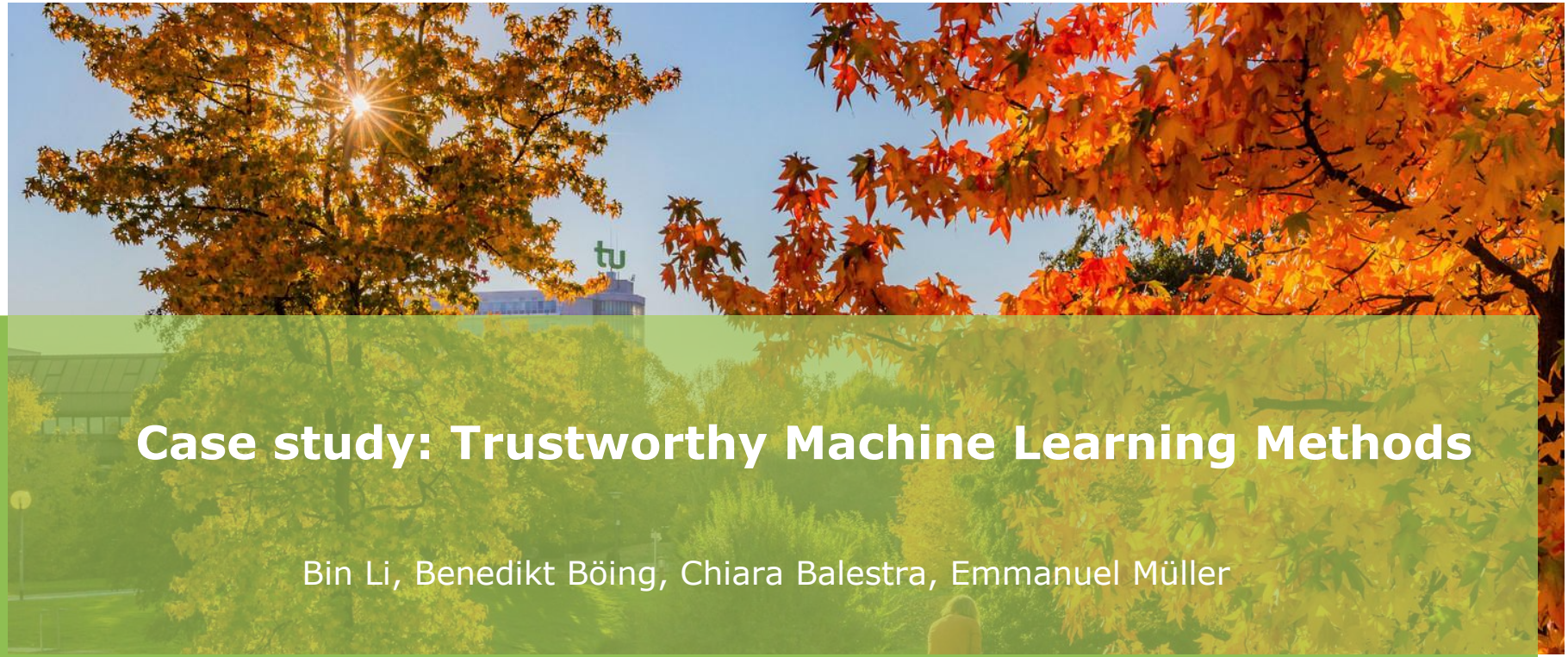
technische universität
dortmund

Chair of Data Science and
Data Engineering
Prof. Dr. Emmanuel Müller



Case study: Trustworthy Machine Learning Methods

Bin Li, Benedikt Böing, Chiara Balestra, Emmanuel Müller

Topic

- **Trustworthy anomaly detection on sensor data in safety-critical infrastructures**
  - **Target**: identify data patterns that are *highly deviating*, *unexpected* or *unusual* in comparison to the overall data distribution or local (temporal) context
  - **Existing solution**: Recurrent Neural Networks + Autoencoders
  - **Challenge**: Lack of secure verification for predicted anomalies

Objectives

- **Design and implement an Autoencoder for anomaly detection**
  - EndDec-AD [Malhotra et al. 2016]
  - DAGMM [Zong et al. 2018]
- **Raise trust into the autoencoder**
  - Adaptivity: update model according to data distributional changes
  - Interpretability: Enrich predicted anomalies with explanation (e.g., feature importance, causal reasoning, uncertainty quantification)
- **Empirical evaluation**
  - Experimental analysis and comparison with selected competitors
  - Human-understandable model interpretation

Organization

- **Target group**
  - Master student of Data Science
  - Max. Capacity: 16 participants
- **Conditions of participation**
  - Interest in anomaly detection and explanable machine learning
  - Attendance of the lectures “Big Data Analytics”, “Machine Learning Paradigms for Complex Data”
  - Advanced programming skill in Python, familiar with deep learning frameworks (e.g., PyTorch)
- **Registration**
  - Email to bin.li@tu-dortmund.de (before September 20) with name, matriculation number, subject and TU Dortmund email.
- **Evaluation**
  - Weekly pitch and discussion (active participation)
  - Final report (20-25 pages) and oral presentation (20 min. + discussion)