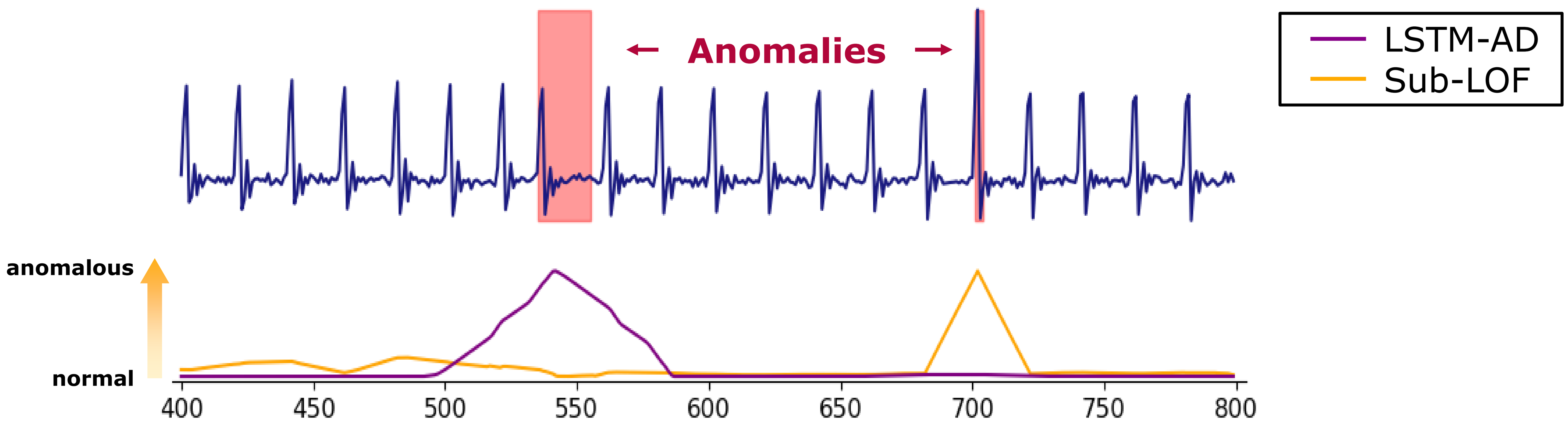
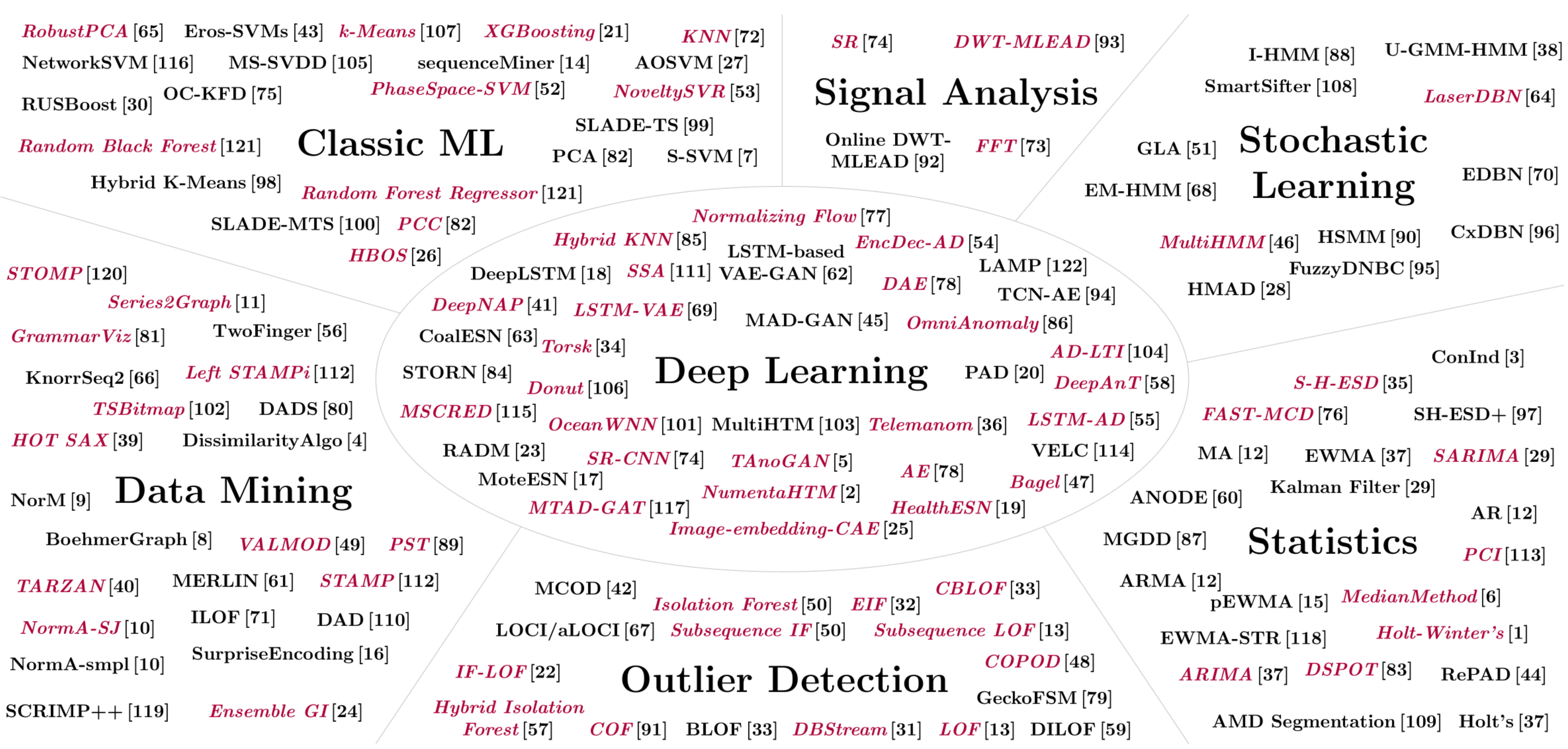
Anomaly Detection in Time Series: A Comprehensive Evaluation

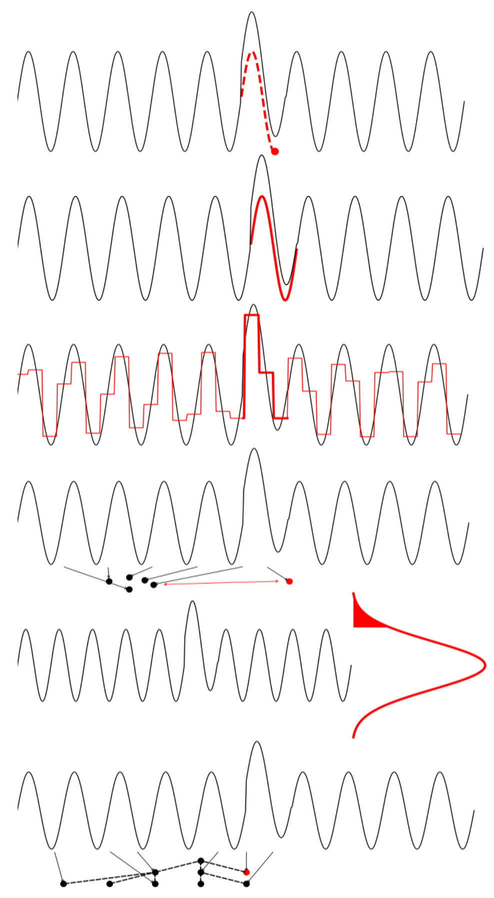
Detecting anomalies in time series is of central interest in many areas because anomalies can indicate important events, such as production faults or heart flicker. Data scientists have developed more 150 specialized algorithms for the automatic detection of anomalous subsequences to deal with the time series' size and complex patterns. However, choosing good algorithms for specific use cases is difficult because no comprehensive study that systematically evaluates the different approaches exists. In our comprehensive study, we carefully evaluate 71 state-of-the-art anomaly detection algorithms on 967 time series datasets.



71 out of 158 Algorithms



6 Method Families



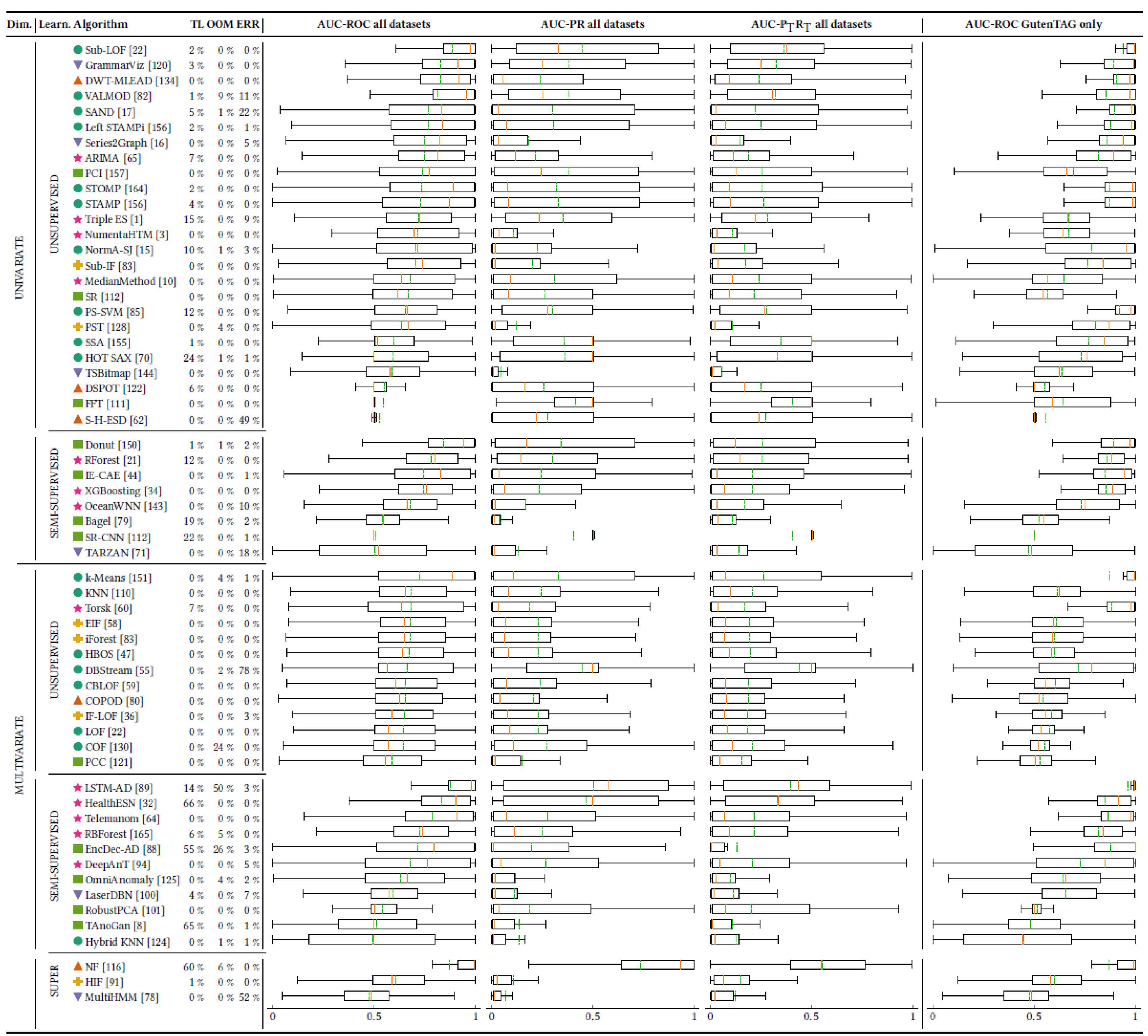
- ★ Forecasting
- ■ Reconstruction
- ▼ Encoding
- ● Distance
- ▲ Distribution
- ✚ Isolation Tree

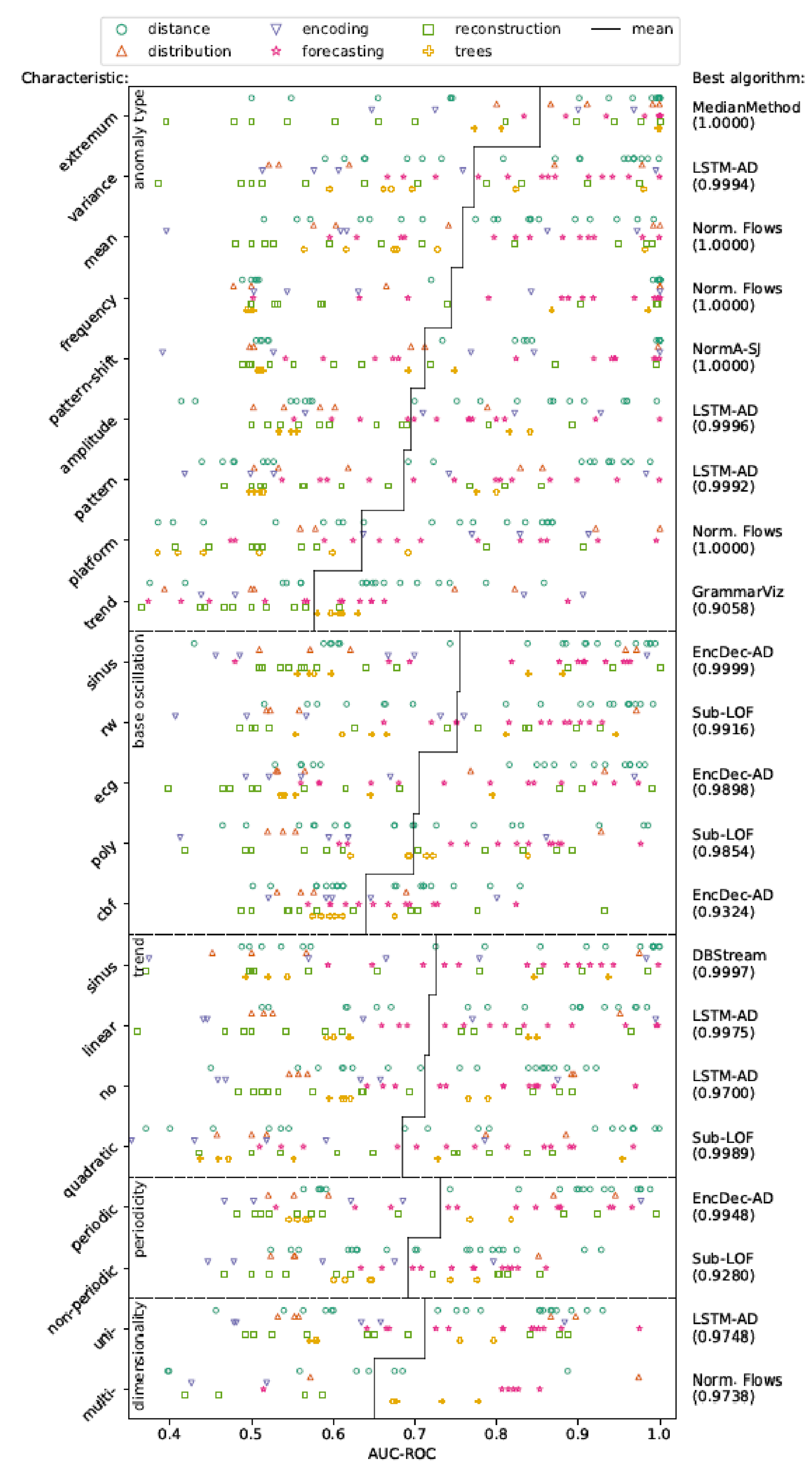
967 Datasets

Collection	Datasets
CalIt2	1
Daphnet	3
Exathlon	2
Genesis	1
IOPS	4
KDD-TSAD	249
MGAB	10
MITDB	4
NAB	56
NASA-MSL	16
NASA-SMAP	35
SMD	23
SVDB	16
WebscopeS5	360
GutenTAG	187

Evaluation Results

- Judgement heavily depends on metric, dataset, and use case. Thus, there is no clear winner!
- Overall high error rates despite our strong investment.
- Hyperparameter optimization is important, but also very difficult.
- Algorithms have their individual strengths and weaknesses.
- Simple and fast algorithms are very competitive.
- Robust and effective algorithms: Sub-LOF, GrammarViz, DWT-MLEAD, k-Means, and Telemanom.
- Deep learning algorithms are not superior to classic algorithms.
- Supervised algorithms are not superior to semi-supervised or unsupervised algorithms.





Sebastian Schmidl
Phillip Wenig
Thorsten Papenbrock

Information Systems Group
Hasso Plattner Institute, University of Potsdam
Potsdam, Germany
E-Mail: {firstname.lastname}@hpi.de

https://github.com/HPI-Information-Systems/TimeEval

doi:10.14778/3538598.3538602

HPI Hasso Plattner Institut