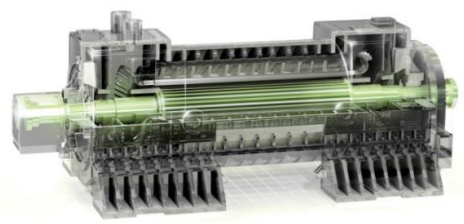
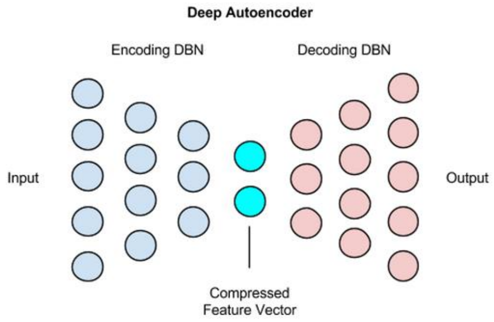
1st European COST Conference on Mathematics for Industry in Switzerland

Deep Fault Detection

-- An Unsupervised Representation to Detect Anomalies in Raw Condition Monitoring Signals





Yang Hu[1], Thomas Palmé[2], and Olga Fink[1*]

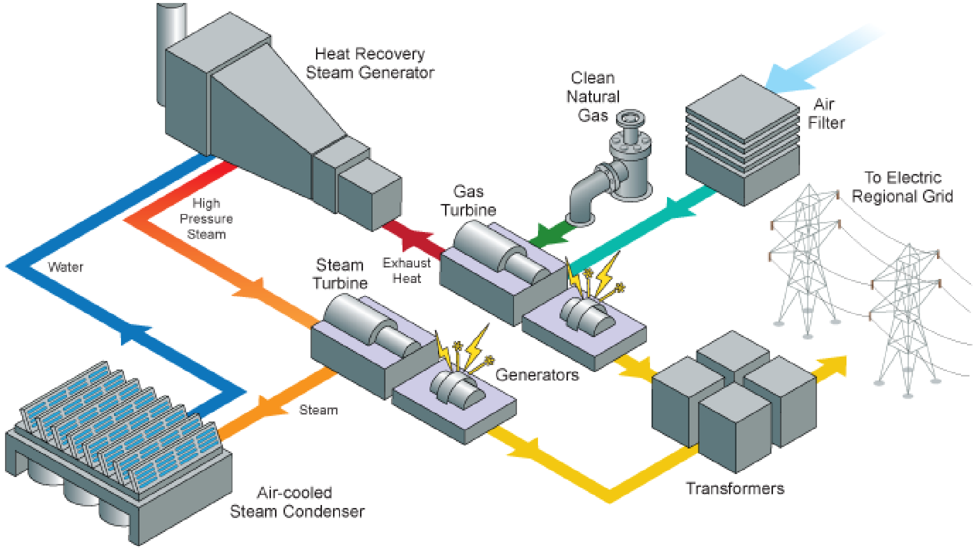
[1]*Zurich University of Applied Sciences, Rosenstr. 3, Winterthur, 8401, Switzerland,*

[2]*General Electric (GE) Switzerland, Brown Boveri Str. 7, Baden, 5401, Switzerland*

2016.09.12

Relevance of the problem

Early warnings and fault detections
of the key components in industry systems



Generator

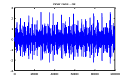
Bearing

- Key component of power plants
- Penalties in case of power supply shortage
- Increase efficiency, reduce operation costs (without sacrificing reliability)

Responsible of about **40%** of the failures in industrial motors.

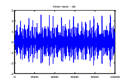
Data-Driven approach

Sensor system



Sensor 1

. . .



Sensor n

Feature Extraction & Selection

High level informative feature

Detection Model
- Normal
- Anomaly
- Fault

Diagnostic Model
- Malfunction Type1
- …
- Type k

Prognostic Model
- Remaining Useful Life

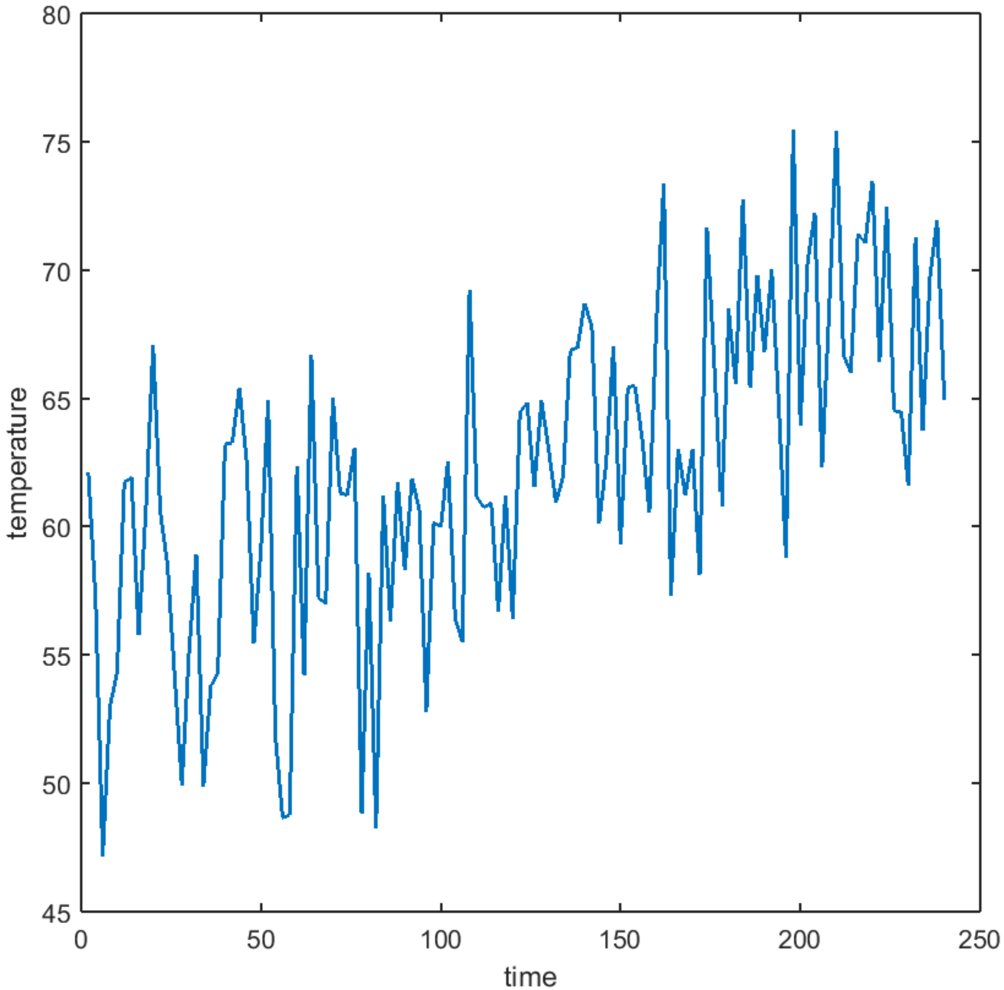
Step 1
Signal acquisition

Step 2
Feature Engineering

Step 3
Model Development

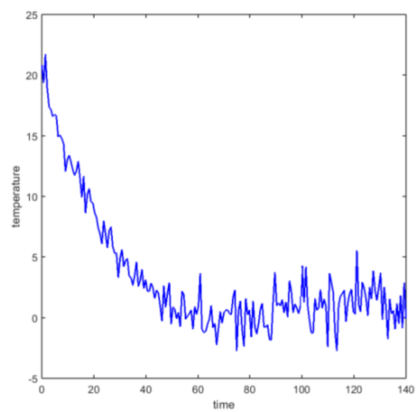
Feature engineering: Issues

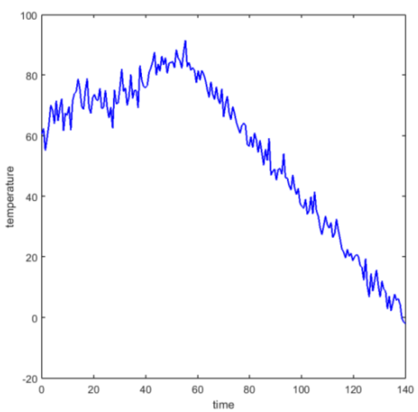
- Raw signal measurements affected by noise

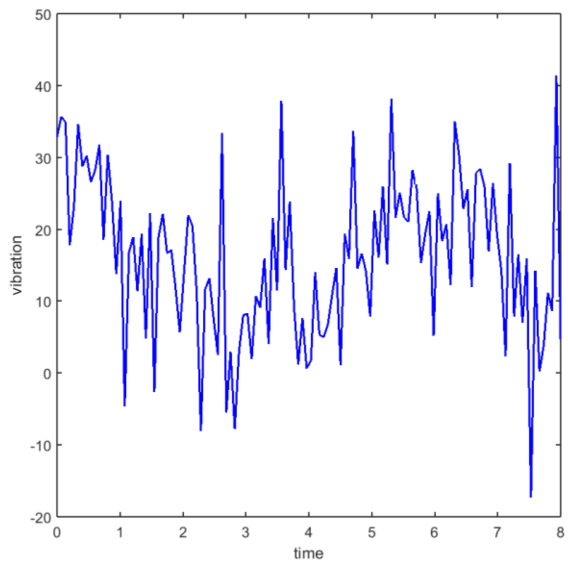


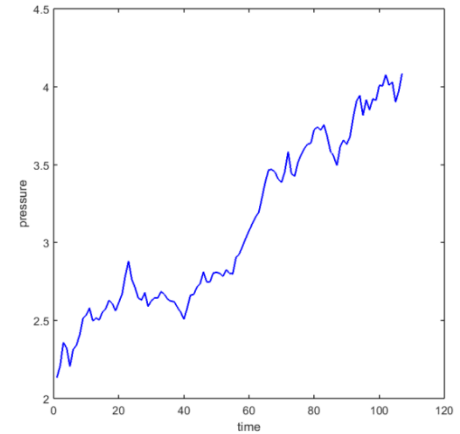
Feature engineering: Issues

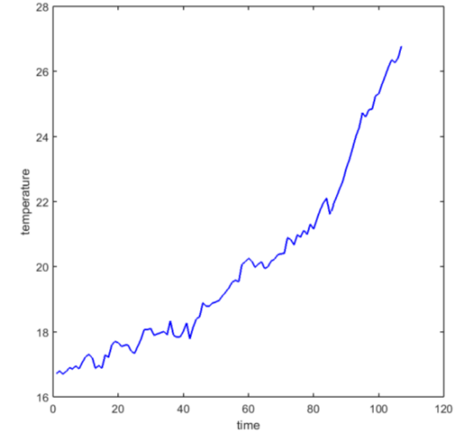
- Measurements affected by noise
- Available measurements not directly related to degradation / fault progression
- Measurements partially non-informative, or redundant



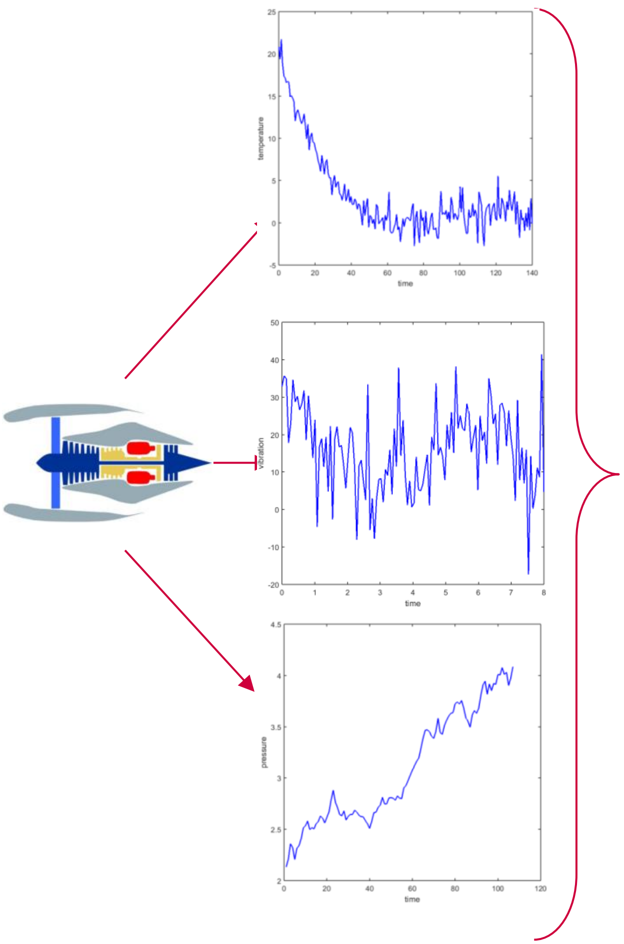




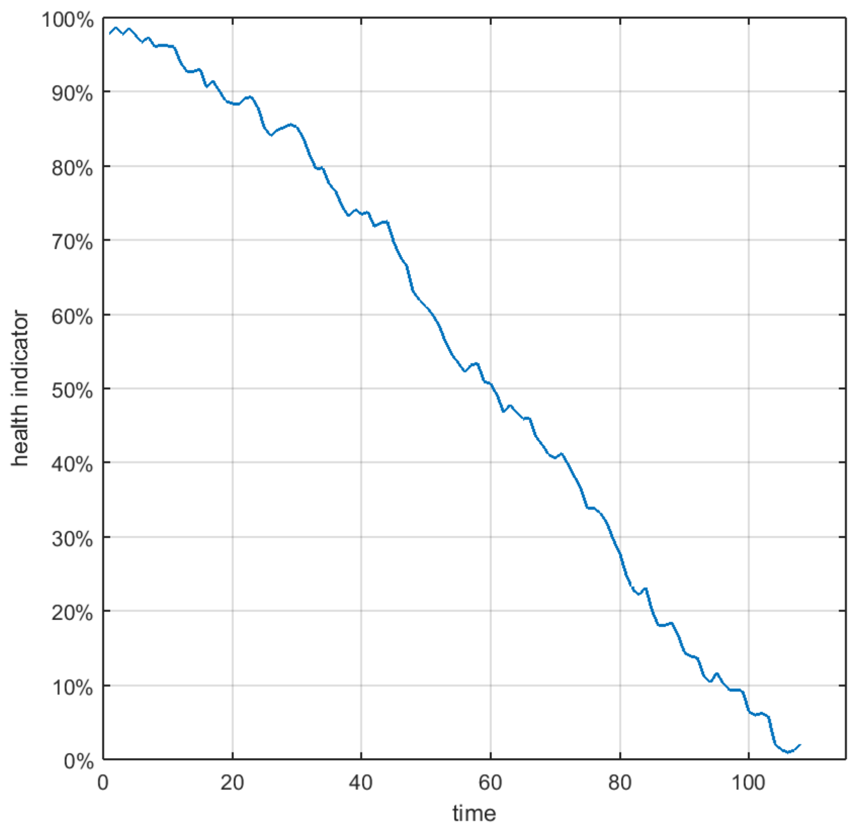


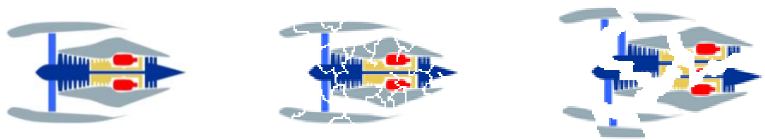


Feature engineering: Target



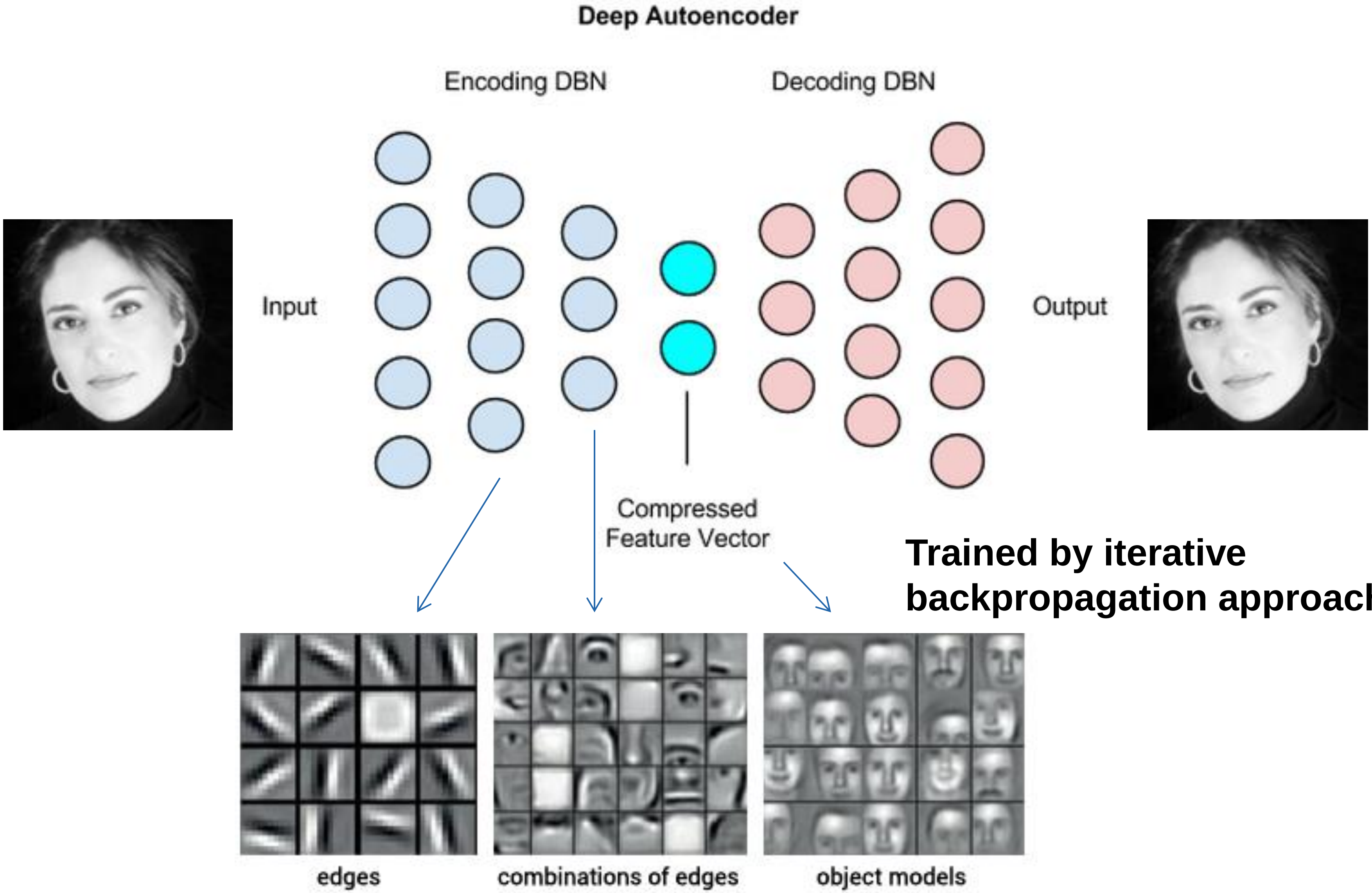
Feature extraction: Health Indicator (HI)





- **Specific for the system or equipment**
- **Intensive time and human resource consumption**

From feature engineering to feature learning



Trained by iterative backpropagation approach

Hierarchical Extreme Learning Machines (HELM)

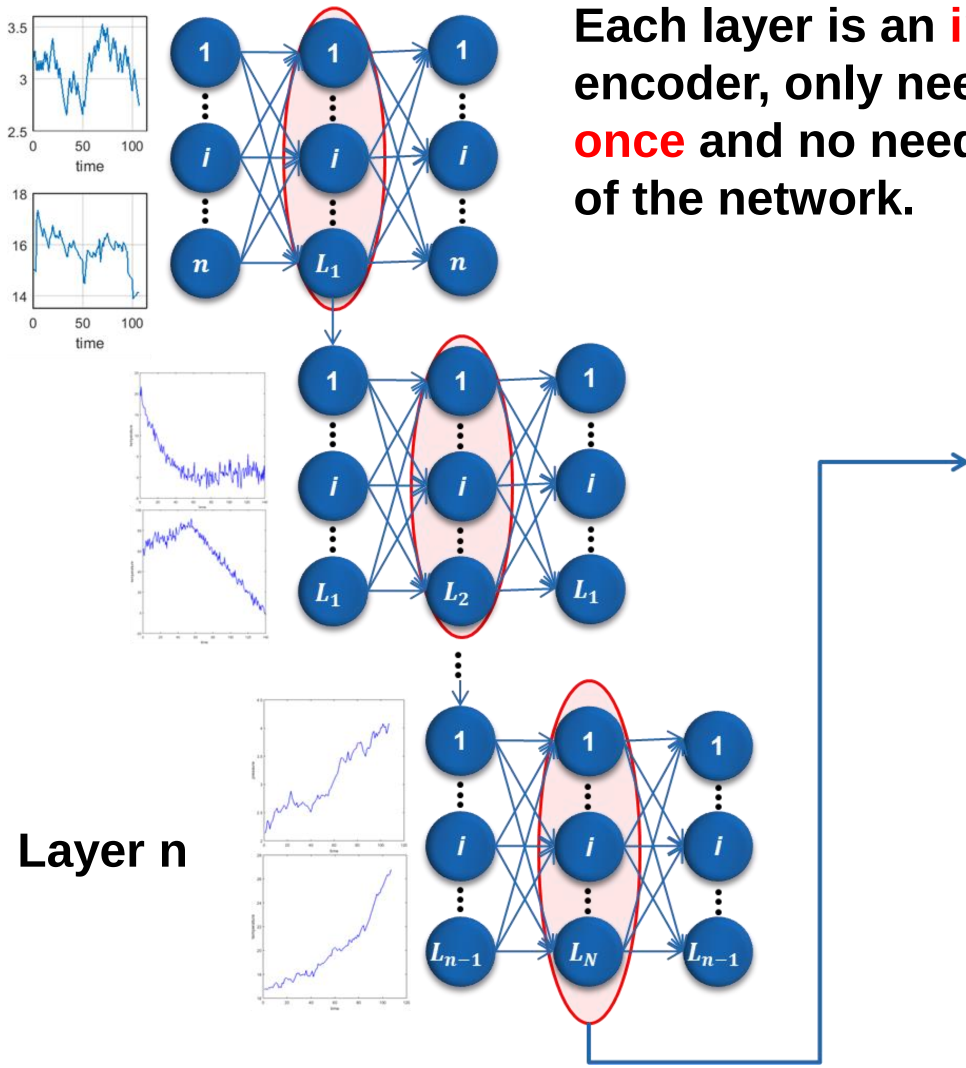
Layer 1

Raw signals

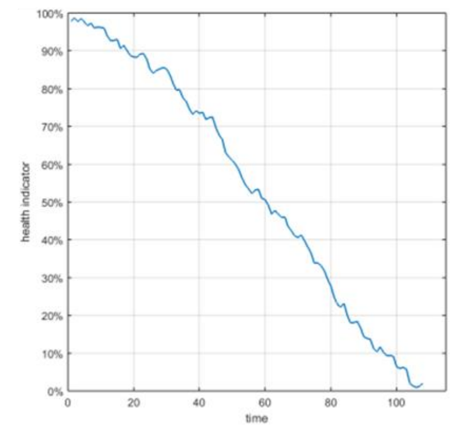
Layer 2

•
•
•

Layer n



Each layer is an independent auto-encoder, only need to be trained for once and no need for iterative tuning of the network.

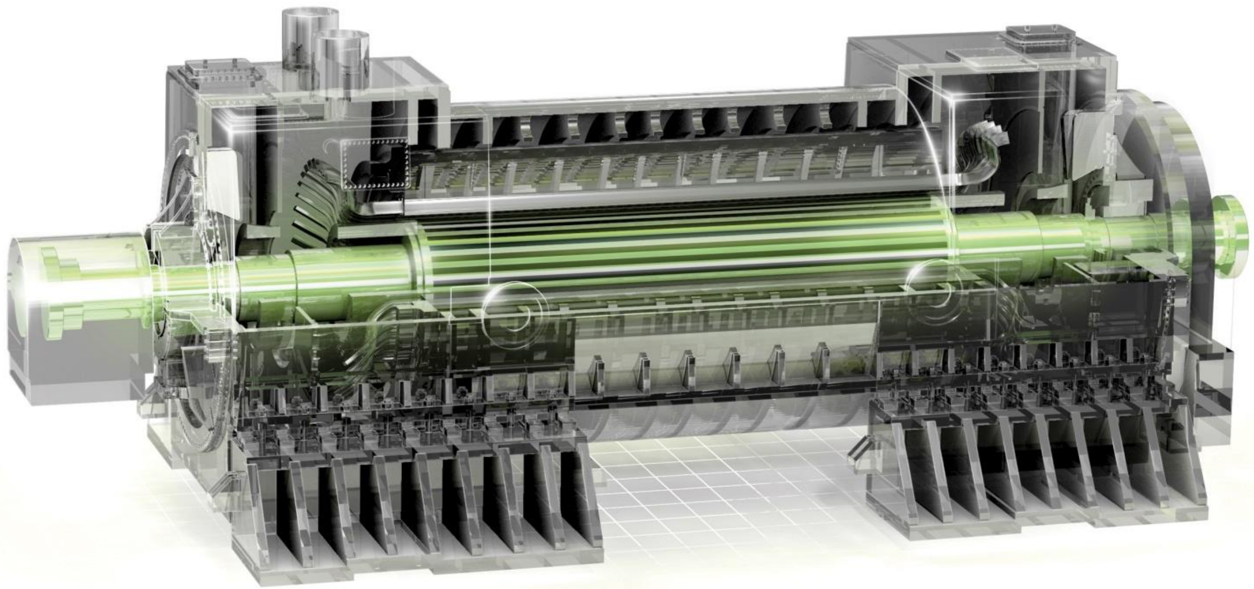


Learned feature

Reference Case Study 1:

Fault detection of generator in combined cycle power plant

Generator condition monitoring



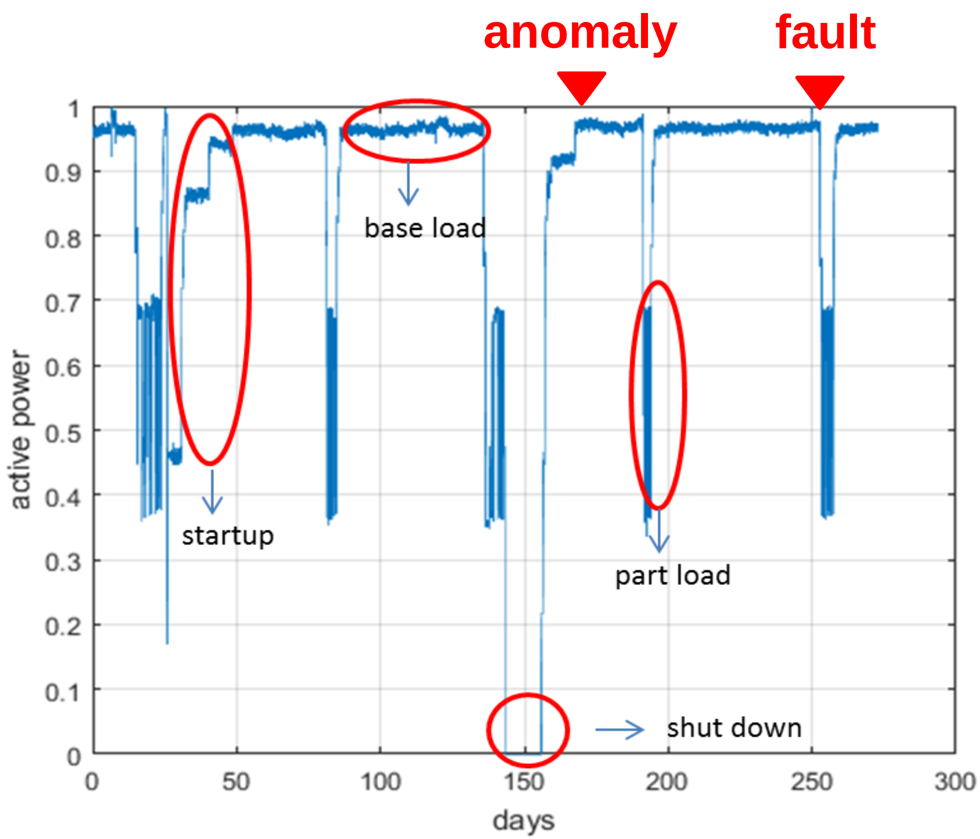
Large quantity of health related signals

Sensing and Data Transmission Devices

188 signals

Sensor	Tasks	Number
partial discharge monitoring **PD**	detect aging of the main insulation, loose bars or contact as well as contamination	16
rotor shaft voltage **RSV**	detect shaft grounding problems, shaft rubbing, electro erosion, bearing isolation problems and rotor inter-turn shorts	7
stator end winding vibration **WV**	detect deterioration in mechanical stiffness of overhang support system	133
rotor flux **RF**	detect the occurrence, the magnitude and the location of the rotor winding inter-turn short circuit	32

Application



Based on the operation records of the generator, an anomaly, and a fault were observed on day **169** and **247** respectively

Only the "base load" observations (totally 55'774 observations) are considered in this case study
→ the overall dimension of signal set is **55'774x188**

Evaluation of feature learning approach: a comparison

Perform the anomaly/fault detection of generator using different signals and feature extraction approaches

partial discharge monitoring PD (16)
rotor shaft voltage RSV(7)
stator end winding vibration WV(133)
rotor flux RF(32)
all signals (188)

Different signals as input

Raw signal
Manual
Principle Component Analysis
HELM

Different feature extraction approaches

Zürcher Hochschule
für Angewandte Wissenschaften

Experiment design

Without signal processing

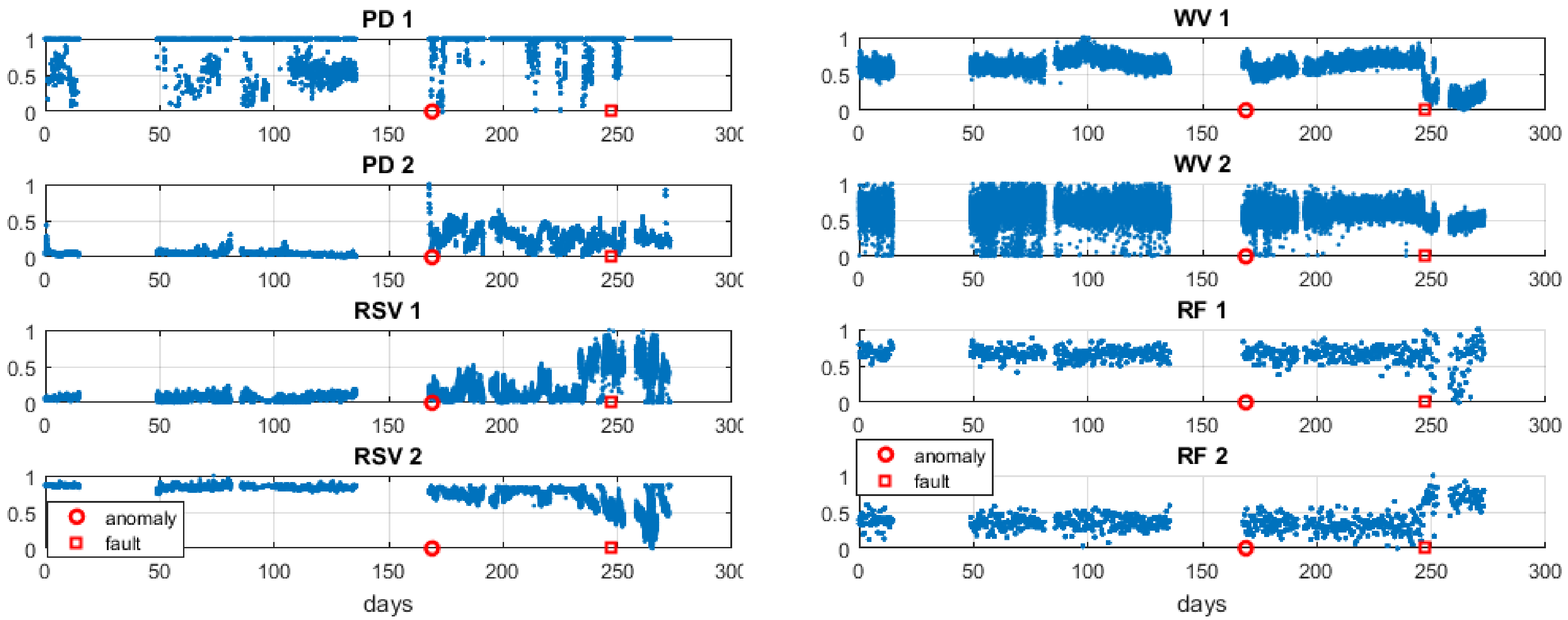
Expert knowledge

Traditional feature extraction approach

Our approach

	Raw signal	Manual	Principle Component Analysis		HELM	
	Input dimension	Selected input dimension	Number of input PCs	Variance explained	Number of layers	Hidden neurons in each layer
PD	16 signals	Select 3 from 16 signals	3	88.6%	2	10,50
RSV	7 signals	Select 2 from 7 signals	3	82.2%	2	3,20
WV	133 signals	Select 5 from 133 signals	30	85.3%	3	50,10,50
RF	32 signals	Select 4 from 32 signals	15	80.8%	2	10,50
All signals	188 signals	Select 8 from 188 signals	30	81.1%	3	50,10,50

Manually selected signals

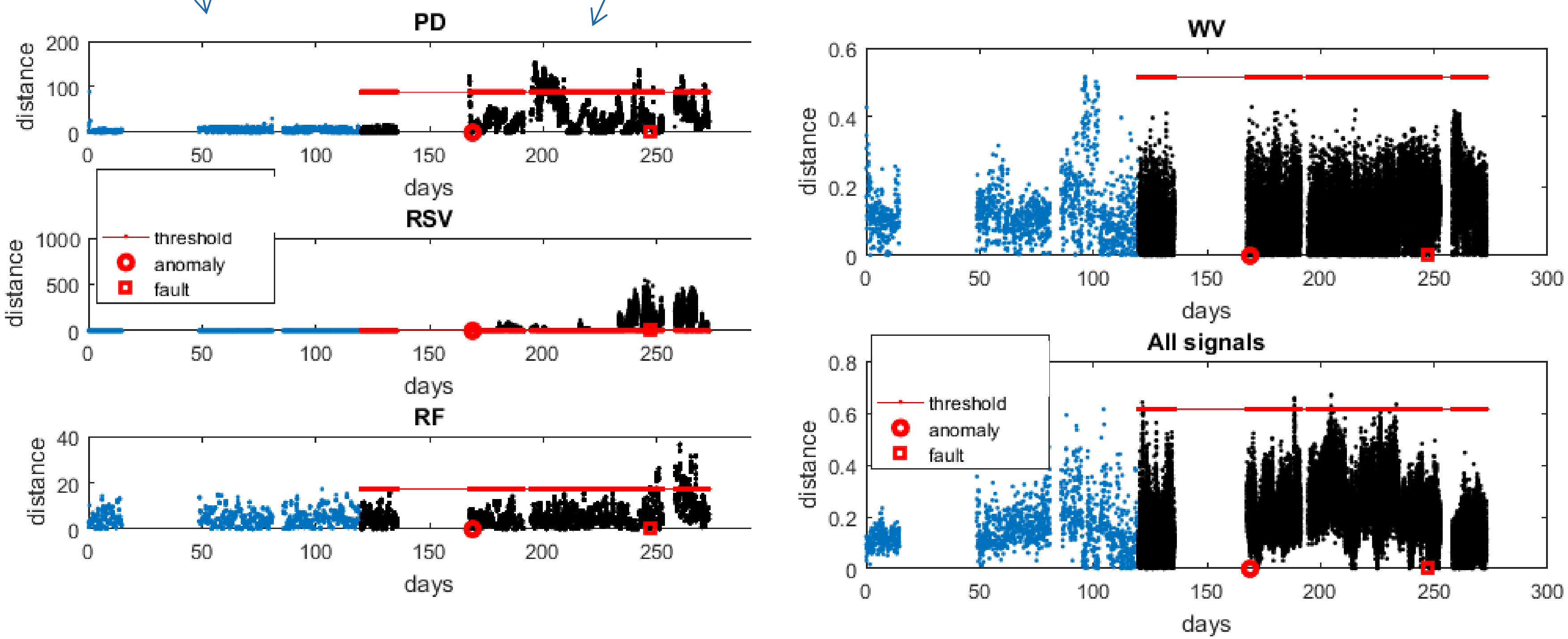


Very time consuming process: 8 “good” signals from 188 raw signals by experienced engineers

Detection results:
Principle Component Analysis

Reference set
Day 1 to 120

Test set
Day 121 to 271

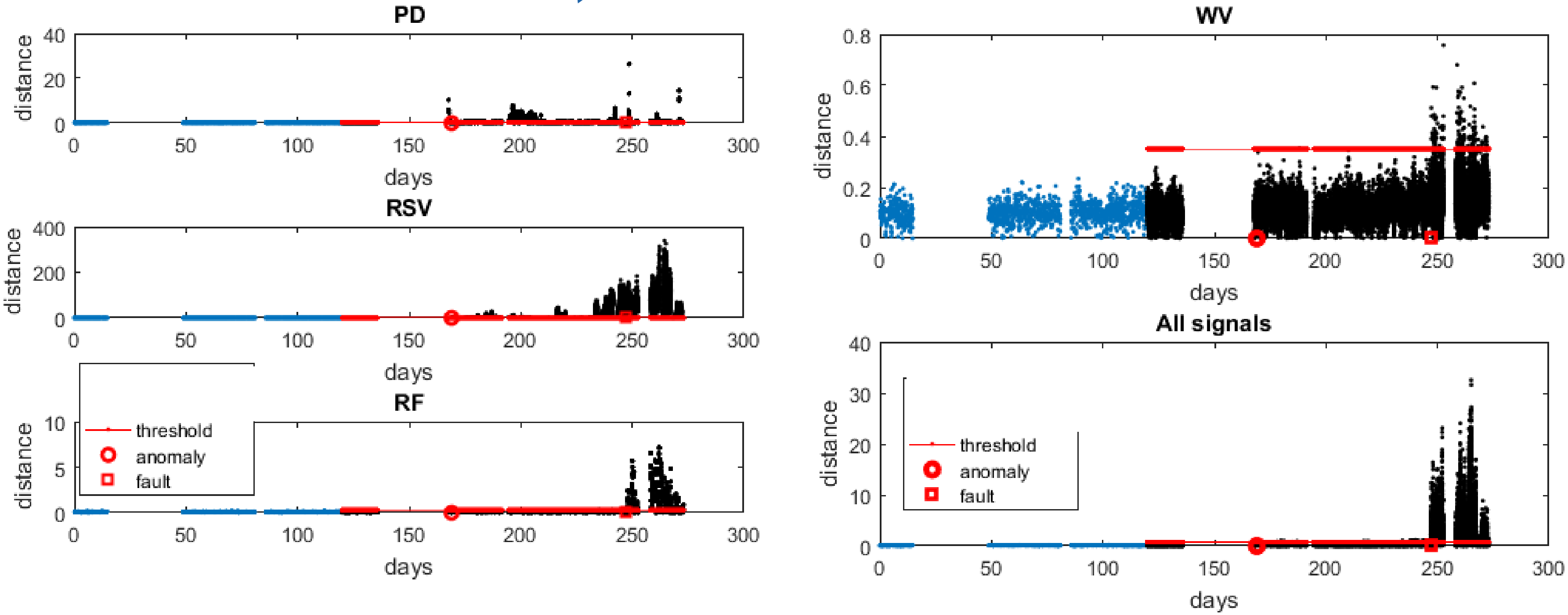


Detection results: HELM

Zürcher Hochschule
für Angewandte Wissenschaften

Reference set
Day 1 to 120

Test set
Day 121 to 271



Detection results: overall

	Raw signal		Principle Component Analysis		Manual		HELM	
	Anomaly	Fault	Anomaly	Fault	Anomaly	Fault	Anomaly	Fault
PD	Yes	No	Yes	No	Yes	No	Yes	No
RSV	Yes	No	Yes	No	Yes	No	Yes	No
WV	No	**No**	No	**No**	No	Yes	No	Yes
RF	No	**No**	No	Yes	No	Yes	No	Yes
All signals	**No**	**No**	**No**	**No**	Yes	Yes	Yes	Yes

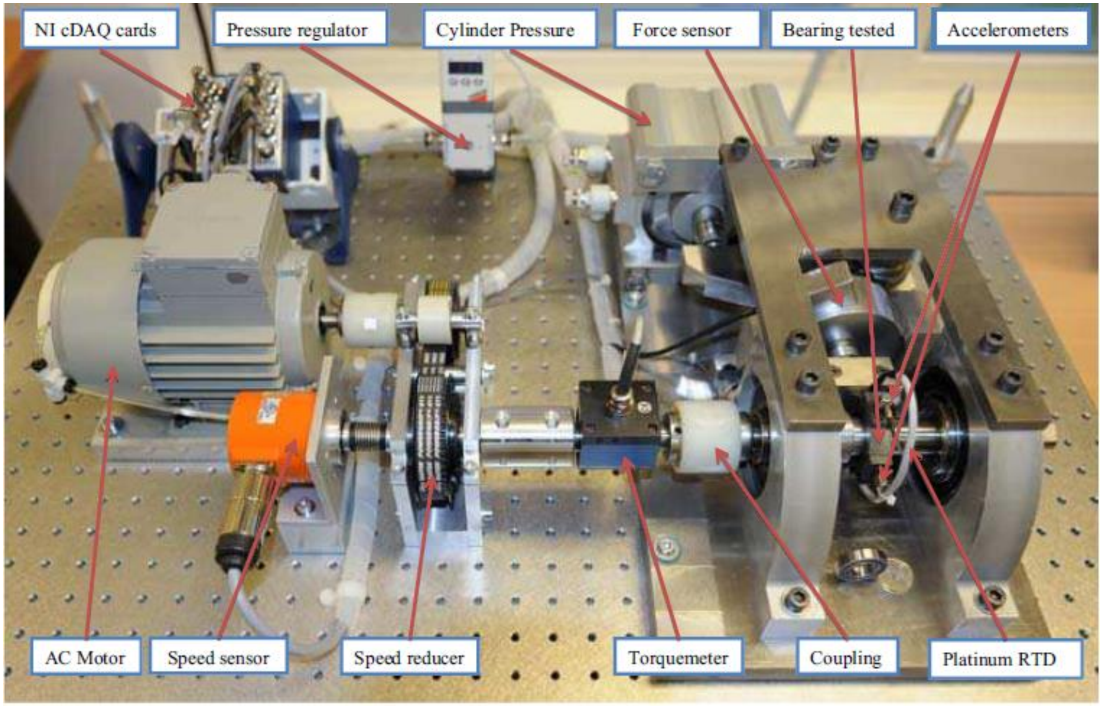
Performance

Raw signal < Principle Component Analysis < Manual = HELM

Reference Case Study 2:

Health indicator extraction of bearings under accelerated degradation tests

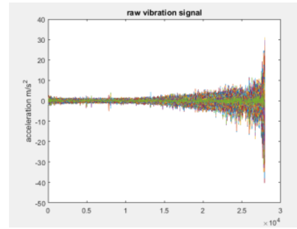
Accelerated degradation test



Experimental Platform of “PRONOSTIA” IEEE PHM 2012 Data Challenge

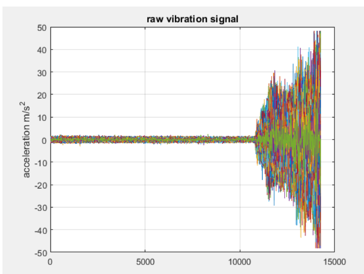
Signals from **three bearings** are considered:
Working condition: 1800 rpm, 4000 N.
Life: 27990, 23740 and 14270 seconds.
vibration signals frequency: 2560 samples per 10 seconds (with some time gaps).
dimensions of the raw signals: $1629 \times 2560, 2371 \times 2560, 1425 \times 2560$

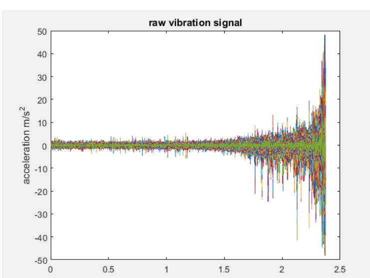
Experiment design



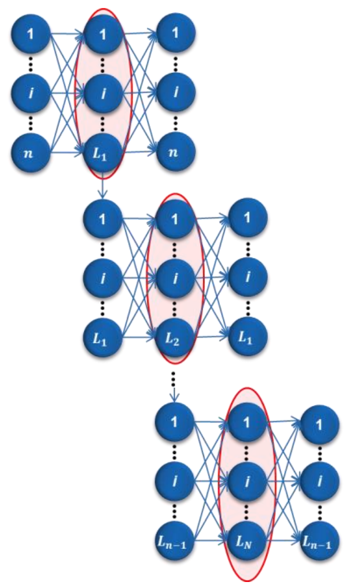
Bearing 1

Training





Testing

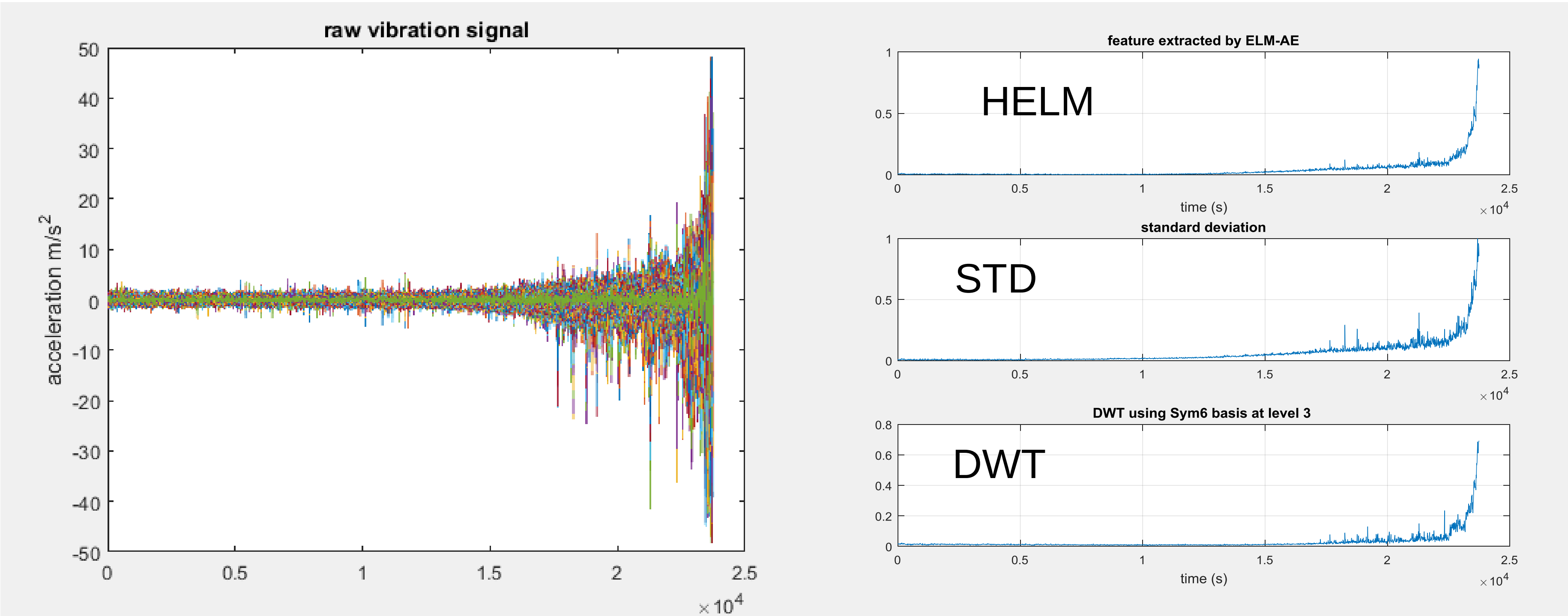


Health indicator

HELM model

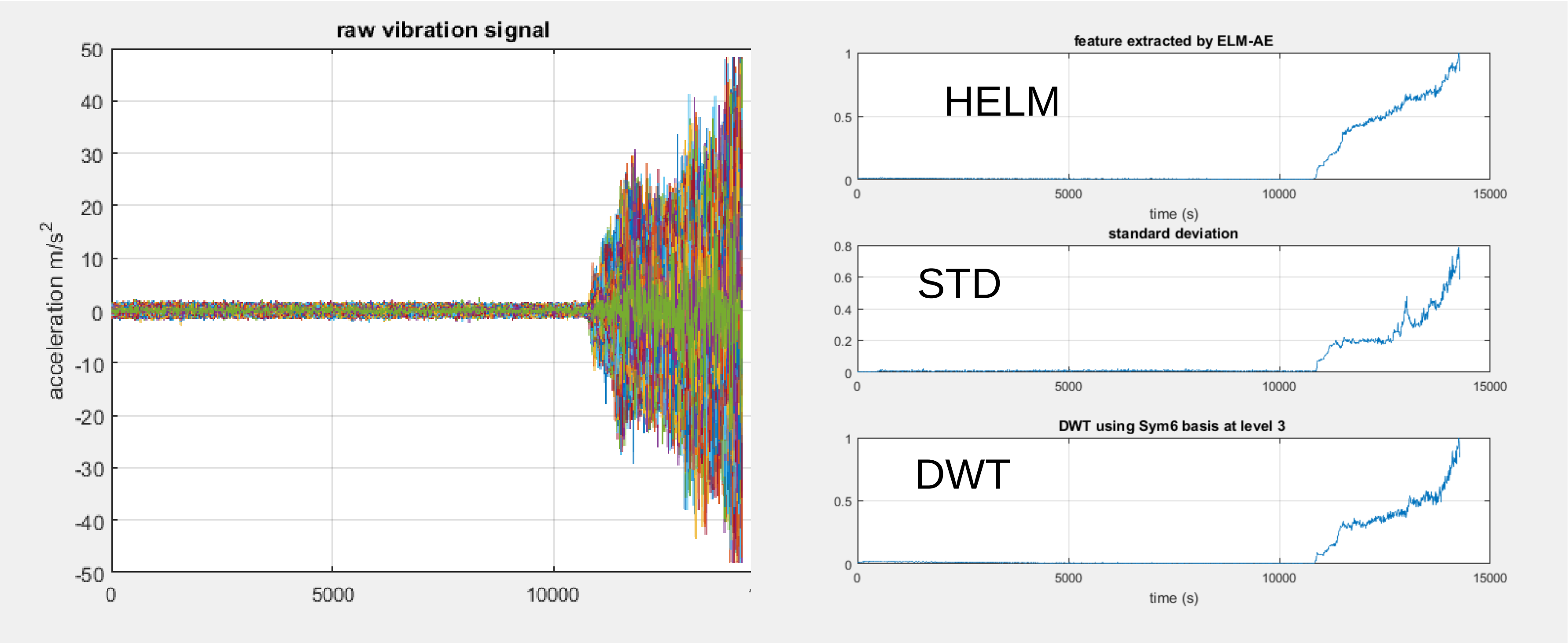
Bearing 2 and 3

Learned Features (bearing 2)



STD: standard deviation of vibration signal
DWT: Discrete wavelet transform using Sym6 basis at level 3

Learned Features (bearing 3)



STD: standard deviation of vibration signal
DWT: Discrete wavelet transform using Sym6 basis at level 3

Conclusion

Pros:

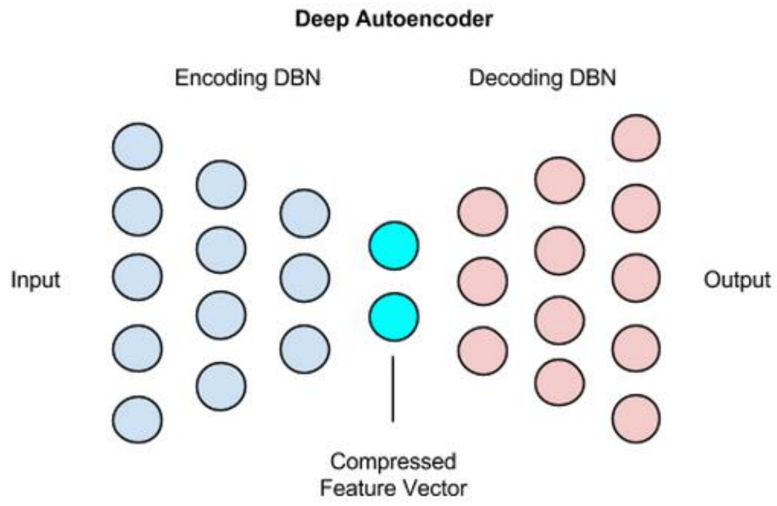
1. Capable to handle high-dimensional complicated raw signals;
2. No need for expert knowledge for feature engineering;
3. Strong adaptability to any learning problems

Cons:

1. Large training datasets required
→ however, just require data representing **“normal”** system state
2. “Black box”-approach → difficult to explain the results from the physical point of view
→ additional approaches required to identify signals with the biggest influence on the anomaly for physical interpretation, a **“pre-selection”** for the engineers.

Deep Fault Detection

-- An Unsupervised Representation to Detect Anomalies in Raw Condition Monitoring Signals



THANK YOU!

Yang Hu[1], Thomas Palmé[2], and Olga Fink[1*]