Leibniz Universität Hannover

Wirtschaftswissenschaftliche Fakultät

Institut für Wirtschaftsinformatik

Cloudbased Anomaly Detection in Periodic Signals with Artificial Neural Networks

Masterarbeit

Zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Aaron Brinker

[redacted]

Prüfer: Prof. Dr. Michael H. Breitner
Betreuer: Jens Passlick

Hannover, 22.November 2018

Contents

1 Introduction

1.1 Introduction

Operational robustness of industrial machines and production systems is a decisive factor for the profitability and competitiveness of an industrial company. Condition-based maintenance (CBM) of machines and equipment can significantly lower the cost for maintenance and risk for unexpected failure. The prediction of anomalous states and potential failures allows the management to schedule maintenance efficiently. However, still to this date a widely adopted policy in manufacturing is system overhaul and replacement of critical components at fixed time intervals [1].

A series of recent developments in different fields, like cloud computing, Internet of Things (IoT) and Machine Learning (ML), creating a paradigm shift that is opening up significant new business opportunities. Cloud computing enables delivery of computing services over the internet, on demand from a remote location. The main economical advantages of cloud computing are the liberation of business owners from the need to invest in hardware infrastructure and the decrease of operating cost. Moreover, several technical benefits such as energy efficiency and hardware optimization are provided by cloud computing [2].

Further, the emerging IoT technologies for interconnected machines and sensors in the industrial environment show promising potential for the realization of CBM. The advancements in microelectronic technology lead to a decreased size and cost in sensors. Moreover, the access to real-time data is increasing due to wireless technologies [3]. One of the most important outcomes of this emerging field is the creation of an unprecedented amount of data. It remains a difficult task to gain relevant insights from the data and implement algorithms for reliable detection. Practical approaches are highly dependent on the domain knowledge of engineers. Decision making rules are crafted manually, and are integrated into an expert to monitor machine behavior. The implementation of such an expert system can be costly with regard to time and required domain knowledge.

In order to generate knowledge from data and support decision making, ML algorithms can be utilized to extract useful information from a complex sensor data stream. A research focus in ML is the area of Deep Learning (DL), which aims to learn features automated at multiple levels of abstraction and therefore allows a system to learn complex functions [4]. Complex problem settings where the time dependency in multivariate sensor data is required to detect anomalous machine behavior, require novel anomaly detection methods. Even though the considerable amount of literature published on DL in different areas, the generalizability of much published research on this issue is problematic. Regarding anomaly detection in multivariate time series, up to now there has been little agreement on which algorithms are effective for which application. Since, the crucial fact in success of CBM is an accurate method to access the machine condition, this research seeks to address the following question:

> How does an anomaly detection scheme for periodic signals look like using Deep Learning algorithms?

In this thesis recent DL algorithms for anomaly detection are examined, within the context of a real world industrial application. Experiments are conducted with multivariate sensor data stemming from a conveyor belt.

1.2 Problem specification

The aim of this thesis is the examination of DL anomaly detection algorithms, using a particular data set. In order to clarify the task and specify the objective, this section defines the term anomaly detection and related challenges.

Anomalies are defined by Chandola et al. [5] as patterns in data that do not correspond to a defined notion of normal behavior. By applying this definition to machine sensor data, it can be narrowed to detect patterns that do not correspond with the expected behavior of the monitored machine. The task requires a sophisticated understanding of machine behavior and the ability to identify significant characteristics in the sensor data, in order to distinguish normal from anomalous behavior. In the manufacturing industry operational efficiency often suffers from machine downtime and frequent machine failures. Modern industries have sensor integrated machines and can rely on a huge amount of machine health related information. The subsequent step, to properly utilize the knowledge hidden in the data, is currently mostly performed manually by engineers and requires domain knowledge and specific machine expertise of the particular application. The increasing amount and complexity of data makes this process costly in time and money. Further, when facing multivariate time dependent data streams from a variety of sensors, the processing by humans gets increasingly unmanageable. Besides the amount of data, several challenges need to be considered for anomaly detection. Chandola et al. [5] identify the following challenges which prevent a straightforward anomaly detection approach:

- The definition of normal behavior is not a trivial task, since the boundary between normal behavior and anomalous behavior is not always clear.

- The notion of normal behavior is not a fixed entity. Rather, it is evolving and might change in the future.
- The availability of labeled data for normal and anomalous samples.
- Data usually contains a lot of noise, that tends to be similar to actual anomalies.

In view of these challenges, it is complex to figure out a general solution. In fact, most of the existing anomaly detection techniques solve a application-specific formulation of the problem. The scope of this thesis is limited by the following factors. First, the data was recorded in a single session, so there can be no examination of slow degradation processes and statements regarding machine wear states over a longer period of time. Second, in a real world scenario the conveying length and parameter settings are variable. The present data set is confined to a fixed conveying length and one fixed parameter setting.

In this thesis, anomalous patterns are patterns that are distinguishable from normal patterns and relate to real machines or machine operator errors. Hereby, the main focus in order to solve the task is the application of DL algorithms.

After the initial introduction chapter, the second chapter begins by explaining the methodological approach taken in this study. Furthermore, sections 3.1, 3.2 and 3.3 provide the theoretical background. Subsequently, chapter 4 summarizes the findings of the current developments and research. Experiments were conducted with three different algorithms and the results are presented in chapter 5. In the discussion the findings are scrutinized and interpreted, and limitations of the experiments are outlined. Finally, the conclusion gives a brief summary of important key findings and areas for further research are identified.

7 Conclusion and Future Work

Finally, this chapter concludes the results of the thesis and gives an outlook for future work.

The purpose of this thesis was to determine a DL anomaly detection scheme for multivariate, cyclic sensor data. All experiments were conducted with real data from a conveyor belt. The definition of anomaly detection includes the finding of anomalous patterns in data. There are several ways to address this problem. However, based on current literature, promising approaches in various fields of ML and DL for similar tasks were identified and applied.

A ML model solves the task by learning the underlying distribution of normal data and recognize anomalies based on the similarity of new data to the expected distribution. On the other hand, if a sufficient amount of anomalous data exist and only those failures are relevant, a basic classification algorithm can be utilized to recognize the different patterns. Since faulty data is difficult to simulate and therefore not available in a sufficient amount, a simple classification algorithm was no viable option.

Ultimately, two DL approaches were selected for further experimentation. A LSTM forecasting model and an AE reconstruction based method. In addition, a PCA was performed in order to establish a comparison for the proposed anomaly detection methods. In the LSTM forecasting model, the anomaly score was derived on the basis of the prediction error. It was shown that the anomaly score is significantly higher for anomalous patterns. The inclusion of time dependent structure with the LSTM leads to a high sensitivity regarding anomalous patterns. Nevertheless, due to the sequential processing, the model has difficulties to detect anomalous behavior that manifest in a whole cycle, under the condition that the individual patterns of the cycle do not contain much deviation to normal patterns.

The AE calculates the anomaly score based on the reconstruction error. Contrary to the LSTM, the AE processes data cycle wise and is therefore well suited to detect anomalies, which poorly manifest in subsequences. This finding is confirmed by the experiments, where such anomalous data lead to high anomaly scores. Moreover, the AE proofed to handle different transport weights and different input features without loosing classification accuracy. These findings confirm that the established AE anomaly detection scheme is well suited for detecting anomalous behavior in cyclic multivariate time series from this specific data domain.

It would be interesting to assess anomaly detection capacities of the proposed schemes, in a data set containing more heterogeneous samples. More training data will improve the generalization capabilities. As a matter of fact, with more data, the models could be build deeper in order to extract more sophisticated features from complex data. Additionally, with increasing data volume and computational complexity batch normalization between layers could be utilized to reduce training time [57].

Since the AE proved to be the best approach for anomaly detection in periodic signals, it is recommended that further research explores variants of this model. Possible extension include Recurrent Autoencoder [61] and Variational Autoencoder [36].

Further research needs to examine more closely the potential complementation between different algorithms to build a hybrid architecture. Different approaches from ML and traditional time series analysis can be combined to extract meaningful features. Hybrid models like this could combine the advantages of feature extraction in time and spectral domain. In addition, performance can be improved by combining multiple models together, instead of just using a single isolated model. Such combinations of models are called committees or ensembles [15, p. 653].