Data-based models of drive technology for automation in automotive production

Eva Dierkes¹,*, Francesca Jung¹, and Christof Büskens¹

¹ Center for Industrial Mathematics, University of Bremen

The digital revolution, especially in the field of manufacturing, has great potential to change the economy sustainably. In this work, the development of methods for predictive maintenance and condition monitoring is a central focus. Data-based models for drive technology in automotive production are investigated in order to generate adequate models, to make statements about the life cycle of the drives, or to suggest system modifications, such as the adjustment of weights. One task in this field is to detect anomalies or disturbances in given engine data. Since this collected data is variable and noisy, the detection of faults is non-trivial. In this context two different methods for anomaly detection are studied. First, a model based on statistical analyses and second, a machine learning model is evaluated.

© 2021 The Authors Proceedings in Applied Mathematics & Mechanics published by Wiley-VCH GmbH

1 Introduction

Due to the progress of technological possibilities, more and more process data is being recorded in many production companies. A large part of this data is not yet being used profitably. In order to be able to benefit from this collected data, models for the process components are needed. Within the scope of this project, models are developed with special attention to the detection of anomalies. In production plants, the drives and their individual parts are subject to high stress and dirt. Therefore the components wear out and this can lead to disturbances in the process flow. The aim is to detect faults, as early as possible, and thus be able to plan maintenance at an early stage in order to prevent unexpected production downtimes.

The application for this contribution is based on automotive production with focus on the conveyor technology. The engines that are used in conveyor belts are normally planned to be maintained at an equidistant frequency, but they are stressed differently depending on location and load. Usually it is not known which vehicle (and in which configuration) is currently being carried so that the load is unknown. Only with great experience it is possible to distinguish between changes in the load or signs of wear and tear. This fact formulates the goal of being able to detect anomalies regardless of the load.

2 Data set

For a simulation of appropriate data, an experimental setup consisting of two engines acting in opposite directions was used. The main engine, which represents the conveyor unit, runs in each recorded run with a constant power between 0 % and 100 % in 10 percent steps. The second one simulates the load, which is constant within one run and is either 0 % or 20 %. For each combination around 33 runs consisting of 818 time samples with a sampling rate of 1000 Hz were recorded. The tests described below are based on the four channels shown in Fig. 1. Another data set is generated to simulate impacts, where the load changes abruptly for a very short period of time. The standardized mean for both data sets are shown in Figure 1. To test the trained models, data sets with power between 0 % and 100 % in 3 percent steps are collected.

---

* Corresponding author: e-mail eva.dierkes@uni-bremen.de

---
3 Anomaly detection

The anomalies in the data are simulated via interruptions from the load engine. That means the load engines run for example constantly with 20% and switches quite regular to 0% for a very short duration. For each power configuration described above also anomalous data were collected. In total this leads to 726 normal and 726 anomalous data sets for training and 506 normal and 520 anomalous data sets for testing.

For the detection of an anomalous data sets we focus on two different approaches. The first one used data analysis to define criteria for the normal behavior and the second uses a simple neural network to learn the classification.

Anomaly detection via statistical analysis As a statistical approach, we define the normal behavior based on the frequency spectrum of the five channels of the signal mentioned above. For the frequencies up to 109.8 Hz the mean value and the standard deviation were calculated for the normal test data set. A signal is classified as an anomaly if for at least two channels at least one value lies outside the fivefold standard deviation. The described threshold is manually adjusted for this data set, which may lead to less generalizability for other situations. It is important to note that the normal behavior is defined without knowledge of the faulty data sets.

Anomaly detection via neural networks Deep learning has achieved high accuracy in many areas of signal processing such as audio and image recognition [2]. In many research contributions convolutional neural networks are used to predict defective engines, e.g. [1]. To keep it simple, a fully connected network is trained in this application. To train the model the whole frequency spectrum is used to demonstrate neural networks are able to focus on relevant data. In this case the input dimension is 540 times 409 frequencies for five channels. The architecture is kept simple and contains two fully connected layers with 256 neurons each. The activation function ReLu is used between the layers and the sigmoid function is used for the outputs. The data set is divided into training (70%) and validation data (30%), moreover the fitting runs over 100 epochs.

4 Results

In Table 1, the accuracy for both approaches for the training- and the two test data sets are shown. The statistical data analysis classifies 100% of the normal data sets as normal and 98.08% of the faulty ones. With an accuracy of 99.21% for the normal test data set and 98.27% on the faulty data set, the neural network performs equally well.

This result is interesting, because the simple network performs well even without handcrafted knowledge of the data. To adapt the model using data analysis, a lot of effort is investigated into setting the hyper parameters, such as how many frequencies are used or how large the threshold should be to differ from faulty to normal data. Nevertheless the supervised learning can only work well if faulty data is available, the statistical approach can work without faulty data.

Sometimes the engine temperature has an influence on the behavior and thus on the selected data. To show that the presented results are independent of temperature influences, data sets with different temperature settings were recorded. The results show that the prediction works well regardless of the temperature.

5 Conclusion

In this paper the comparison of a statistical approach with a deep learning approach to anomaly detection was presented. It turned out that both work equally well, but since deep learning does not require as much manual work as the statistical approach, it might be preferred. Since different loads are contained in the training and the test data the detection of anomalies is independently of the loads.

Acknowledgements This project was funded by the European Regional Development Fund (EFRE). The authors acknowledge Encoway GmbH for creative discussions and the generation of the simulation data. E. Dierkes acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG) – Project number 281474342.

Open access funding enabled and organized by Projekt DEAL.

References

[1] N. Günnemann and J. Pfeffer, Predicting Defective Engines using Convolutional Neural Networks on Temporal Vibration Signals, (PMLR 74:92-102, 2017)
[2] J. Schmidhuber, Deep learning in neural networks: An overview, (Neural networks 61:85–117, 2015)