of arrival
  - advanced BSS
  - adaptive filter for denoising
  - Blind De-Reverb (delay estimation)



Methodology



- Model Building: ID00 and ID02 50% of the dataset:
  - is used in pick the best algorithms and pre-processing for the ensemble
  - is used to define the hyper parameters of the algorithm
  - is used to find best available feature extraction, common pre-filter, pre-filter and feature
- Model Evaluation: ID04 and ID06 50% of the dataset
  - used to verify that the the picked algorithms and the hyperparameters work
  - derive the final performance results

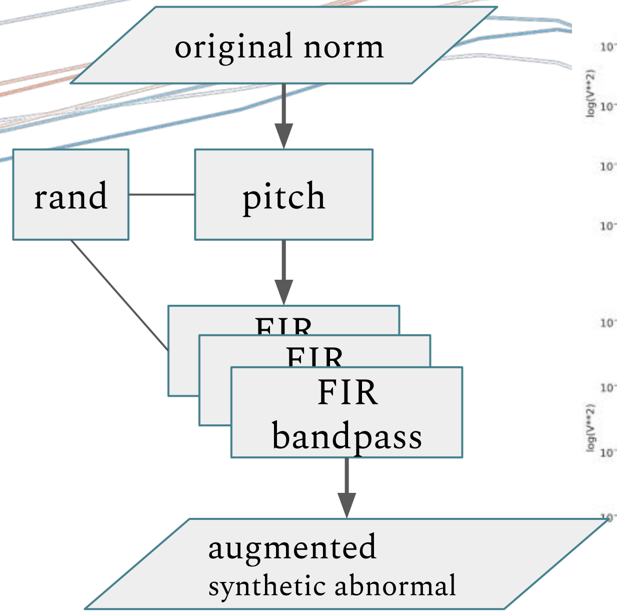
Machine Learning Score

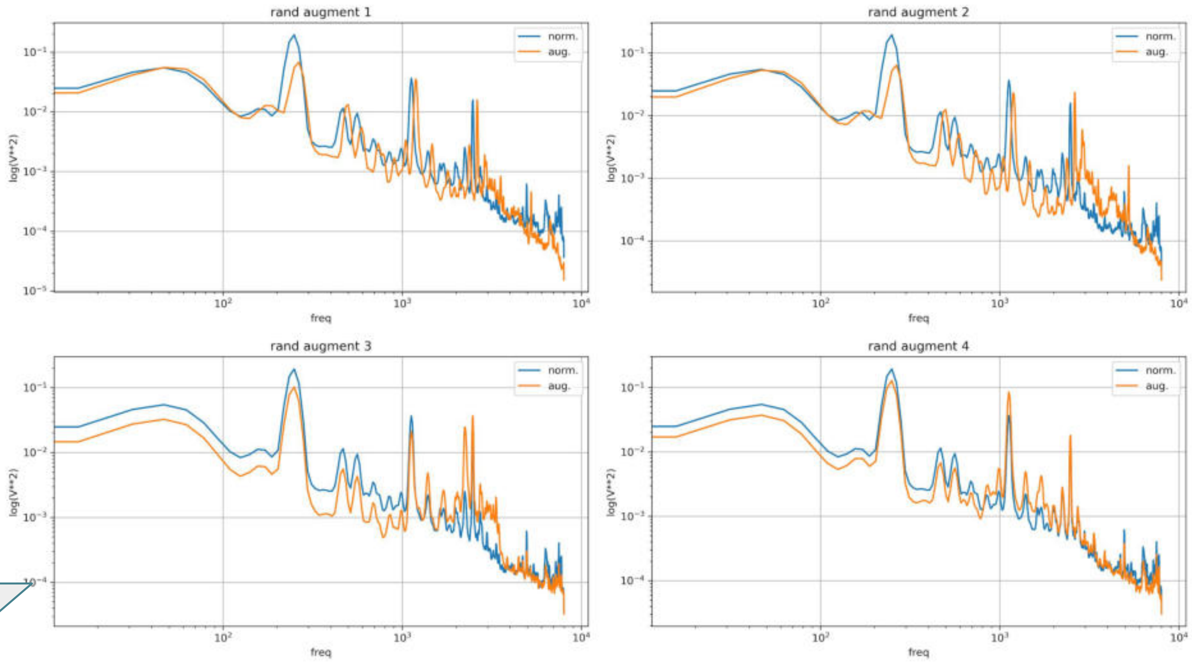
- How to best describe the quality of a model in one digit?
- Tradeoff: False Alarms ←→ Missed Events
- **ROC-AUC Score:**
  - independent of tradeoff
  - represents degree or measure of separability
  - It tells how much model is capable of distinguishing between classes

must be at least >50%
100% is perfect
>80% is realistic good

Pseudo Supervised - synthetic abnormal

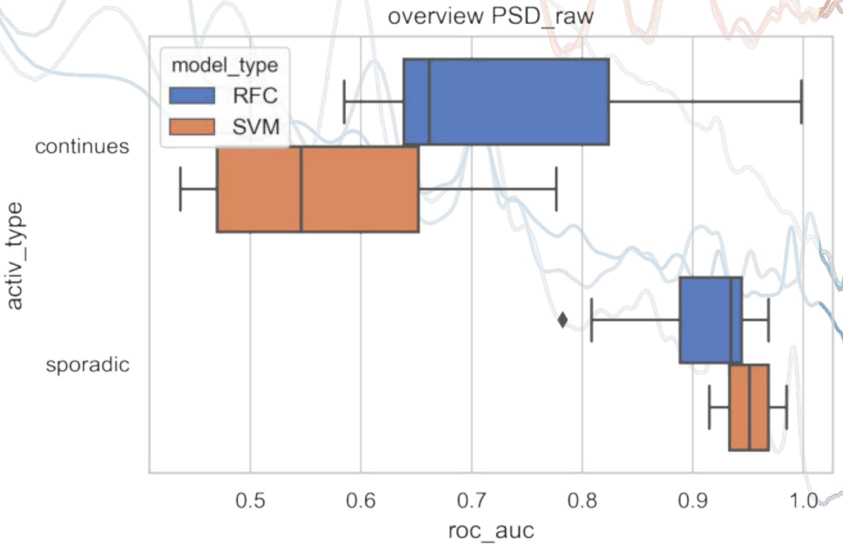
- Distorting normal to create synthetic abnormal - **abnormal is not known**
- train supervised classifier on synthetic abnormal, and validate on real abnormal ...





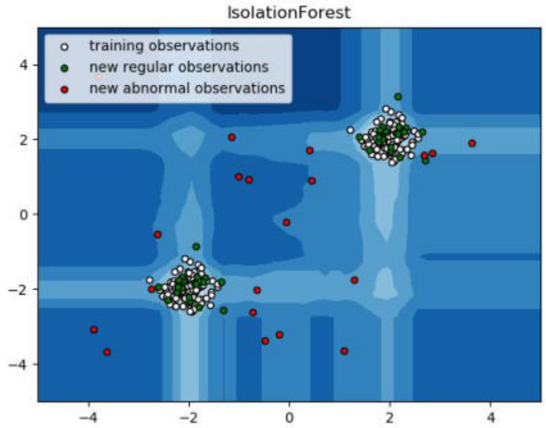
Pseudo supervised - classification

- Multiple classic algorithm tested
  - Random Forest, Logistic Regression, Support vector machine, K nearest neighbour, ...
- The approach is promising
  - general better than the baseline
- Discovered an optimisation limit
  - because the training is overfitting
  - the synthetic anomalies are to easy to spot
  - more work is needed ...
- For sporadic activity types the SVM is a candidate for the ensemble



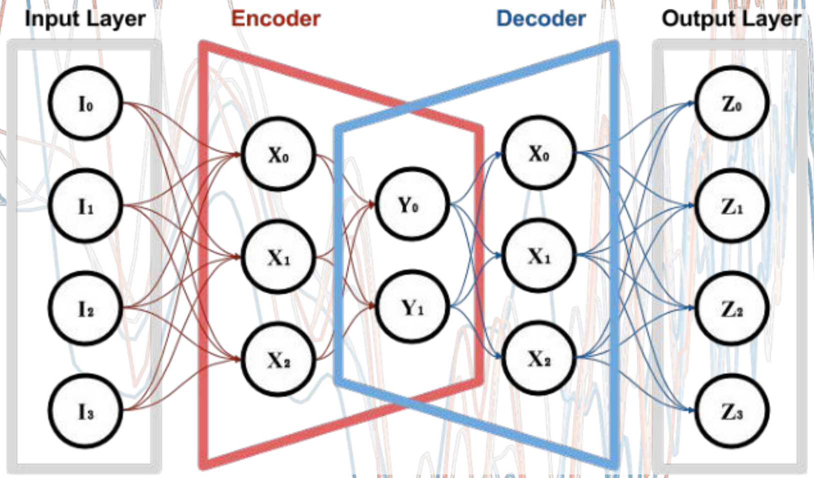
Stochastic Models

- Gaussian Mixture & Elliptic Envelope
  - models represent the stochastic distribution of the data
  - anomalies will be detected by likelihood that they belong to the distribution
- Isolation Forest
  - splits data into random subsets
  - anomalies can be described more easily
  - candidate for the ensemble
- In general:
  - performance is sensitive to noise
  - profit from denoising
  - fast training process
  - small model sizes



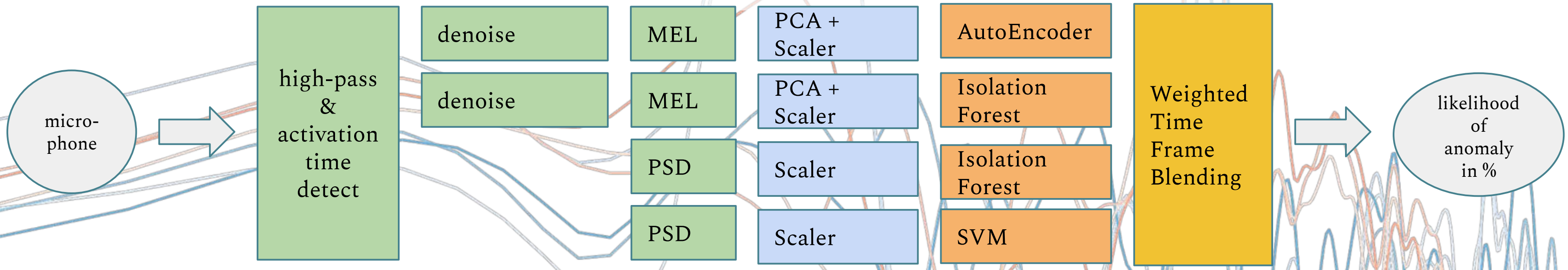
Autoencoder

- artificial neural network model
- is being trained to reproduce the data (data as input *and* as target)
- metric is the reproduction error
  - can be transformed to be a certainty of abnormality
- condenses the data into its fundamental components
- outperforms stochastic models
- robust against noise → not nec. denoise needed



Ensemble

- Architecture of the ensemble



- The Ensemble
  - generally performs better than individual models
  - ensembling increases robustness

ROC AUC				
	low noise		high noise	
machine	baseline	ours	baseline	ours
Valve	67%	91%	53%	79%
Pump	81%	99%	68%	88%
Fan	94%	99%	70%	80%
Slider	90%	99%	70%	83%
Avg.	**83%**	**97%**	**65%**	**82%**

Performance Results - Detection

- Results
  - tuned for less false alarms (tradeoff: more missed events)
- Implication for the application
  - false alarms cause distrust
  - too many missed events cause failing of the product
  - at current performance it can never be safety or mission critical
  - helpful addition for remote supervision and root cause analytics and preventive maintenance
- The level of difficulty
  - might be very different in the missed and false alarms
  - this can be confirmed with spot checking
  - if the anomaly is minor maybe the fault is minor as well (naive assumption)

	false alarms	**missed events**
low noise	1.8%	11.3%
high noise	9.0%	21.5%

Future Work

- Feature extraction
  - testing more pre-filter : blind dereverb MPE, direction of arrival DOA filter, noise cancel with adaptation filter on non active time frames, testing more denoising algorithms, using cepstrum and MFCC etc. as features, active time frame multi cluster and by onset in the hilbert envelope. Testing blind source separation on concunated samples.
- Modeling
  - **pseudo supervised** training with augmentation in the loop, trying general advisory neural network by focusing on the discriminator training. Exploring U-Net to mark abnormalities in spectrum.
  - **stochastic** boosting stochastic models to find subsequent distributions of residuals, research deployability of model types
  - **Autoencoder** investigate more complex network structures and DSP-embedding layers, implement LSTM-Autoencoder
  - **Ensemble** optimize ensemble blender, additional models

Thank you for the attention

GitHub: https://github.com/BA-HanseML/NF_Prj_MIMII_Dataset