Linköping University | Department of Computer and Information Science
Master thesis, 30 ECTS | Computer Science
202018 | LIU-IDA/LITH-EX-A--2018/XXX--SE

BEAUTIFUL TITLE, ENGLISH

– SUBTITLE, ENGLISH

VÄLDIGT FIN TITEL

Linus Kortesalmi

Supervisor : Mattias Tiger
Examiner : Fredrik Heintz

External supervisor : Simon Johansson

LINKÖPING UNIVERSITY

Linköpings universitet
SE–581 83 Linköping
+46 13 28 10 00 , www.liu.se

Abstract

Abstract.tex

Acknowledgments

`Acknowledgments.tex`

Contents

Todo list

1 Introduction

Maintenance of devices/vehicles/systems/machines (hereinafter referred to as “units”) is generally done by planned scheduling. Typically this is done when some parameter of the system reaches a threshold value. For example, car maintenance can be scheduled after 30,000 kilometres, after a year or perhaps after a certain number of operating hours. One problem with planned scheduling is the reliance on experience and statistics from many units. For a single unit, the planned scheduling will either be executed too early (could have waited longer before service) or too late (problems encountered before threshold reached).

Internet of Things (IoT) permits the streaming of continuous data from multiple units. The data is typically the state of the unit in the shape of many different variable values. A rule framework can be built incorporating the continuous data. The framework can then give an informed service alert based on the actual state and need of a unit.

Machine Learning (ML) models can be used to, for example, capture dependencies in large-scale data sets (ref needed), anomaly detection (ref), clustering (ref), image recognition (ref), and decision making (ref). Anomaly detection could be used together with continuous real-time data from a system to find unusual changes or behaviour, which could be the basis for a service alert and/or the gathering of new knowledge about a system. ML-algorithms can also be used to perform Predictive Maintenance (PdM).

1.1 Problem Description

1.2 Aim

The expected results is a study on how ML-algorithms can be used for Predictive Maintenance (PdM). The study shall compare at least two different ML-algorithms. The goal is to implement the ML-algorithms and use real-world data from an existing IoT system. The aim of the study is also to compare how well the improved DcM performs with the current solution used in the IoT system (preliminary, more info needed from company as well as ideas how to compare the solution with the practice in use today).

1.3 Research Questions

This work explicitly answers the following questions:

The RQs below are just an initial draft.

1. Which methods can be used to detect anomalies in a dynamic real-time system?
2. How can the need for maintenance be detected from real-time processing of (domain-specific) data?
3. How can predictive maintenance be achieved by processing real-time (domain-specific) data?
4. How far in advance can maintenance be predicted?

1.4 Delimitations

2 Theory

3 Related Work

Predictive Maintenance (PdM) approaches have been studied and tested for many different industries. This chapter covers some of the approaches proposed in the literature.

3.1 Multiple Classifier Approach

Results of the Multiple Classifier approach?

Requirements in order to use it? (type of data, dimensionality, environment (R2F), supervised/unsupervised etc.

In [2], Susto et al. propose a multiple classifier approach to PdM. They use the approach to predict problems which stem from the ”wear and tear” effects of equipment used for semiconductor manufacturing. Each classifier is trained on a different failure horizon m, which results in a different classification problem for each classifier. The failure horizon is the number of iterations in a maintenance cycle where the fault has taken place. In a traditional R2F (Run to Fail) environment only the last iteration would be faulty ($m = 1$). Instead, the dataset is transformed for each classifier so that the last m iterations are marked as faulty. A larger m reduces the skewness of the dataset and enables a more conservative PdM policy. The multiple classifier approach thus enables the implementation of a cost optimisation policy as well as a fault prevention policy.

3.2 Framework based on ARMA and data-driven techniques

Baptista et al. propose in [1] a different approach to PdM utilising usage data instead of sensor data. A framework is built in order to predict the next fault event based on previous events. The usage data (past failures and past scheduled events) is given to the Auto-Regressive Moving Average (ARMA) model, which outputs predictions on future failure events. The predictions from the ARMA model is fed to the data-driven model and transformed, using statistics features and PCA, in order to output a more informed prediction. The data-driven model trains five different classifiers: k-nearest neighbours (k-NN), random forest (RF), neural networks (NN), support vector machines (SVM), and generalised linear regression model (GLM).

The framework is then compared against a baseline approach using a standard life usage (LU) model with the Weibull distribution . Baptista et al. show in their case study [1], that almost all data-driven models outperform or perform comparably with the LU model. The only model to perform worse was the NN model, due to over-fitting [1]. The SVM model

source?

source?

achieves the best results in almost all metrics which, according to Baptista et al. [1], proves that it is possible to build more sophisticated and improved models than the LU model.

Vi kan nog inte använda denna approach, eftersom det inte riktigt är den typ av data som vi har tillgång till. Vi har tillgång till sensor data (och kanske även lite usage data). Dock är usage datan baserad på den tidigare statistiska modellen, så man vet ju inte riktigt ifall alla scheduled events of failure events hände precis då som de står i loggen. Det hade kanske varit intressant att undersöka huruvida man kan kombinera sensor data med usage data?

3.3 title

4 Method

Bibliography

[1] Marcia Baptista, Shankar Sankararaman, Ivo. P. de Medeiros, Cairo Nascimento, Helmut Prendinger, and Elsa M.P. Henriques. “Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling”. In: *Computers & Industrial Engineering* 115 (2018), pp. 41–53. ISSN: 0360-8352. DOI: `https://doi.org/10.1016/j.cie.2017.10.033`. URL: `http://www.sciencedirect.com/science/article/pii/S036083521730520X`.

[2] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi. “Machine Learning for Predictive Maintenance: A Multiple Classifier Approach”. In: *IEEE Transactions on Industrial Informatics* 11.3 (June 2015), pp. 812–820. ISSN: 1551-3203. DOI: `10.1109/TII.2014.2349359`.