

Active Learning and Transfer Learning for Process Analytics

Chen Song, Ruomu Tan, Marco Gärtler, Martin Hollender, Sylvia Maczey (ABB) & Franz Baehner (Bayer) & Bram Bamps (Covestro)

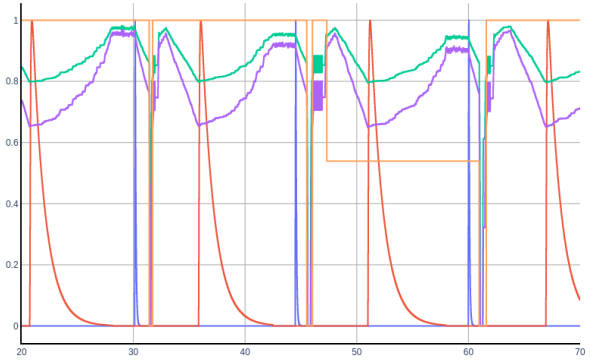
25.05.2023

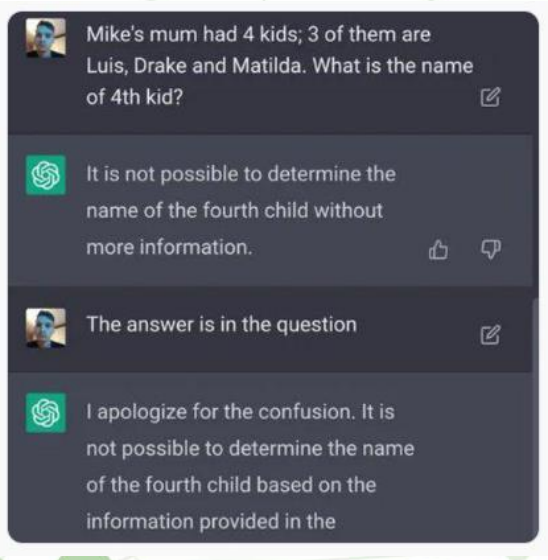
Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

KE-3N



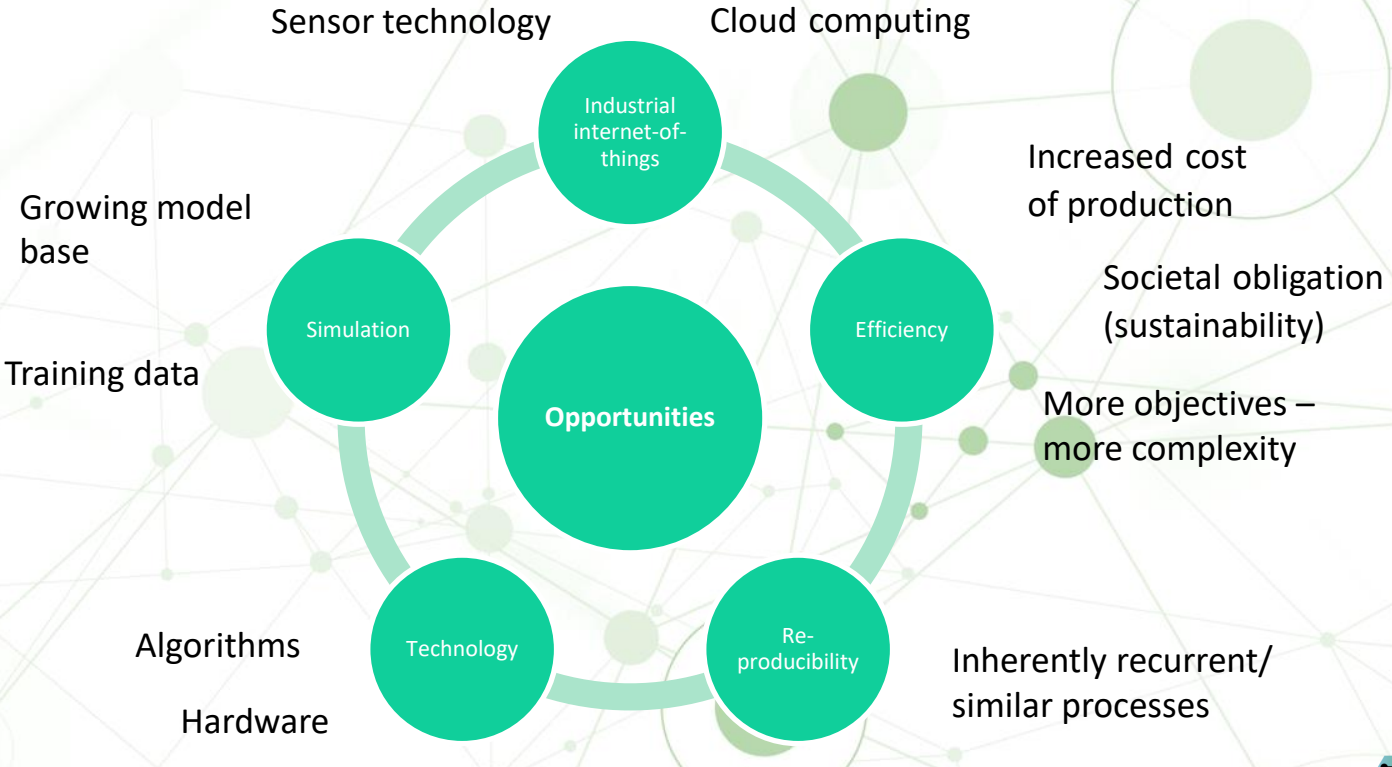


Process data analytics with ML

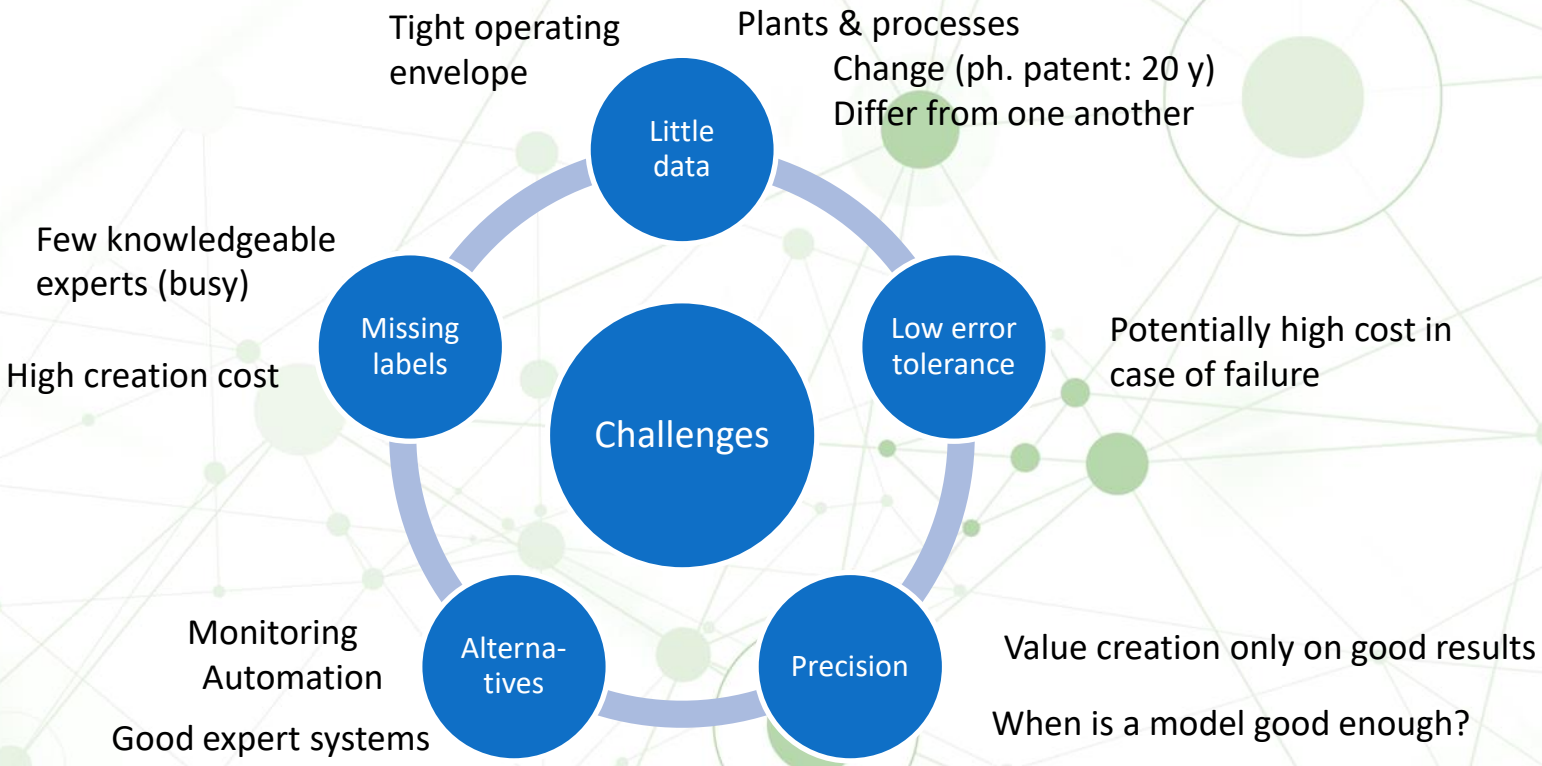
"A fool is very dangerous when in power."

Denis Fonvizin

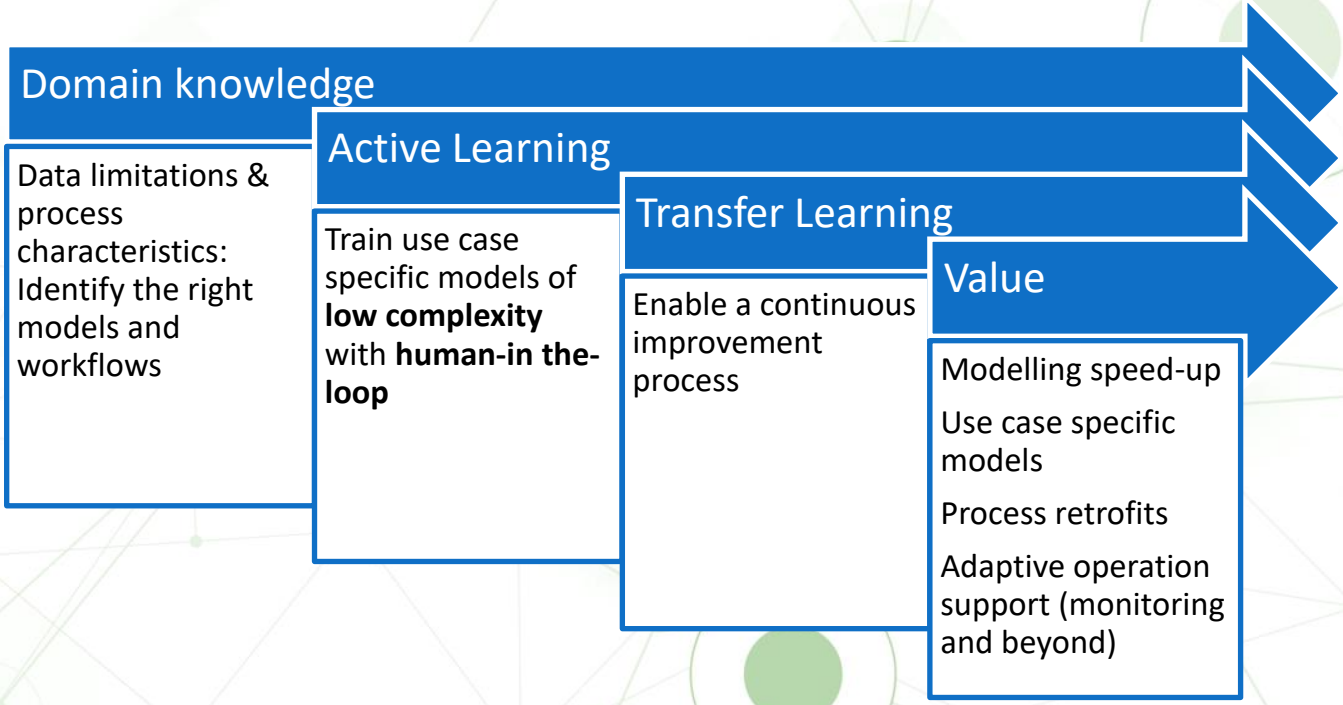
ML for process data analytics

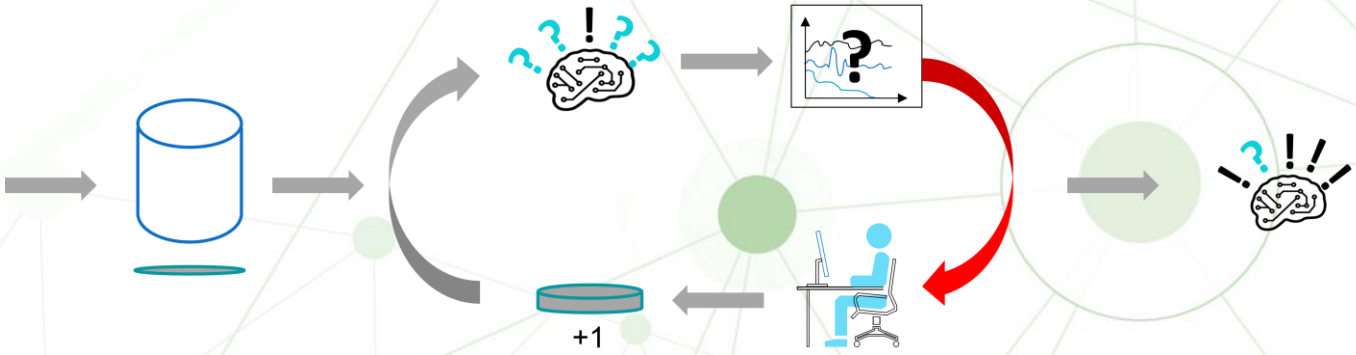


ML for process data analytics



Motivation

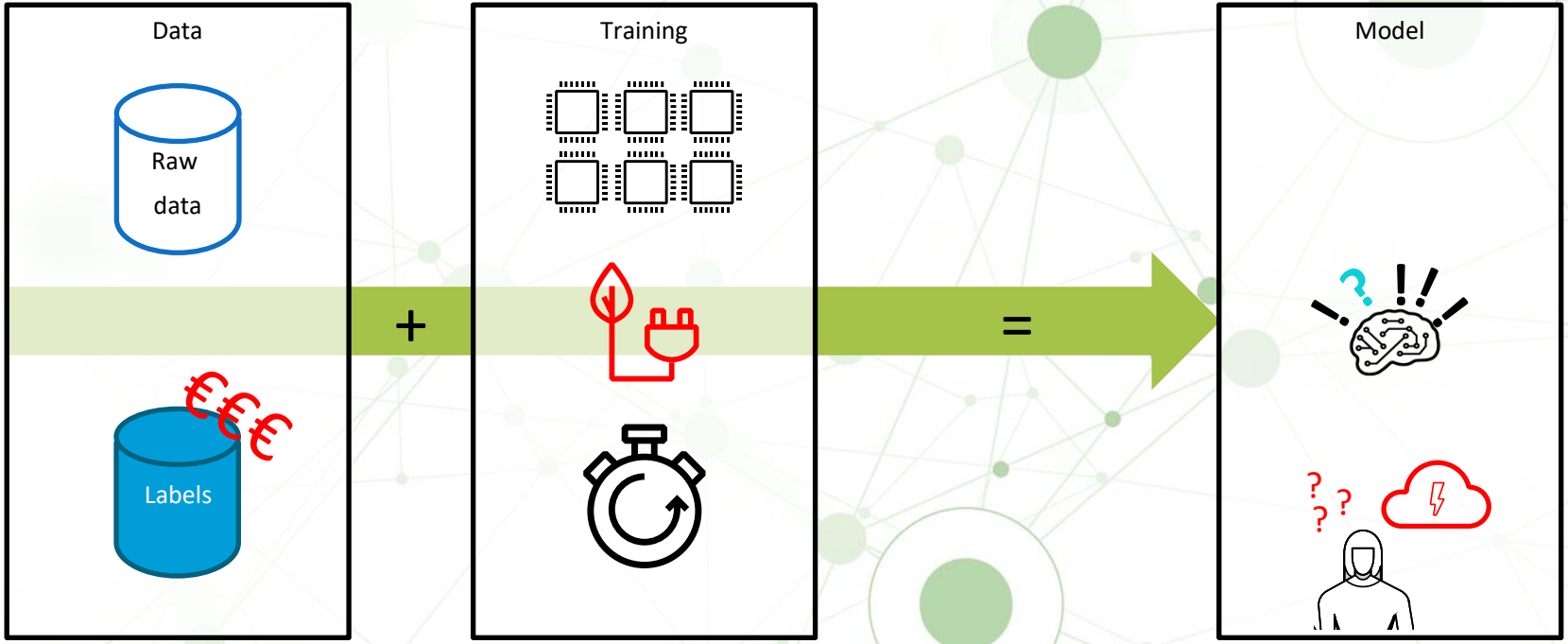

Domain knowledge
Data limitations & process characteristics: Identify the right models and workflows
Active Learning
Train use case specific models of **low complexity** with **human-in the-loop**
Transfer Learning
Enable a continuous improvement process
Value
Modelling speed-up
Use case specific models
Process retrofits
Adaptive operation support (monitoring and beyond)



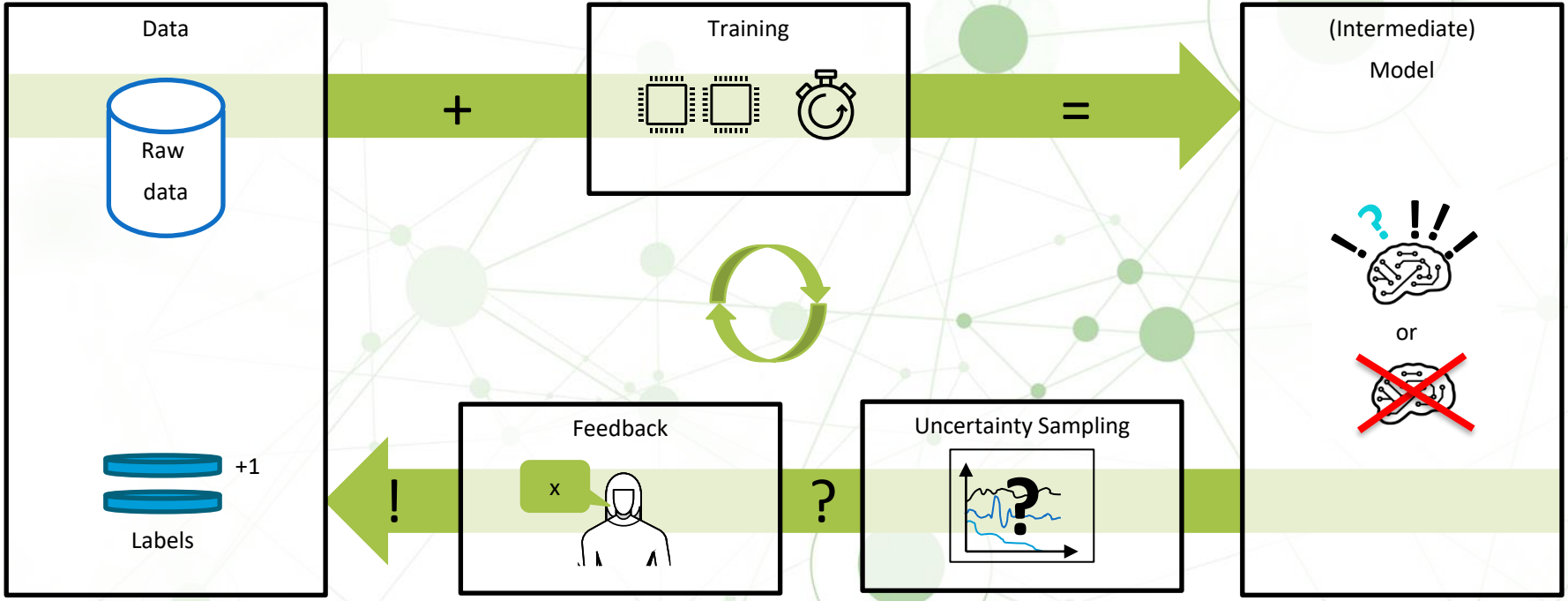
Technical background

Brief overview of methods

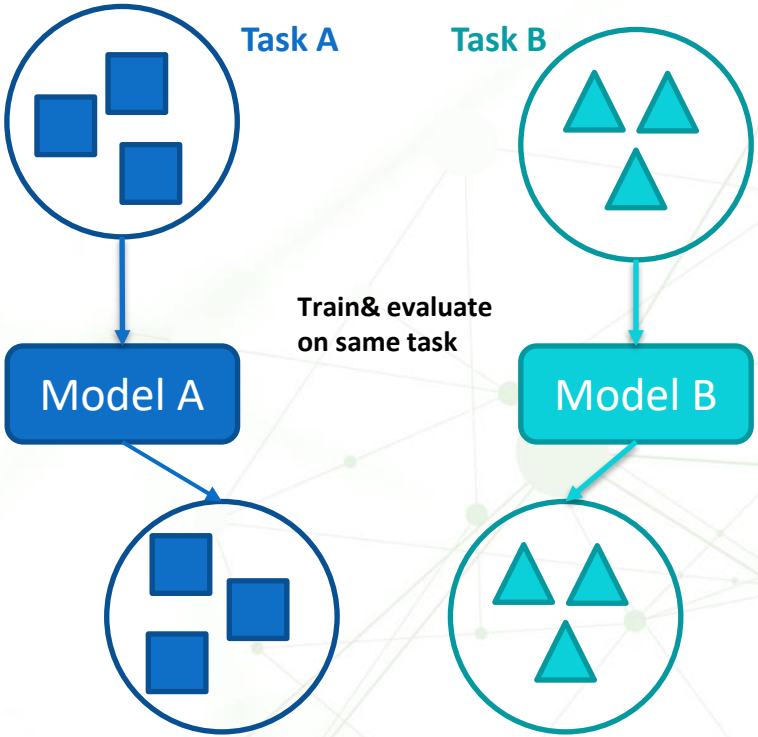
Traditional ML



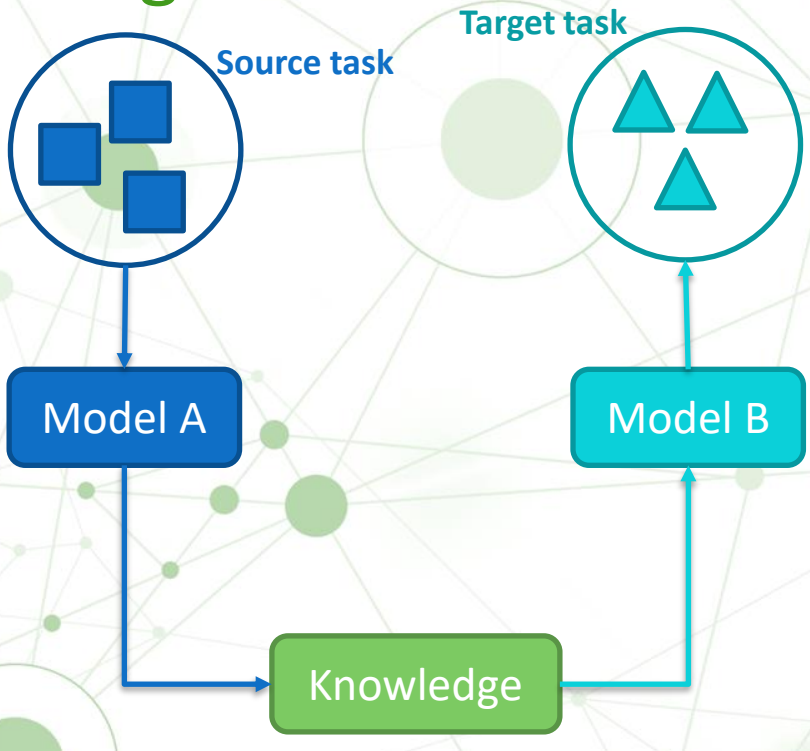
Active Learning



Transfer Learning

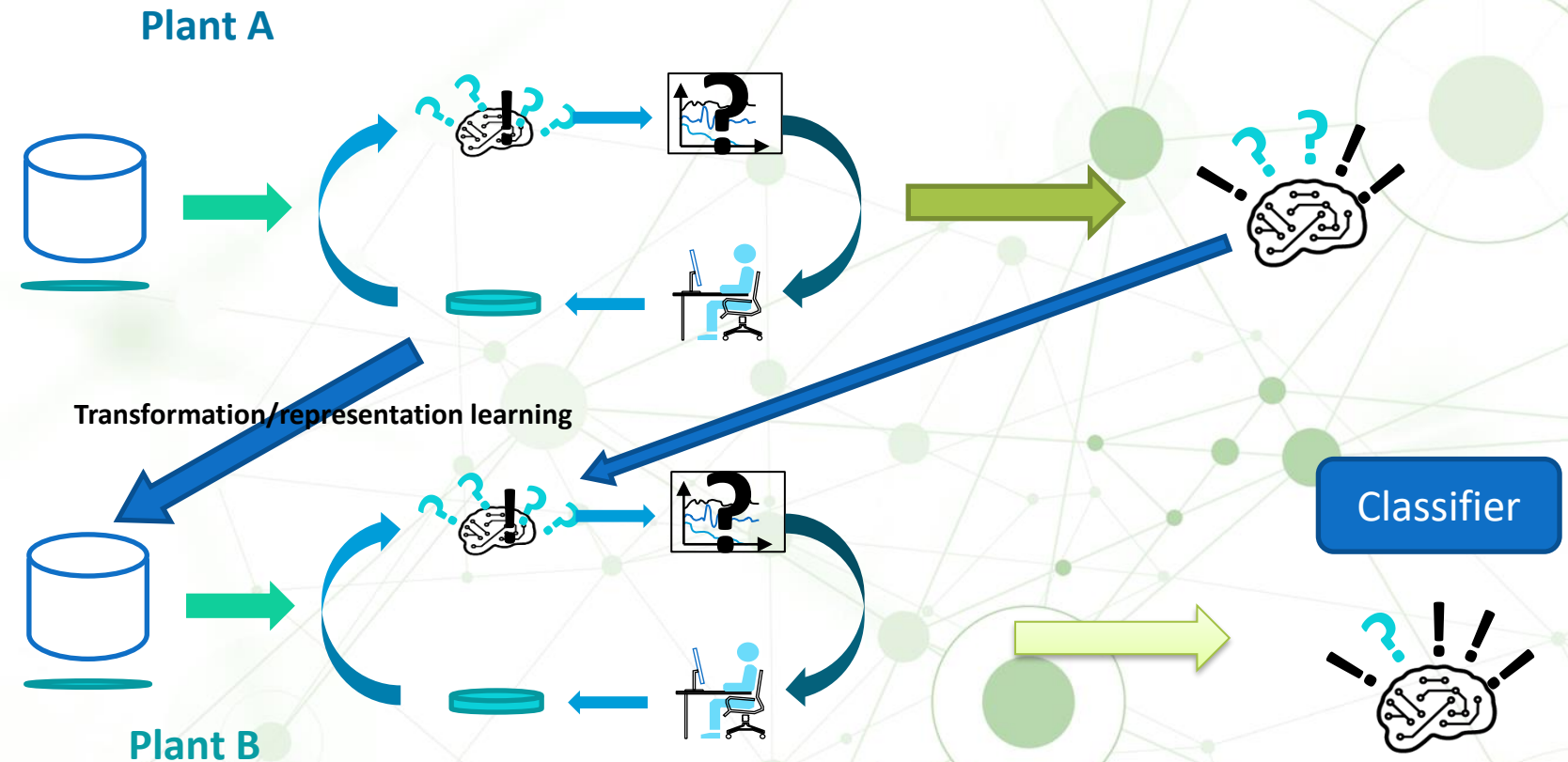


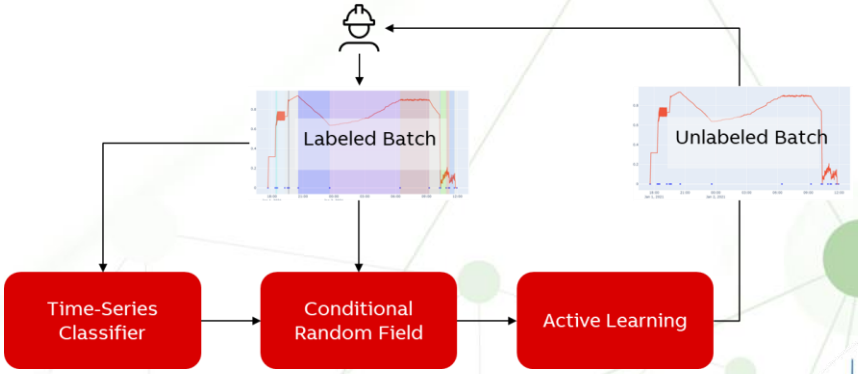
Traditional ML



Transfer Learning

Active Learning & Transfer Learning

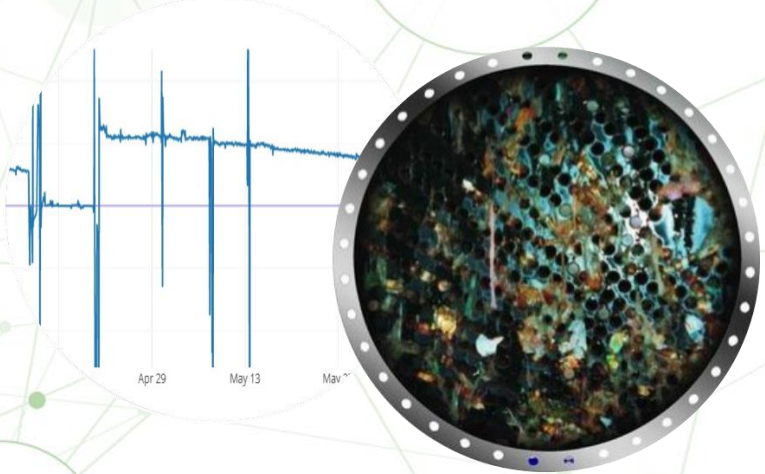




Use cases

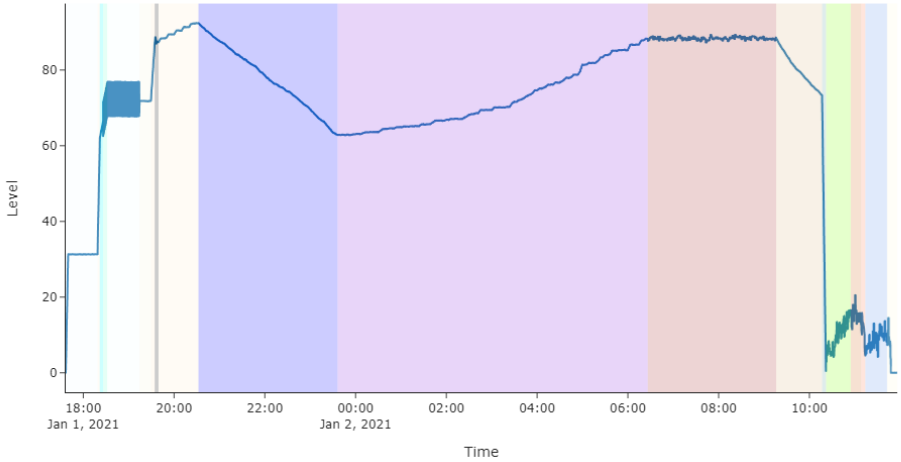
Batch labelling (Bayer)

Heat exchanger fouling cycle (Covestro)



Use case: batch labelling

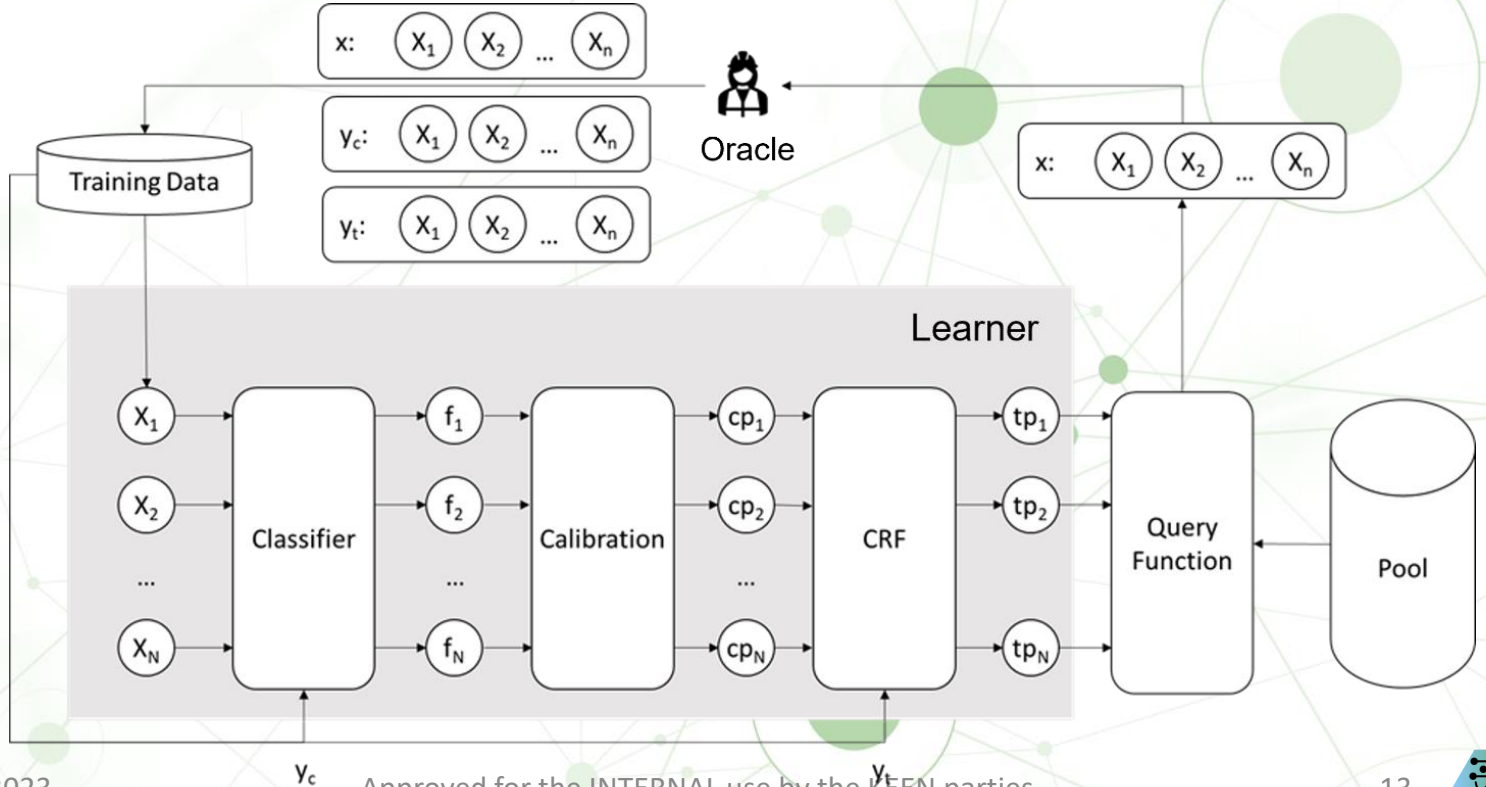
- Missing data labels in old and semi-automatic plants
- Method: sequence-sensitive timeseries classification
- Data
  - KEEN open dataset with multiple benchmark cycles[1]
  - Historical process data (confidential)



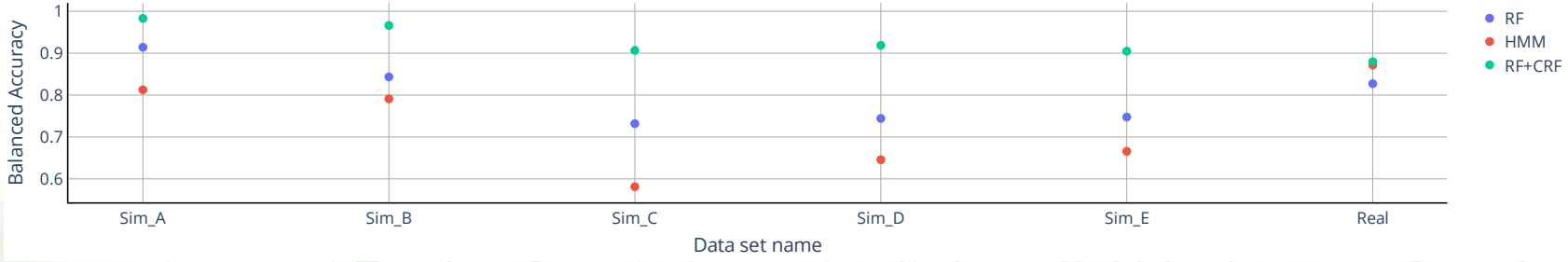
Level measurement of example batch from simulation

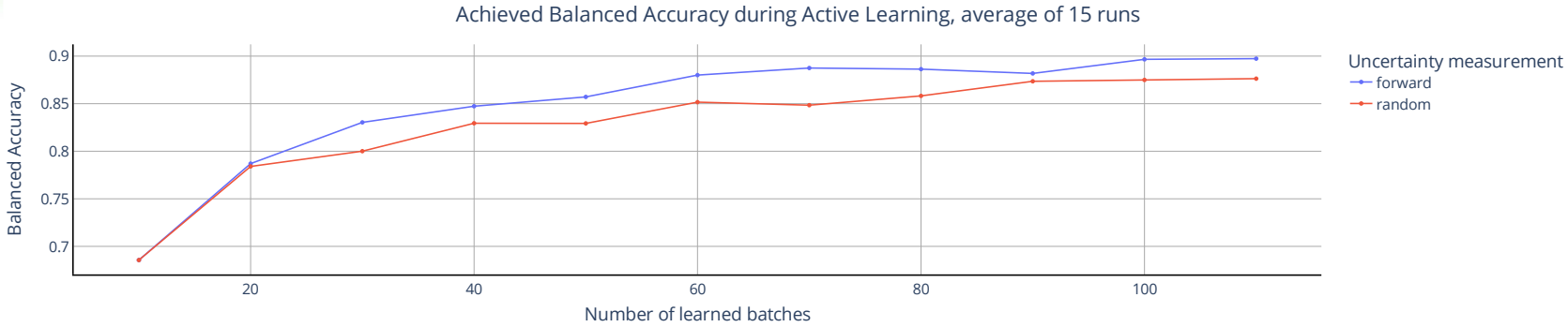
[1] https://doi.org/10.57826/KEEN/ODU6MA

Combining graphical model and active learning



Numerical Experiment





Use case: Heat Exchanger Fouling Cycle

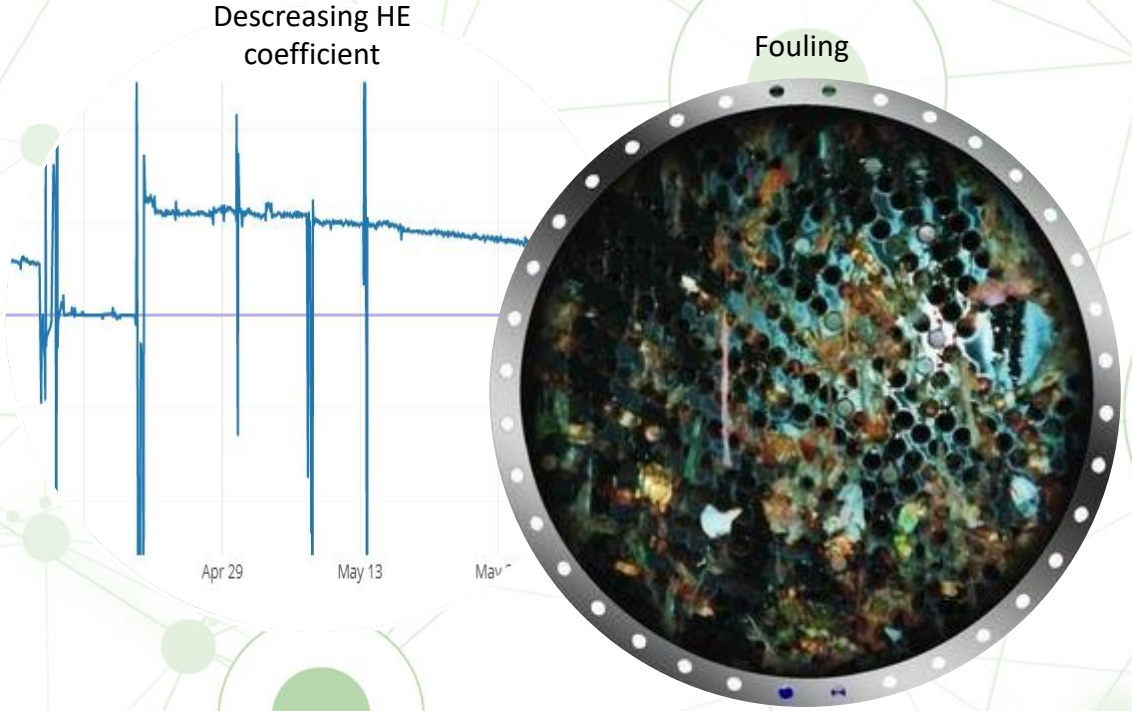
Heat exchanger fouling is a very common problem in the industry

For analysis and root cause determination it is important to analyses the segments between cleaning – ”fouling cycles”

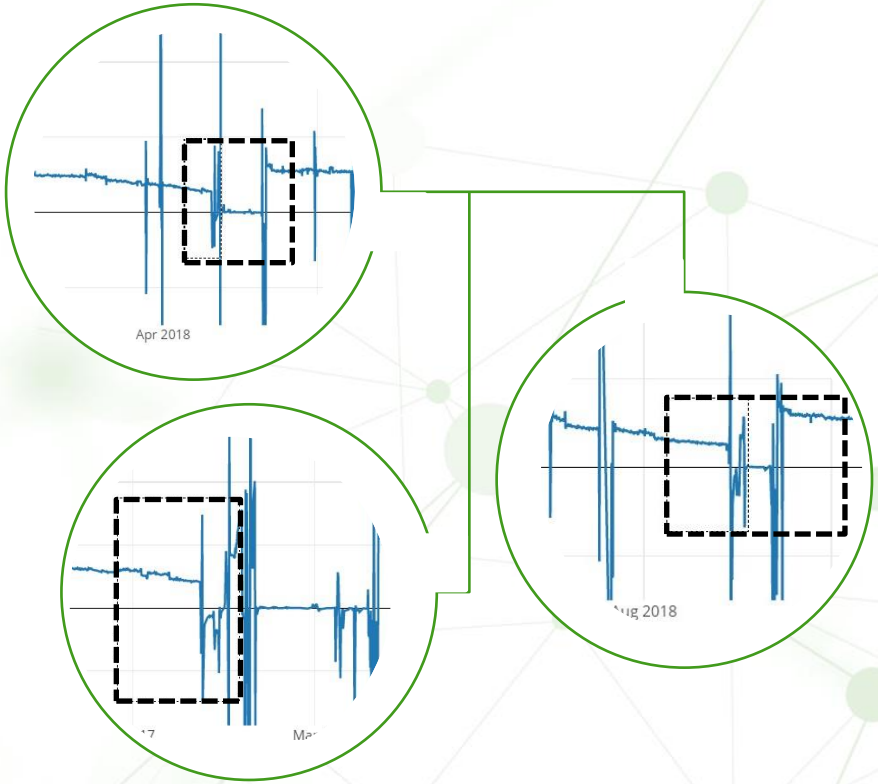
Two challenges exist to determine those segments efficiently at scale:

1. High oscillation and intermediate shutdowns (without actual cleaning taking place) rule out the use of “rule-based” segmentation
2. Operators register cleaning, but in a non-standardized way

Assess the use of the Active Learning ML algorithm to segment our data and perform analytics at scale.



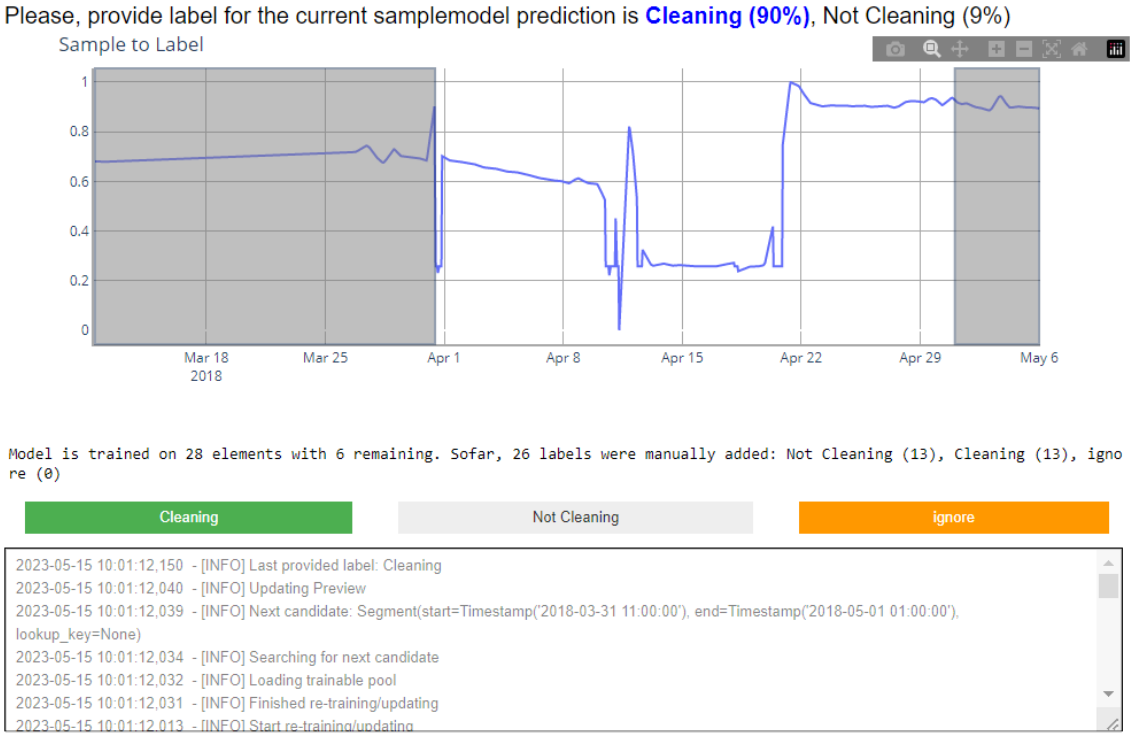
Use case: Heat Exchanger Fouling Cycle



Initial sample selection

- Initial samples and segments are chosen to include a few days of operational activities before and after the actual potential cleaning sample
- Note that the samples are selected based only on the Heat Transfer Coefficient, yet the Active Learning Model could also include other HE attributes
- Data within the potential segments are scaled between 0 and 1.

Use case: Heat Exchanger Fouling Cycle

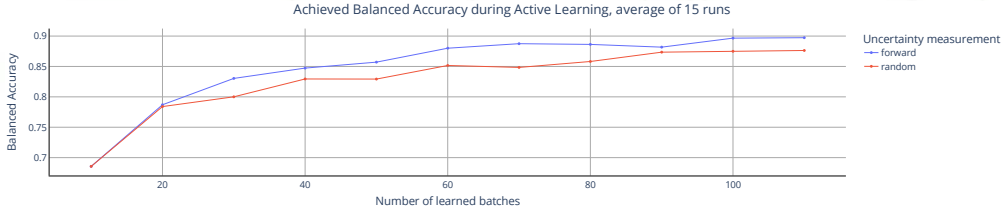


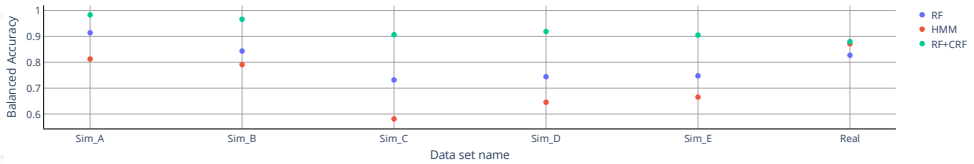
Active Learning Loop

- Yet iterating through 10 more potential segments yields satisfying probabilities of 90%.

Scalability

- Application of the data model to other heat exchangers needs to be further evaluated.





Conclusion

Promising technologies, difficult to research

Conclusion & Outlook

- Active Learning
  - Exploit domain knowledge
  - New/improved economics
  - Expert-in-the-loop
- Transfer learning
  - Persistent process knowledge management
- True value unfolds in an interactive environment
  - Emulated oracles
  - Offline trial with KEEN members
- Outlook
  - Integration into analytics tooling
  - Simulation models (pre-training)
  - Benchmark data


Questions?
Comments!