Foto: Volker Emersleben

Condition Monitoring & Transfer Learning

Good predictions in situations with (initially) almost no data

DB Systel GmbH | Dr. Daniel Germanus, Felix Bert | FOSDEM | February 2019

DB Systel
Digital bewegen. Gemeinsam.

Background

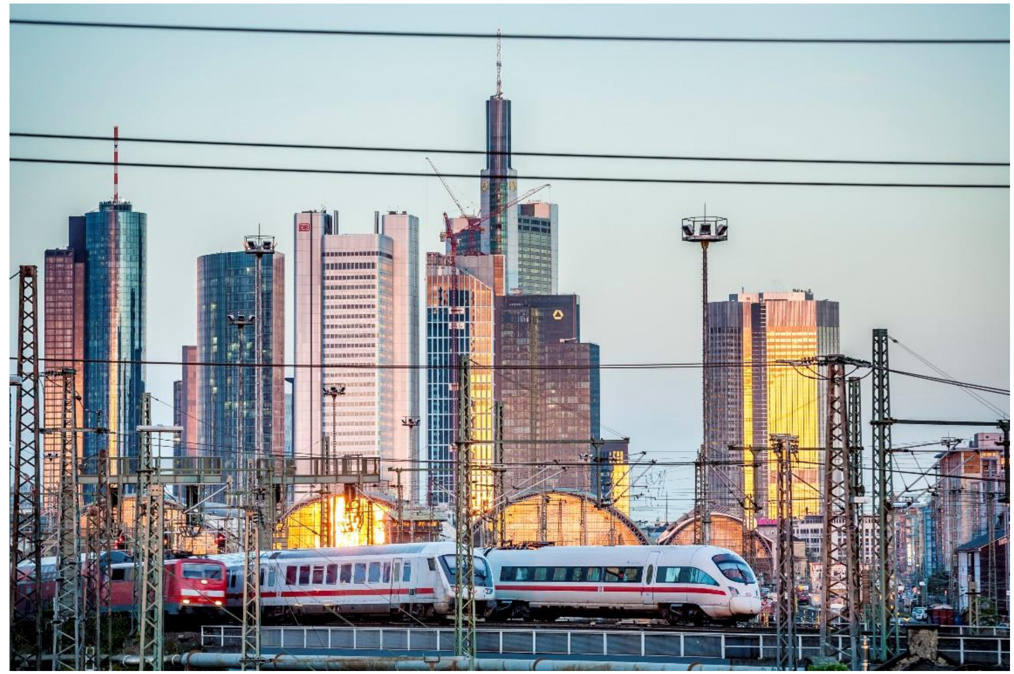
- Condition Monitoring is a precondition to achieve predictive maintenance!
- What kind of Deutsche Bahn equipment could be monitored?
- What kind of sensor seems universal?
- We've founded a DB Systel Venture called Acoustic Infrastructure Monitoring and listen to our equipment!

AIM

Challenges galore

- Generalization
- Little data (in the beginning)
- Little expert time
- Immediate expectation of cost savings

- We chose a machine learning approach
  - But: machine learning is also a tricky subject!
  - Today we present transfer learning to leverage a quick start with the customer



- tl;dr: equipment breaks, we detect it early on using microphones and apply transfer learning to do it even better than w/o ;-)

AIM

Condition Monitoring Goals:

1. **Decrease maintenance costs**
2. **Optimize personnel placement**
3. **Increase availability**

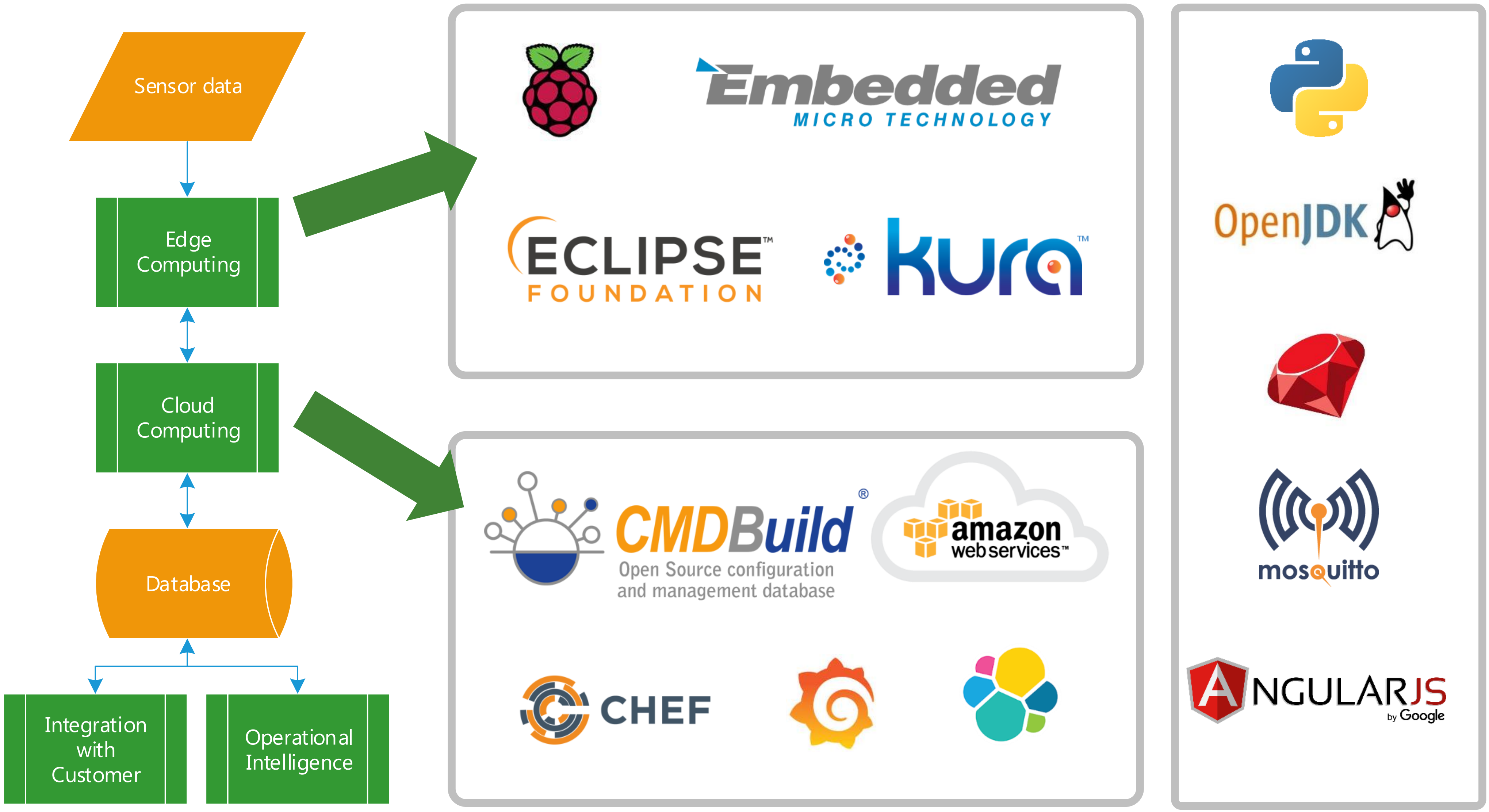
[1] „Condition Monitoring and Diagnostics of machines - Vocabulary“ in ISO 13372
[2] „Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines; Bearings, Pumps, Gearboxes, Engines and Rotating Structures” in The Shock and Vibration Digest, Vol 38(1), 2006, David Mba and Raj B. K. N. Rao

Transfer Learning Goals:

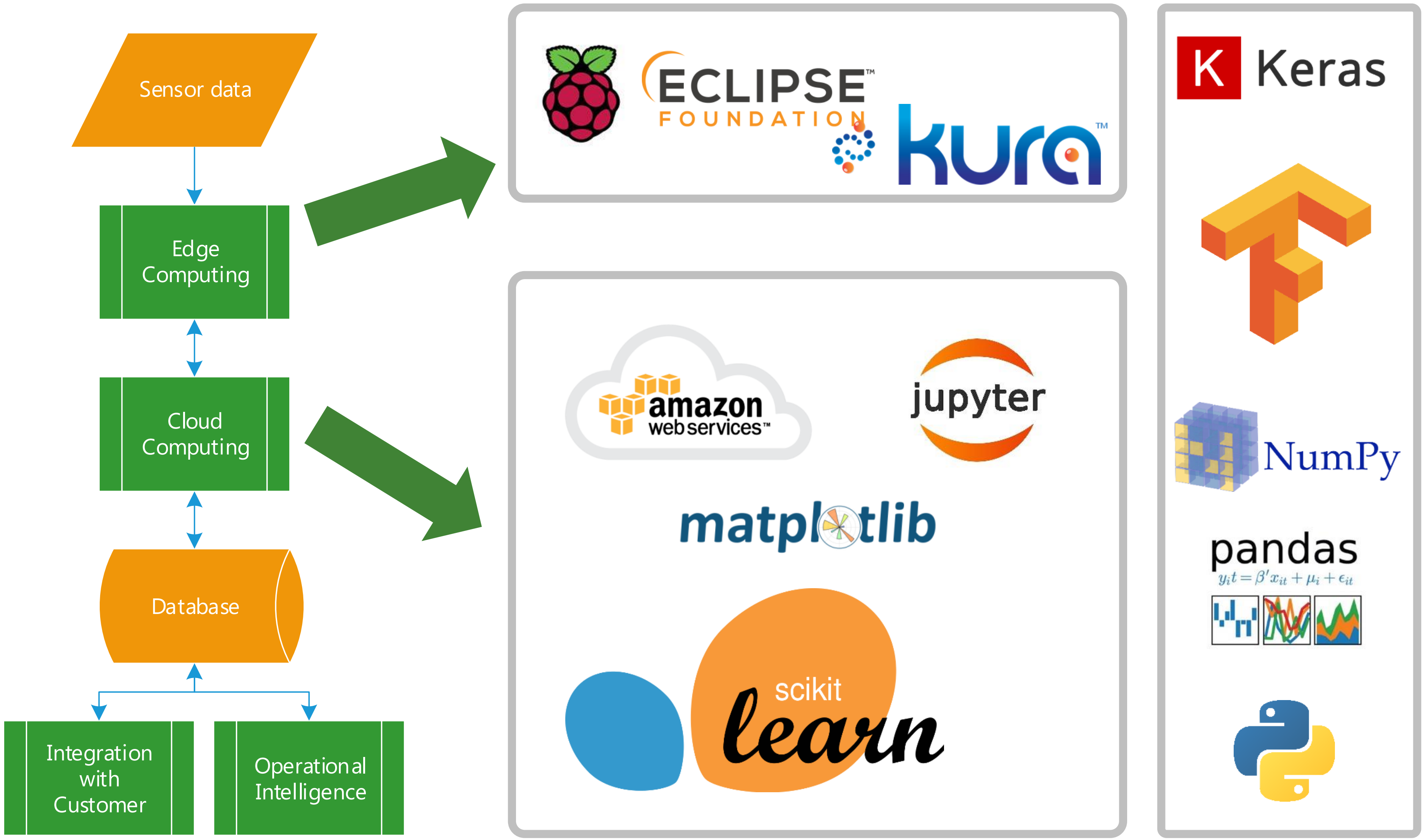
1. **Increase prediction accuracy**
2. **Quick start with customer**

[3] „Transfer Learning will radically change machine learning for engineers" direct quote of Andrew Ng at NIPS 2016
[4] „Deep Learning" MIT Press, Ch. 15, 2016, Ian Goodfellow, Yoshua Bengio, and Aaron Courville

System architecture: service delivery

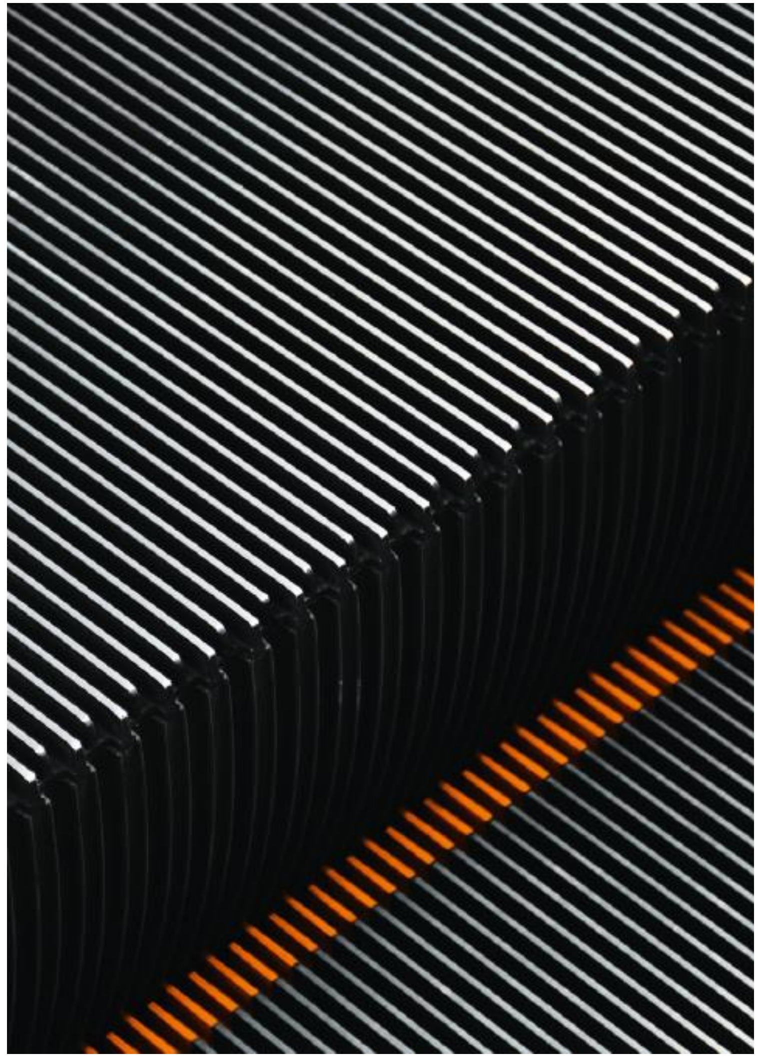


System architecture: data & analysis pipeline



Example equipment: escalators

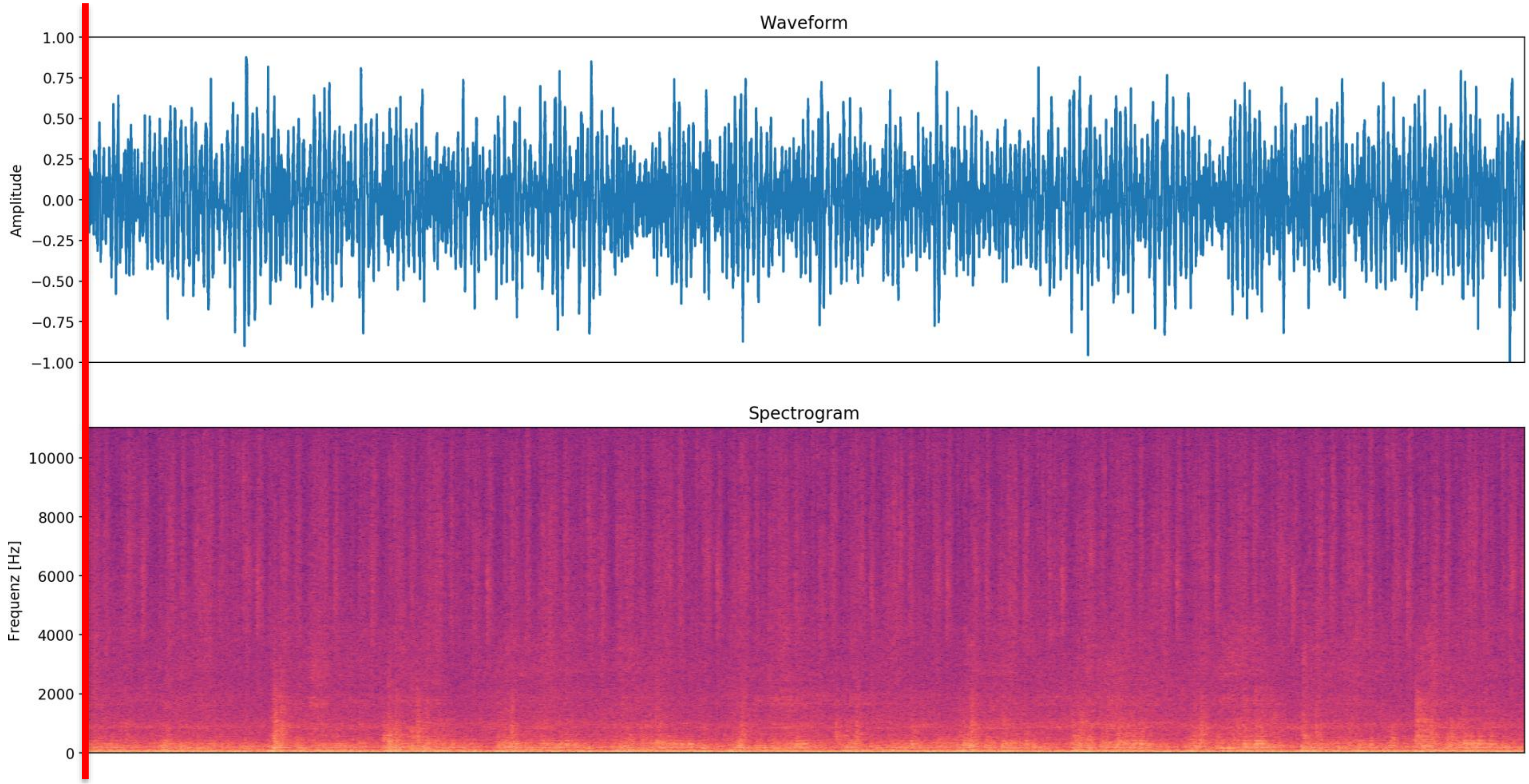
- DB operates ~1000 of those in .de
- Escalator failures result in high material and personnel costs
- Also, due to accessibility, contractual penalties are raised in case of inavailability 0600-2200
- Some failures kick in really fast → immediate detection important!



Example equipment: escalator failures

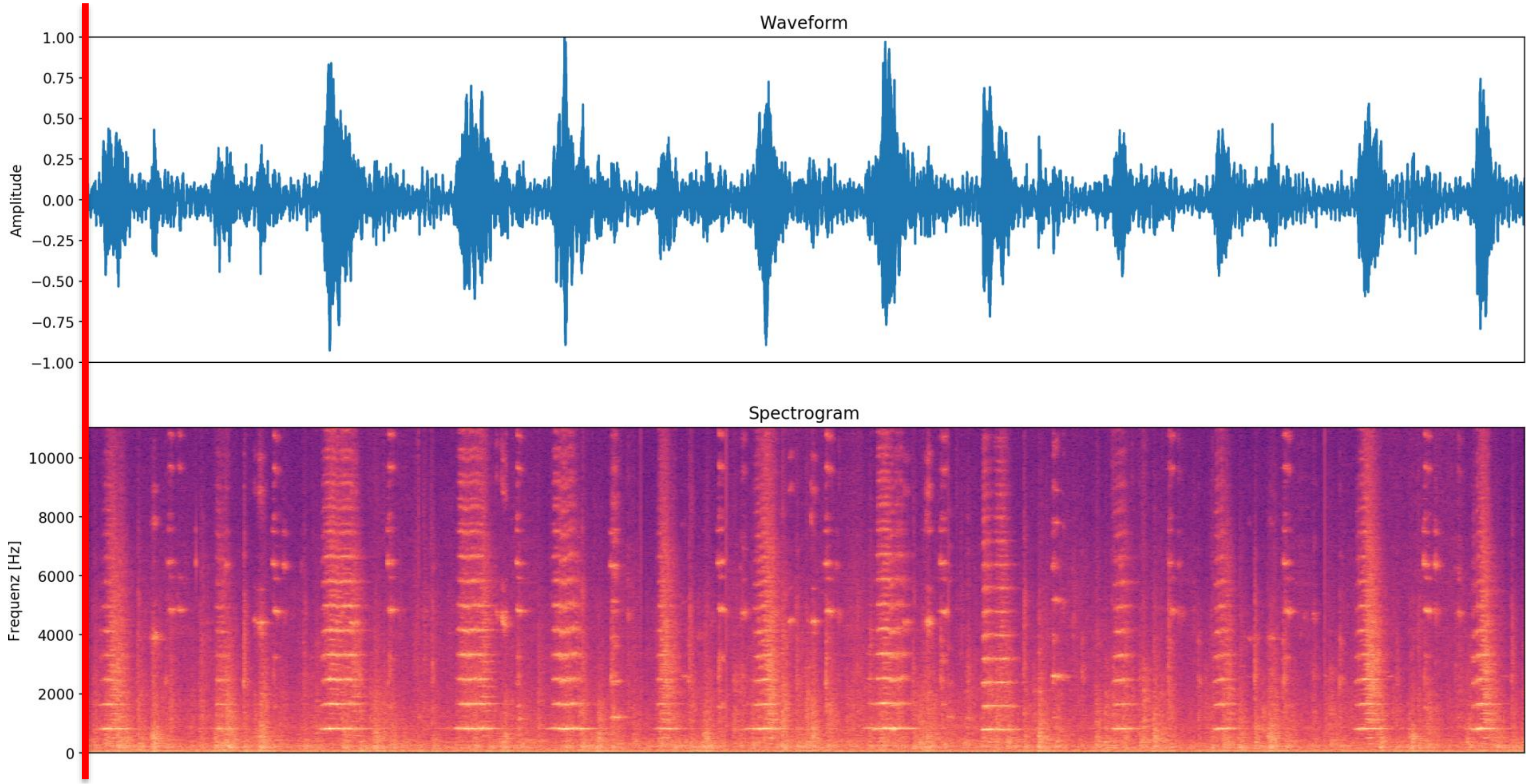
- Failures include
  - Foreign bodies intrude steps/combs
    - Coins
    - Glass
    - Crushed gravel
    - Screws
  - Steps and guiding rails wear off
    - Heavy lifting for years
  - Propagation to other parts of the machinery

Example equipment: escalator sound sample



Hamburg: good case

Example equipment: escalator sound sample

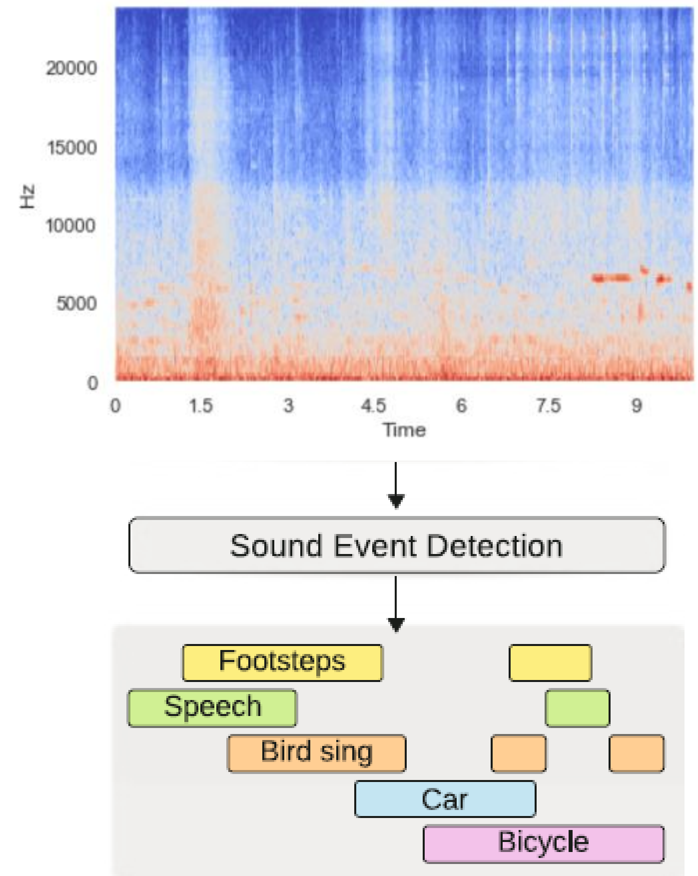


Hamburg: squeaks due to poorly adjusted steps

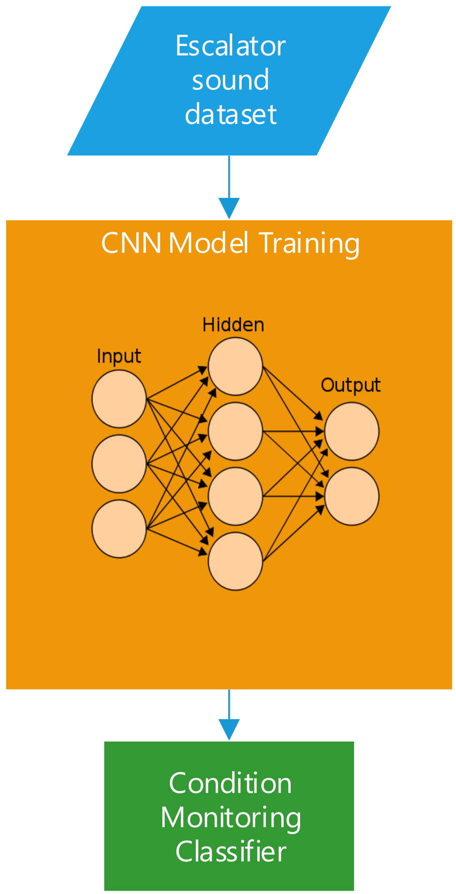
Machine learning approach: sound event detection using convolutional neural networks (CNNs)

- CNNs established for predictions on images
  - we feed spectrograms
  - Annotations do exist for severe failures and their (audible) preconditions
- CNNs provide classification
  - Likelihood of a failure precondition being active
- Do postprocessing in order to reduce oscillation!

❖ Now, how could transfer learning (TL) help?
  - ❖ Little data, grouching customers!
  - ❖ Data collection is lengthy and expensive



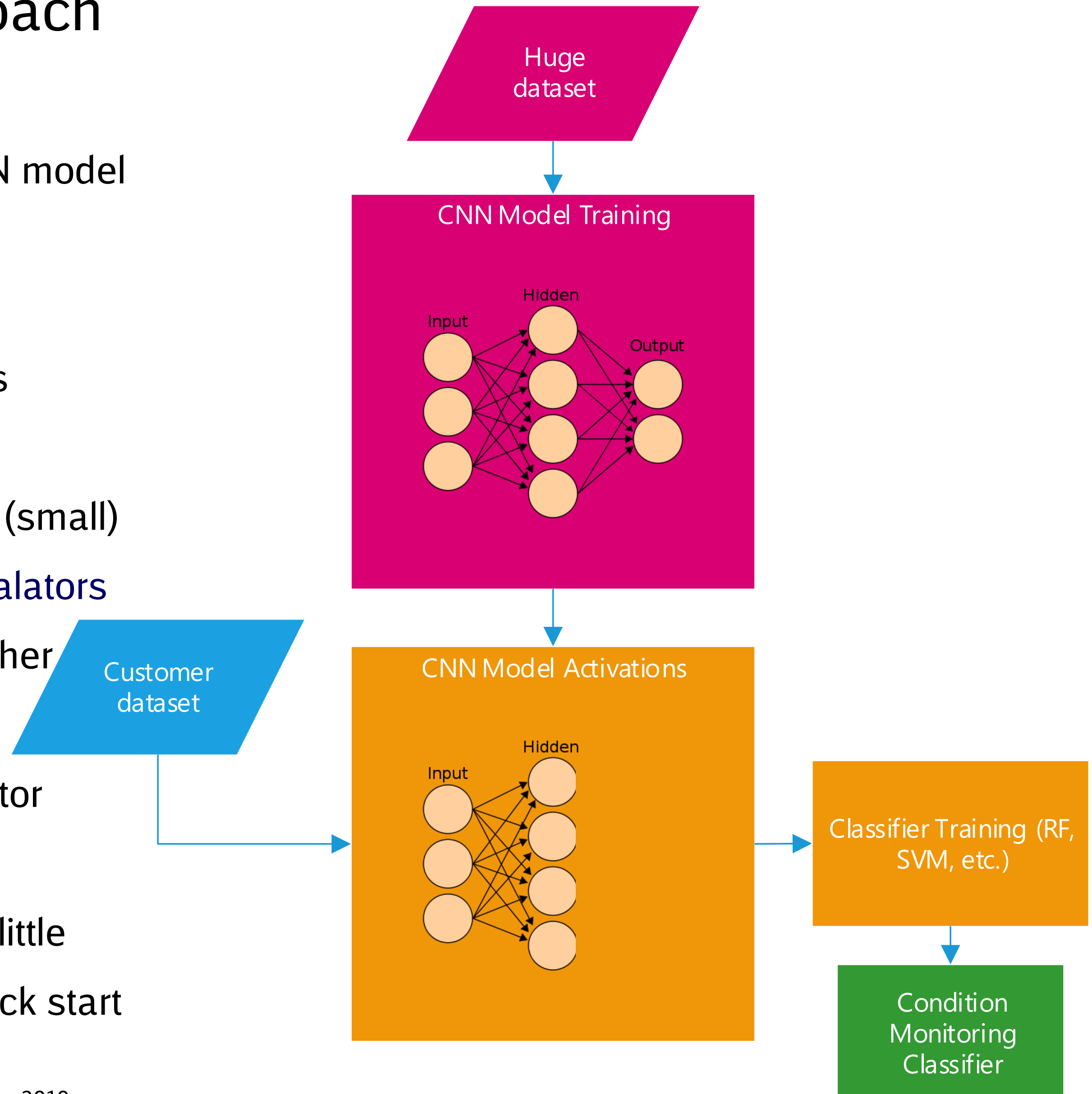
Back to the escalator case: deep learning opposed to our transfer learning approach



- CNN training and prediction drawbacks:
  - Requirements on minimum dataset size
  - Retraining required for
    - new sound events
    - new/adjusted annotations

Back to the escalator case: deep learning opposed to our transfer learning hybrid approach

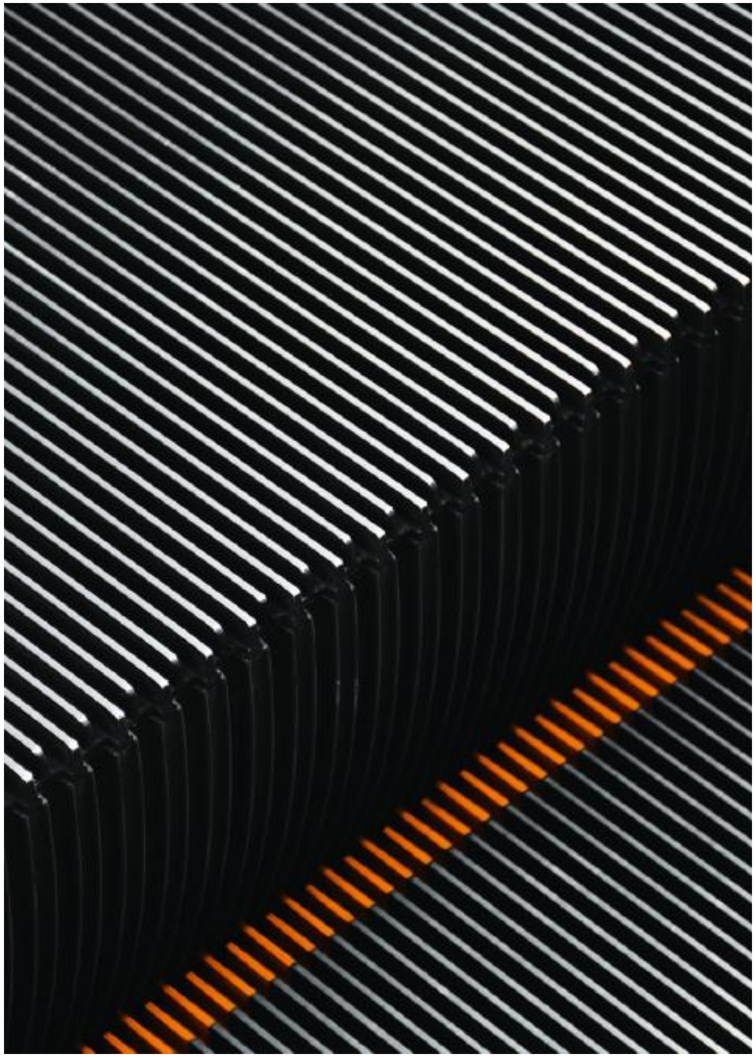
- Transfer Learning
  - Train using huge dataset for base CNN model (train once, no recent customer data): Imagenet, AudioSet
  - Variety of evaluated CNN architectures include: InceptionV3 and VGG16
  - Pick CNN model's activation on actual (small) customer data set: DCASE17, DB escalators
    - Pick activations in order to train another classifier
      - Random Forest (RF), Support Vector Machine (SVM), etc.
      - Predictions possible even for very little customer data, allows ramp up/quick start



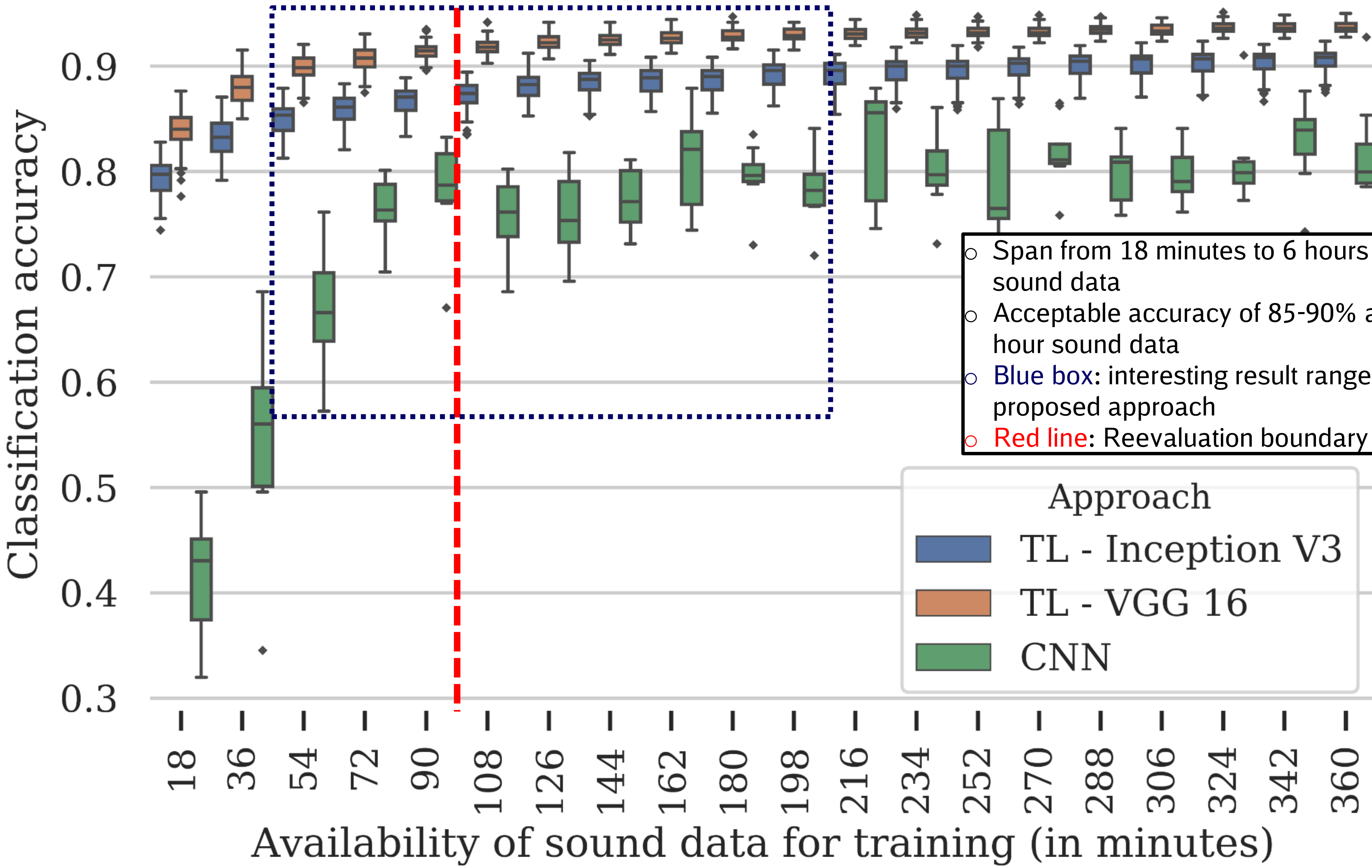
Evaluation overview

- Chosen parameters for the evaluation experiments shown on the next slides:
  - Huge dataset: ImageNet
  - Network architectures: InceptionV3, VGG 16
  - Customer dataset: DB Escalators
  - Classifier: Random Forest

- Overall evaluation goals:
  - Identify accuracy of pure NN and TL hybrid approaches
  - Identify dataset size ranges for which either of the two approaches is preferrable

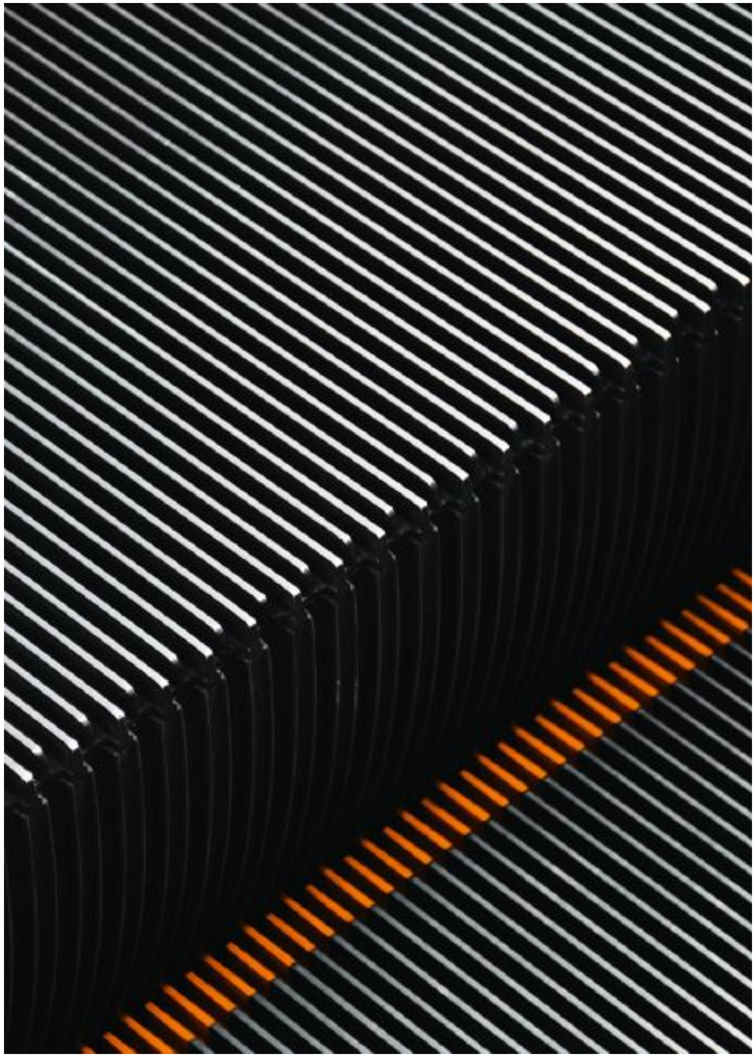


Cross-validated evaluation results on DB escalators dataset



- Span from 18 minutes to 6 hours of training sound data
- Acceptable accuracy of 85-90% achieved at ~1 hour sound data
- Blue box: interesting result range for the proposed approach
- Red line: Reevaluation boundary (NN vs. hybrid)

Conclusion & limitations



- Customer perspective: reduce time to market significantly (for appropriate use cases)
- Business perspective: less expert time required for initial data labelling
- Technical perspective:
  - improved accuracy on small datasets
  - Possibility of choosing classifiers insensitive to overfitting
- Limitations:
  - high variance for *very* small datasets (< 30m)
  - hybrid approach’s advantageous range is use case dependent

Next steps

- Determine the approach's suitability for various use cases over 2019/2020
- Preparation and provision of a dedicated "huge audio data set" based on DB condition monitoring use cases
- Assess the approach's suitability for IoT-like edge computing (learning at the edge, low bandwidth scenarios)

Comment to the Program & Session Chair:

We've prepared a Jupyter Notebook featuring a tutorial, there are at least two possibilities:

1. **We just provide a link to github (no additional time)**
2. **Walk through with audience (+10 minutes)**

So either we stick with 20 minutes for the talk or extend it to 30 minutes in total including the walk through. Let's get in touch.

Thanks for your attention – time for Q&A!

M.Sc. Computer Science
Dr. Daniel Germanus

Chief Architect Machine Learning
Strategic Architecture Management

Tel. +49 69 265-28267

daniel.germanus@deutschebahn.com

DB Systel GmbH
Jürgen-Ponto-Platz 1
60329 Frankfurt am Main
www.dbsystel.de

B.Sc. Engineering Management
Felix Bert

Data Scientist
Application Architecture

Tel. +49 69 265-28267

felix.bert@deutschebahn.com

DB Systel GmbH
Jürgen-Ponto-Platz 1
60329 Frankfurt am Main
www.dbsystel.de