MINIMAL-CONFIGURATION ANOMALY DETECTION FOR IOT SENSORS

A Systematic Analysis of Deep Learning Approaches

Clemens Heistracher
Anahid Jalali
Bernhard Haslhofer
AIT - Austrian Institute of Technology

Axel Suendermann
Jana Kemnitz
Siemens AG Austria

Sebastian Meixner
Daniel Schall
Siemens AG Austria

MOTIVATION

Maintenance and repair of equipment are responsible for up to 30% of production costs.

Traditional Condition Monitoring and Predictive Maintenance approaches are expensive and only viable for mission-critical assets.

Industry requires tools to avoid unnecessary maintenance and to prevent failures of production equipment.

MOTIVATION

Modern IoT devices made data acquisition more affordable.

However, ground truth is still hard to obtain in relevant quantity.

Manual annotations

Most data from a healthy state

No data available for new equipment

OVERALL GOAL

Observations

Unlabeled sensor data are often available in large quantities because of low marginal costs.

Labeled sensor data are rare due to expensive manual interventions and annotations.

Industry requires

Unsupervised Anomaly Detection

Minimal configuration effort

→ Unsupervised anomaly detection for universally deployable IoT sensor tag

METHOD

1. **Dataset Creation**
   - Use IoT sensor Tag.
   - Rotation Pump.
   - Lab environment.

2. **Baseline Definition**
   - Supervised Anomaly Detection.
   - Does the dataset work?

3. **Experimental Setup**
   - Unsupervised Anomaly Detection.

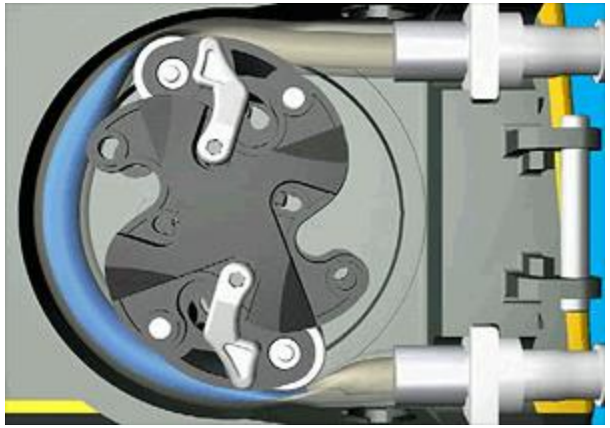
DATASET CREATION

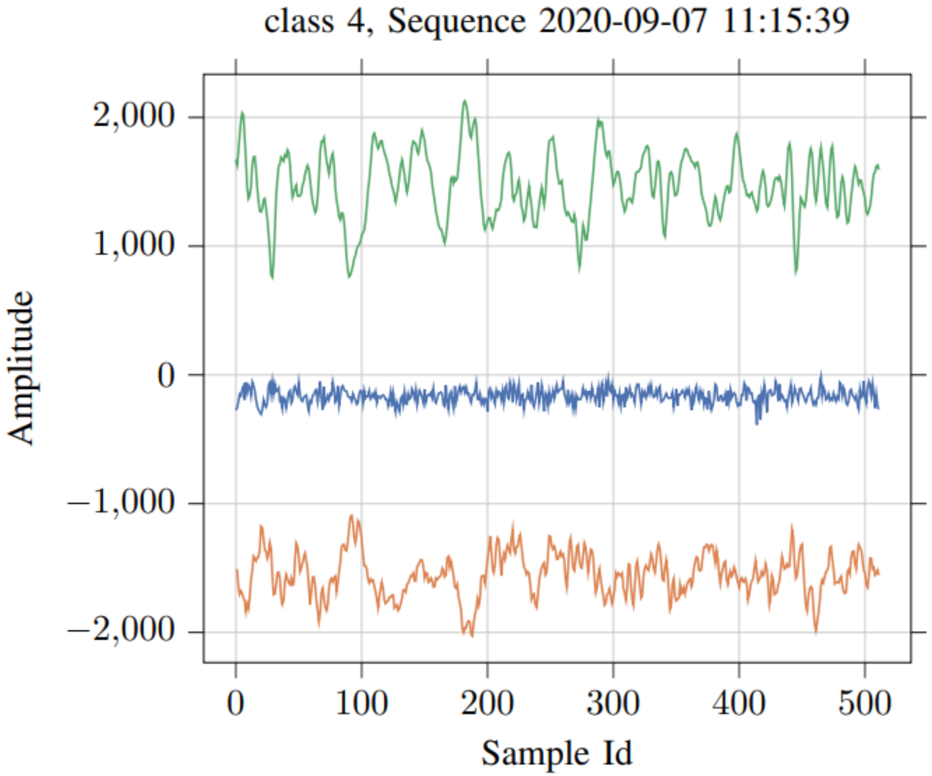
AUSTRIAN INSTITUTE OF TECHNOLOGY
TOMORROW TODAY

METHOD

DATASET

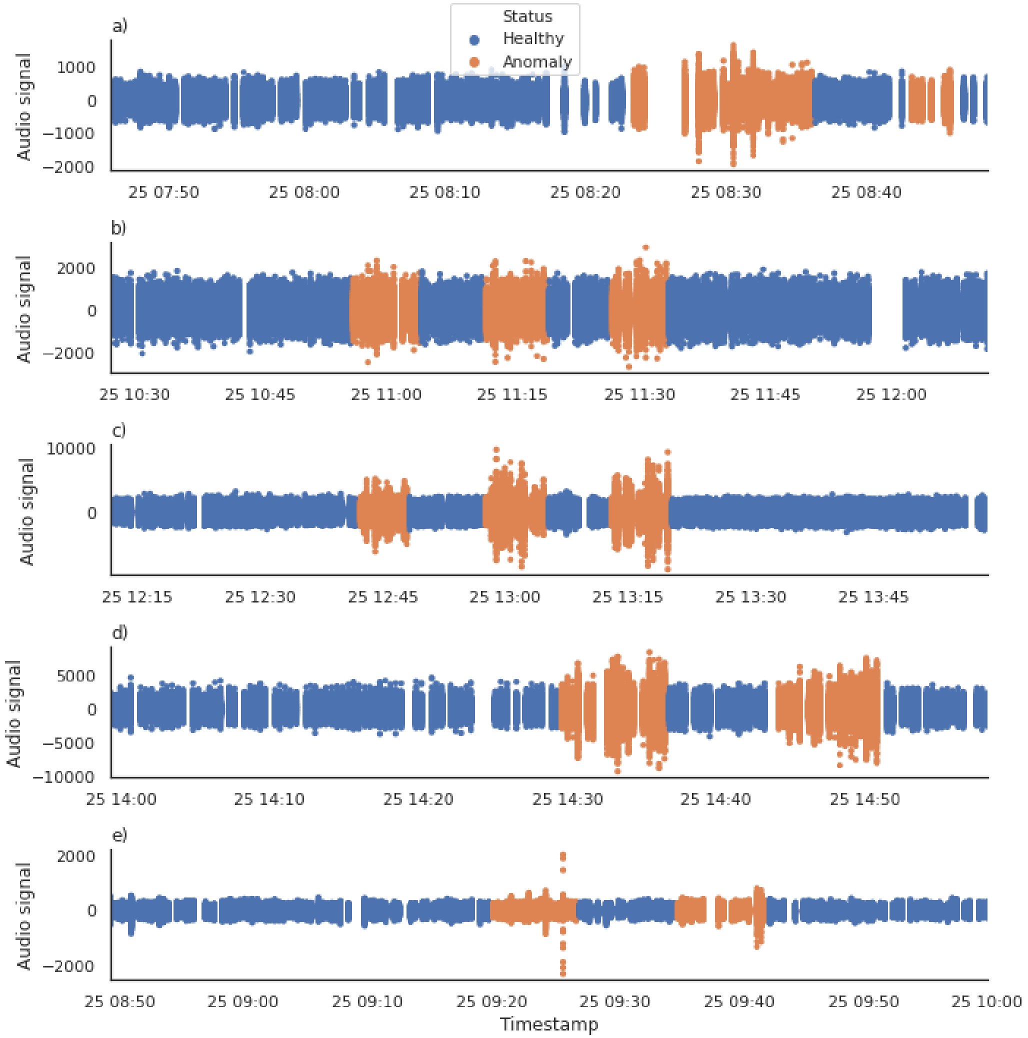
- Recorded in a lab environment
- On peristaltic pumps
- 3D vibration signal (6644 Hz)
- Audio signal (16k Hz)
- Under varying operating conditions





METHOD
DATASET

The recorded dataset for varying operating condition and healthy and anormal state.



BASELINE DEFINITION

BASELINE DEFINITION

METHOD

Supervised anomaly classification	To validate Dataset To select Features
Preprocessing	Norm 3D Vibration signal Balance , Shuffle, Split (70:30) Z-normalize
Features	Short Term Fourier Transform (STFT) Mel Frequency Cepstral Coefficients (MFCC) Mel Spectrogram
Models	CNN, LSTM, KNN, SVM

BASELINE DEFINITION
RESULTS

- Deep Neural networks are compared to conventional machine learning classifiers,
  - Support Vector Machine (SVM)
  - K-Nearest Neighbors (KNN)
  - Long Short-Term Memory Units (LSTM)
  - Convolutional Neural Networks

	Precision	Recall	F1
SVM + FFT	96.38%	83.13%	87.96%
SVM + MFCC	92.04%	89.15%	90.09%
SVM + Melspectrogram	15.66%	14.45%	15.03%
KNN + FFT	**95.78%**	**95.18%**	**95.18%**
KNN + MFCC	95.32%	95.18%	95.16%
KNN + Melspectrogram	83.02%	68.67%	67.96%

	Precision	Recall	F1
Best Benchmark (KNN + FFT)	**95.78%**	**95.18%**	**95.18%**
LSTM + FFT	75.89%	84.33%	79.47%
LSTM + MFCC	82.40%	84.33%	82.75%
LSTM + Melspectrogram	7.02%	26.50%	11.10%
CNN + FFT	**96.69%**	**96.28%**	**96.38%**
CNN + MFCC	52.70%	57.83%	47.24%
CNN + Melspectrogram	22.73%	45.78%	29.82%

BASELINE DEFINITION
SUMMARY

Supervised approaches achieve high accuracy on our dataset

Best results: 0.96 F1-Score

Best model is convolutional neural network

Best feature is FFT

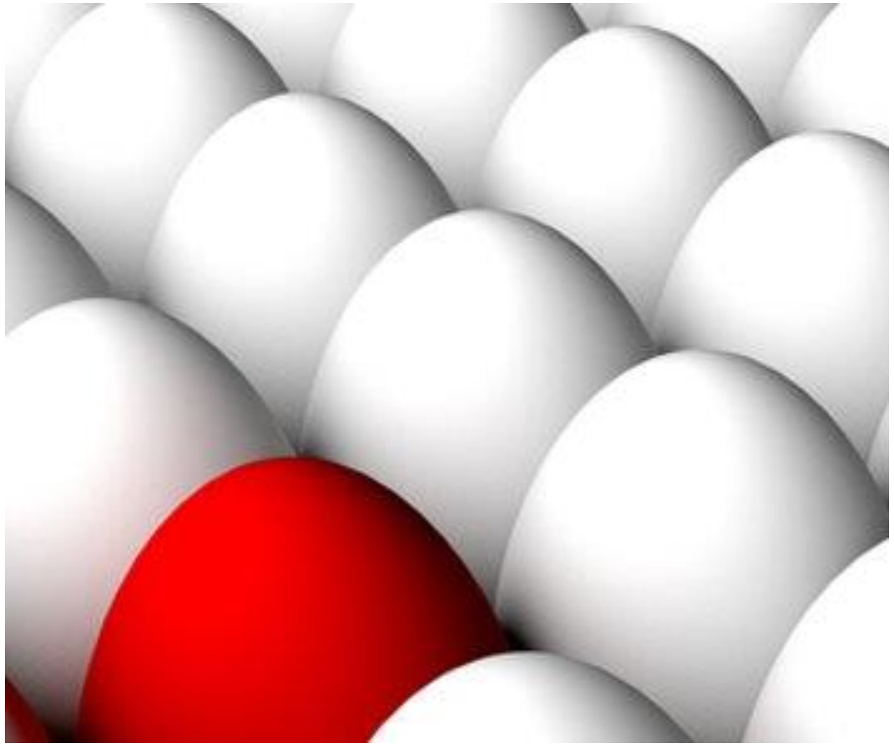
Model + Feature	Precision	Recall	F1
(KNN + FFT)	95.78%	95.18%	95.18%
CNN + FFT	96.69%	96.28%	96.38%

EXPERIMENTAL SETUP

Unsupervised Anomaly Detection

METHOD
UNSUPERVISED ANOMALY DETECTION

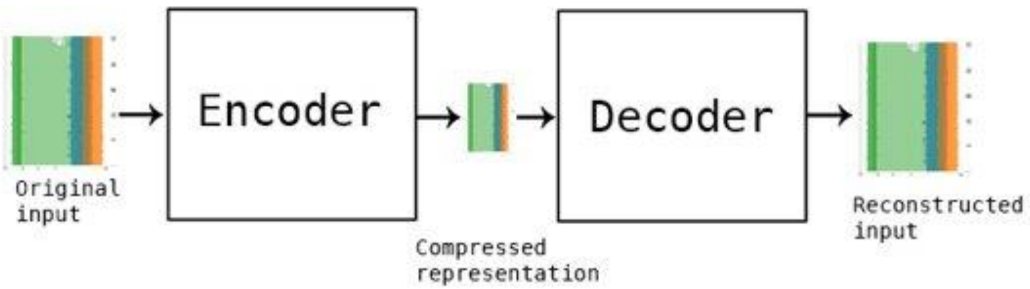
- Anomaly detection (or outlier detection) is the identification of rare items, events, or observations that raise suspicions by differing significantly from the majority of the data.
- In our case, anomaly anomalies are undesired conditions of the equipment.
- A priori, **these conditions are usually unknown**. Therefore, unsupervised approaches are necessary.



METHOD
AUTOENCODER

- An **autoencoder** is trained to compress and then reconstruct the signal on **healthy data**.
- The **reconstruction error** is the difference between the original and the reconstructed signal.
- The model will fail to reconstruct abnormal data and therefore yield a high reconstruction error
- The model is evaluated on the test set and compared to the **benchmark** results.

Approach	Reconstruction Error
	Dense Autoencoder
Model	LSTM Autoencoder
	CNN Autoencoder

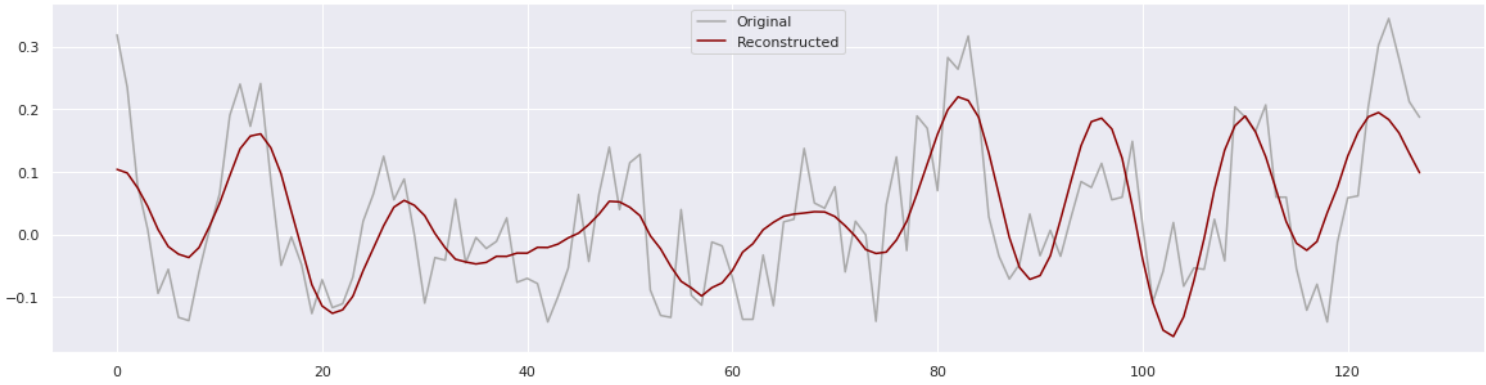


METHOD

ANOMALY SCORE

The anomaly score measures how well the signal was reconstructed. The higher the anomaly score, the likelier it is an anomaly.

The anomaly score is defined as the root mean squared error of the original and the reconstructed signal.



METHOD
FEATURES AND MODELS

Models

- Stacked LSTM autoencoder
- Stacked CNN autoencoder
- Fully connected autoencoder

Features

- FFT/Raw
- Vibration/audio
- 1D/3D

METHOD
BENCHMARKS

BM 1 – PCA

- Reconstruction Error with PCA. Signal is compressed and reconstructed form principal components.

BM 2 - IQR

- Anomalies are defined as data points that lie outside mean ± (iqr*1.5). Mean and iqr are calculated using the training set only.

BM 3 - Mean

- Deviation from mean signal is the anomaly score

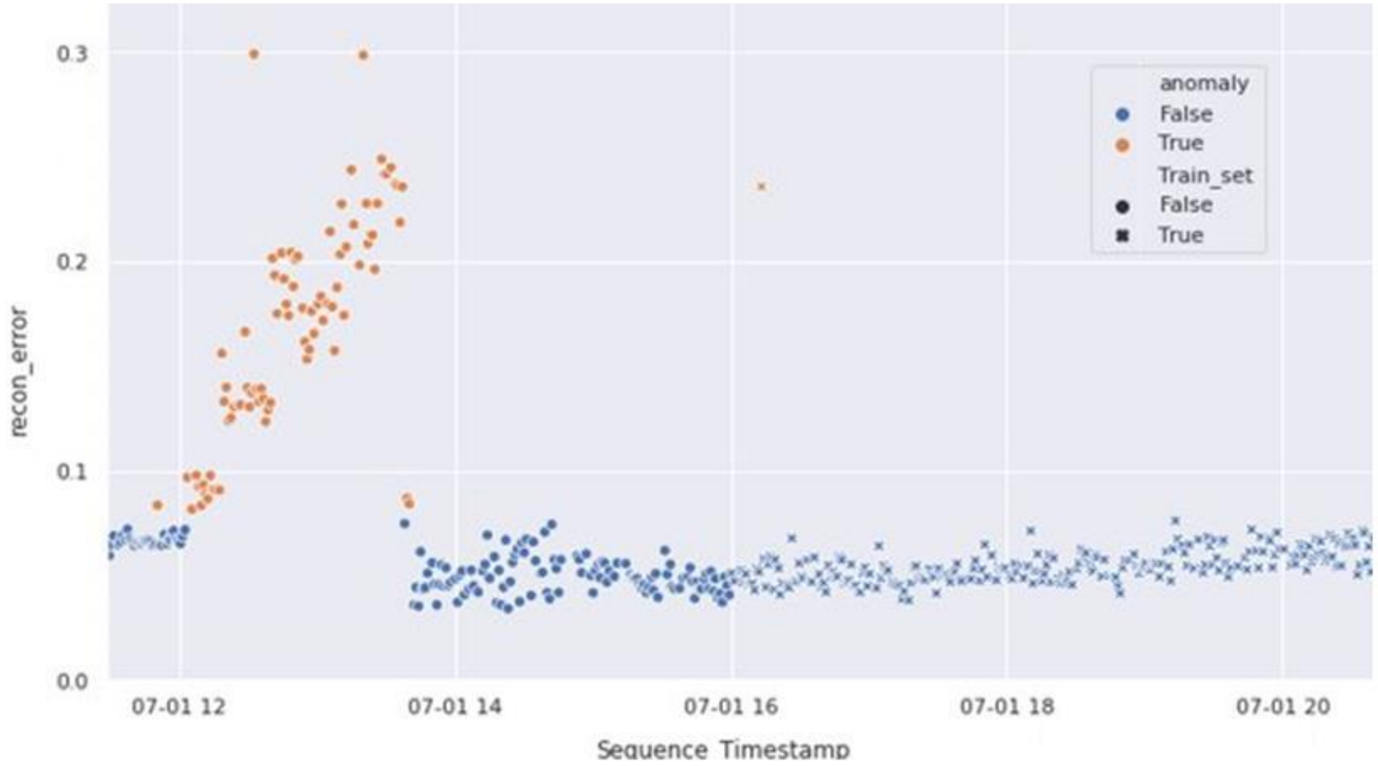
BM 4 – Range

- Data points outside the range of the train set are labels as anomalies

RESULTS

- The autoencoder fails to reconstruct the signal for anormal samples.

- Thus, anormal samples yield a higher anomaly score.



RESULTS

	Benchmarks																Models											
	BM ICR				BM Mean				BM PCA				BM Range				CNN				DNN				LSTM			
	Acc.	F1	P	R	Acc.	F1	P	R	Acc.	F1	P	R	Acc.	F1	P	R	Acc.	F1	P	R	Acc.	F1	P	R	Acc.	F1	P	R
Vibrations 1D	0.48	**0.63**	0.47	0.94	0.49	0.43	0.44	0.41	0.51	0.48	0.47	0.48	0.49	0.45	0.45	0.41	0.47	0.47	0.44	0.5	0.49	0.53	0.46	0.61	0.47	0.48	0.44	0.53
Audio	0.56	0.31	0.55	0.21	0.52	**0.54**	0.48	0.61	0.48	**0.54**	0.45	0.66	0.54	0.32	0.5	0.61	0.46	0.53	0.44	0.67	0.46	0.54	0.45	0.68	0.46	0.53	0.44	0.67
Vibrations 3D	0.52	**0.63**	0.49	0.89	0.51	0.52	0.47	0.57	0.5	0.45	0.46	0.44	0.55	0.25	0.56	0.57	0.54	**0.62**	0.5	0.8	0.54	**0.62**	0.5	0.81	0.55	**0.62**	0.51	0.79
Vibrations 1D & Audio	0.47	0.62	0.46	0.94	0.52	**0.54**	0.49	0.62	0.48	**0.54**	0.45	0.66	0.48	**0.48**	0.44	0.62	0.46	0.53	0.44	0.67	0.47	0.54	0.45	0.68	0.46	0.53	0.44	0.67
FFT Vibrations 1D	0.48	**0.63**	0.47	0.94	0.49	0.43	0.44	0.41	0.51	0.48	0.47	0.48	0.49	0.45	0.45	0.41	0.54	0	0	0	0.59	0.47	0.59	0.39	0.53	0.02	0.4	0.01
FFT Audio	0.56	0.31	0.55	0.21	0.52	**0.54**	0.48	0.61	0.48	**0.54**	0.45	0.66	0.54	0.32	0.5	0.61	0.53	0.46	0.49	0.44	0.5	0.45	0.46	0.44	0.52	0.48	0.48	0.48
FFT Vibrations 3D	0.52	**0.63**	0.49	0.89	0.51	0.52	0.47	0.57	0.5	0.45	0.46	0.44	0.55	0.25	0.56	0.57	0.54	0.04	0.57	0.02	0.49	0.43	0.44	0.41	0.6	0.37	0.71	0.25
FFT Vibrations 1D & Audio	0.47	0.62	0.46	0.94	0.52	**0.54**	0.49	0.62	0.48	**0.54**	0.45	0.66	0.48	**0.48**	0.44	0.62	0.56	0.08	1	0.04	0.5	0.45	0.45	0.44	0.63	0.42	0.74	0.3

RESULTS

1 Best model

Score 0.62 F1

LSTM Autoencoder

Euclidian norm of 3D vibration signal

2 BM1 PCA

Best score 0.54

Models beat BM 16/24 experiments

3 BM2 ICR

Best score of 0.63

Outperforms our models!

SUMMARY

We experimented with	Results	Outlook
• Three different neural network architectures • Unsupervised machine learning models • Without any feature engineering	• Our model outperforms 3 benchmarks • The benchmark based on the interquartile range outperforms our models	• Focus on supervised approaches and transfer learning

AUSTRIAN INSTITUTE OF TECHNOLOGY

TOMORROW TODAY

clemens.heistracher@ait.ac.at