Machine learning and deep learning approaches for multivariate time series prediction and anomaly detection

Thill, M.

Propositions

pertaining to the dissertation

"**Machine Learning and Deep Learning Approaches for Multivariate Time Series Prediction and Anomaly Detection**"

by Markus Thill

1. The problem of anomaly detection (in time series) is harder than other machine learning problems since usually, no target function exists which defines the notion of abnormal and since anomalies are rare events, resulting in highly imbalanced data sets with only a few anomalous examples. [*Chapter 2*]

2. Time series produced under non-stationary conditions require stable online-adaptable anomaly detection algorithms which can track the concept drifts in the data. [*Chapter 4 & 5*]

3. In many cases, it is beneficial to analyze a time series simultaneously at different frequency/time scales in order to detect anomalies reliably. [*Chapter 5 & 7*]

4. Learning long-term dependencies in time series is hard and requires models which can create a large temporal receptive field without significant changes in the model complexity. [*Chapter 7*]

5. Quasi-periodic time series such as electrocardiograms cannot be predicted (in a sense of forecasted) accurately in practice. Allowing a prediction-based anomaly detection algorithm to correct itself even after training to a certain extent by utilizing the true time series values can improve the overall performance. [*Chapter 6*]

6. The practicability and acceptance of anomaly detection algorithms in most real-world applications are strongly dependent on their false alarm rates. A system that produces even moderate false alarm rates will be rendered useless in many cases.

7. The performance of time series anomaly detection algorithms cannot be assessed with common metrics (such as accuracy) known from supervised learning tasks.

8. In the future, the success of anomaly detection algorithms will be decided by how well they can incorporate sparse (expert) feedback into their models.

9. The currently available time series anomaly detection benchmarks are insufficient and only partially suited to evaluate and compare different algorithms. The research community needs an "ImageNet" counterpart for anomaly detection in time series and data streams.

10. The ability to honestly admit one's ignorance and to permit doubt about one's own positions is a virtue that is rather hard to learn but, in the long run, more satisfying than living with wrong answers.