
Project Group:
**Machine Learning for Predictive Maintenance**
(Weekly Status Report)

Supervisors:
**Prof. Dr. Eyke Hüllermeier (eyke@upb.de)**
**Tanja Tornede (tanja.tornede@upb.de)**

Name: Selami Hoxha
Matriculation Number: 6853870
E-mail: shoxha@mail.uni-paderborn.de

# Calendar Week: 10 11 March, 2021

## Completed Tasks

1. Created the first version of the class diagram for HIE
2. Created the first version of the sequence diagrams for the three approaches.
3. Together with the team discussed and made changes to the general class diagram.
4. Wrote a short description of the general class diagram on the system design document.

## Challenges

1. When creating the class diagram it was hard to forsee which parameters will each class have.

## Tasks in-progress

1. Finalize the class diagram.
2. Finalize the sequence diagram.
3. Write a description on the system design document for the class diagram and sequence diagrams.
4. Imporve the description of the general class diagram in the system design document.

## Calendar Week: 9 05 March, 2021

### Completed Tasks

1. Wrote the example part in the data format documentation
2. Together with the team worked on the class diagram.

### Tasks in-progress

1. Finalize the data format chapter.
2. Add the class diagrams for the HI chapter.
3. Create the sequence diagrams for the HI chapter.

## Calendar Week: 8 26 February, 2021

### Completed Tasks

1. Worked together with the team to come up with the initial data format for the ML4PdM framework.
2. Wrote the example part in the initial data format documentation.
3. Together with the team redefined the data format based on the feedback we got on the wednesday meeting of 24 February.

### Tasks in-progress

1. Write the documentation for the defined data format.
2. Help the team in the framework design.

## Calendar Week: 7 19 February, 2021

### Completed Tasks

1. Reviewed survey chapters:
   (a) Introduction
   (b) Formal definition of time series
   (c) RUL
   (d) Conlusion
   (e) References
2. Made the final changes to the sruvey based on the second review.
3. Checked my work for spelling and punctuation mistakes.
4. Fixed the dataset references to show in bibliography as a seperate section of bibliography named Datasets (p. 110).

### Tasks in-progress

1. Think and design a format for the data representation.
2. Think and help the team in design of an architecture for the ML4PdM framework.

## Calendar Week: 6 12 February, 2021

### Completed Tasks

1. Completed changes based on feedback for approaches and publicly avaliable datasets of the HI chapter. The changes were mostly about fixing errors, but also included modifying or adding some content.
2. Added a very brief introduction of Principal Component Analysis.

### Tasks in-progress

1. Review survey chapters:
   (a) Introduction
   (b) Formal definition of time series
   (c) RUL
   (d) Conlusion
   (e) References
2. Implement the changes in the evaluation section.
3. Implement the changes based on the second review to get the final version of the survey.

## Calendar Week: 5 5 February, 2021

### Completed Tasks

1. Completed the revision for the alpha version of ML4PdM survey for chapters:
   (a) Introduction
   (b) Formal definition of time series
   (c) Health state classification
   (d) Conlusion
   (e) References
2. Made changes based on feeback I received for the Publicly Available Dataset chapter.

### Tasks in-progress

1. Make further changes based on feedback for the other parts I have written. That includes three approaches and evaluation for HI chapter; also the paragraph for the history of predictive maintenance in the Introduction chapter.

## Calendar Week: 4 29 January, 2021

### Completed Tasks

1. Completed writing the Kullback-Leibler approach for multi-sensor prognostics [1].
2. Completed the evaluation chapter and the publicly available data chapter.

### Tasks in-progress

1. Evaluation of the alpha version of the survey. The chapters that I will cover are:
   (a) Introduction
   (b) Formal definition of Time series
   (c) Health state chapter
   (d) Conlusion chapter

## References

[1] Oluseun Omotola Aremu et al. "Kullback-Leibler Divergence Constructed Health Indicator for Data-Driven Predictive Maintenance of Multi-Sensor Systems". In: *17th IEEE International Conference on Industrial Informatics, INDIN 2019, Helsinki, Finland, July 22-25, 2019*. IEEE, 2019, pp. 1315–1320. DOI: 10.1109/INDIN41052.2019.8972069. URL: https://doi.org/10.1109/INDIN41052.2019.8972069.

## Calendar Week: 3 22 January, 2021

### Completed Tasks

1. Replaced the third approach with an appropriate one.

### Tasks in-progress

1. Add dataset for GRNN and ANN approaches.
2. Fill the evaluation module with the missing metrics.
3. Revise all the content written for spelling and notation mistakes.
4. Check the references for correctness.

## Calendar Week: 2 15 January, 2021

### Completed Tasks

1. Added new content in Evaluation chapter.
2. Added the first datasets in the the public available datasets chapter.
3. Modified some notation in the LSTM-ED approach to correspond with the latest introduced notation for a single time step in a time series.

### Challenges

1. It is challenging to adapt the notation for subsequence of a time series in the way the approach is presented. It is something that might need a bit more thinking in the review phase.
2. Have decided that the fuzzy support vector machine approach will not work. The HI from expert can not be obtained. Hence a replacement should be found.

### Tasks in-progress

1. Find more data sets that can be used by the approaches.
2. Complete the evaluation setup.
3. Replace fuzzy support vector machine approach.

## Calendar Week: 1 08 January, 2021

### Completed Tasks

1. Completed the Fuzzy support vector machines approach.
2. Completed the hierarchical gated unit network approach.

### Challenges

1. During writing some questions about notation came up, which should be clarified in the upcoming meetings.

### Tasks in-progress

1. Extract the description for publicly available data.
2. Write the Evaluation Setup section.

## Calendar Week: 52-53 30 December, 2020

### Completed Tasks

1. Helped Gourav Prakash complete the formal definition of health index.
2. Completed the writing of the TLSM encoder decoder approach.

### Challenges

1. The fuzzy support vector machine approach for determining health index uses discrete values from 1-5 instead of real values from 0 up to 1 as we defined it in the formal definition. The challenge is to adapt the formal definition so it captures this case as presented in [1]

### Tasks in-progress

1. Write about health index Based prognostics using Ensemble of machine learning algorithms.
2. Write about determining health index of a system using fuzzy support vector machines.

## References

[1] A. D. Ashkezari et al. "Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers". In: *IEEE Transactions on Dielectrics and Electrical Insulation* 20.3 (2013), pp. 965–973. DOI: 10.1109/TDEI.2013.6518966.

## Calendar Week: 51 18 December, 2020

### Completed Tasks

1. Have written history of Predictive maintenance. Some improvements are still necessary.
2. Finalized the plan for the upcoming weeks.

### Tasks in-progress

1. Complete the health index formal definition chapter.
2. Write the first Approach for calculating HI.

## Calendar Week: 50 11 December, 2020

### Completed Tasks

1. Done with reading of the seed papers [1], [2], and [3] and geatherd material that explain different machine learning methods for estimating health index, which will be studied in detail in the upcoming weeks.

2. Completed the partition of the work on the heath index section with Gourav Prakash.

### Challenges

1. A small challenge has been finding papers that focus soley on health index calculations. A lot of papers have health index as an itermedeiate result needed for RUL calculation.

### Tasks in-progress

1. Write down the plan for the upcoming seven weeks.

2. Start with the first two methods for calculating health index, Subsystem-Based GRNN [1] and Neural Networks [2].

3. Write the history of predictive maintenance part for the Introduction chapter.

## References

[1] M. Islam et al. "Calculating a Health Index for Power Transformers Using a Subsystem-Based GRNN Approach". In: *IEEE Transactions on Power Delivery* 33.4 (2018), pp. 1903–1912. DOI: 10.1109/TPWRD.2017.2770166.

[2] F. Yang et al. "Health Index-Based Prognostics for Remaining Useful Life Predictions in Electrical Machines". In: *IEEE Transactions on Industrial Electronics* 63.4 (2016), pp. 2633–2644. DOI: 10.1109/TIE.2016.2515054.

[3] Pankaj Malhotra et al. *Multi-Sensor Prognostics using an Unsupervised Health Index based on LSTM Encoder-Decoder*. 2016. arXiv: 1608.06154 [cs.LG].

## Calendar Week: 49 04 December, 2020

### Completed Tasks

1. Participated in the thursday meeting where we decided the primary structure of the survey document.
2. Had an initial meeting with Gourav Prakash where we discussed our intial thougts for a plan for milestone 2.

### Challenges

1. Deciding on the pipleine elements that will be included.

### Tasks in-progress

1. Seperation of the work for the subtopic.
2. Give a short description for each section of Health Index estimation subtopic.
3. Read the seed papers [1], [2], and [3] and start gathering literature for other methods.

## References

[1] M. Islam et al. "Calculating a Health Index for Power Transformers Using a Subsystem-Based GRNN Approach". In: *IEEE Transactions on Power Delivery* 33.4 (2018), pp. 1903–1912. DOI: 10.1109/TPWRD.2017.2770166.

[2] F. Yang et al. "Health Index-Based Prognostics for Remaining Useful Life Predictions in Electrical Machines". In: *IEEE Transactions on Industrial Electronics* 63.4 (2016), pp. 2633–2644. DOI: 10.1109/TIE.2016.2515054.

[3] Pankaj Malhotra et al. *Multi-Sensor Prognostics using an Unsupervised Health Index based on LSTM Encoder-Decoder*. 2016. arXiv: 1608.06154 [cs.LG].

## Calendar Week: 48 27 November, 2020

### Completed Tasks

1. In this week I looked up some extra information on how feature extraction can be achieved. More precisely I looked at *Vibration Feature Extraction Techniques for Fault Diagnosis of Rotating Machinery - A Literature Survey*. The focus of this survey were techniques for feature extraction from vibration data in rotating machinery. Techniques for extraction in time domain, frequency domain and time-frequency domain were presented. Also, their performance was evaluated. [1]

### Challenges

1. If the subtopic I will be working on is feature extraction, then a few of the techniques presented in [1] will need a bit more attention and understanding.

## References

[1] Saleem Riaz et al. "Vibration Feature Extraction and Analysis for Fault Diagnosis of Rotating Machinery-A Literature Survey". In: *Asia Pacific Journal of Multidisciplinary Research* 5 (Mar. 2017), pp. 103–110.

## Calendar Week: 47 20 November, 2020

### Completed Tasks

1. Read *Development and Performance Evaluation of Prognostic Approaches for Technical Systems* by James Kuria Kimotho up to chapter 4. The work covers in a detailed manner the prognostic approaches. Initially an overview of diagnostic and prognostic approaches is given (chapter 1 and 2). Then, methodologies for preprocessing of the data, feature extraction, and feature selection are presented (chapter 3). Lastly, methods for identifying the health states, estimating health index and calculating RUL are given (chapter 4). Also, applications of these methodologies and their results are presented. [1]

### Tasks in-progress

1. Research focused on feature extraction and health detection/health classification as a preparation for the subtopic selection.

## References

[1] J.K. Kimotho. "Development and performance evaluation of prognostic approaches for technical systems". In: *Ph.D. dissertation* (2016).

## Calendar Week: 46 13 November, 2020

### Completed Tasks

1. Read the survey *A Survey of Predictive Maintenance: Systems, Purposes and Approaches*. This survey is focused on system architectures, purposes and approaches used for developing a PdM system. Firstly, three types of architectures are described: Open System Architecture for Condition Based Monitoring (OSA-CBM), Cloud enhanced PdM systems and PdM 4.0. Then, cost minimization and reliability/availability maximization and are given as the purpose of these systems. Lastly, PdM approaches are categorized in three categories: knowledge based, machine learning (ML) based and deep learning (DL) based. Then an overview of these methods is given, with more focus on the DL methods. [1]

2. The survey *A systematic literature review of machine learning methods applied to predictive maintenance* gives an overview of the latest research being done in PdM. Also, there is given a summary of the journals and conferences where these papers are published. Also, a summary of the most recent ML methods that are used to achieve PdM is given. Furthermore, public data sets for use in predictive maintenance are presented. [2]

### Challenges

1. Understanding some of the more complicated DL methods.

2. Narrow down the choice of the subtopic.

### Tasks in-progress

1. Reading more about some of the mentioned DL methods to get a basic understanding.

2. Explore more on how feature extraction is realized.

## References

[1] Yongyi Ran et al. "A Survey of Predictive Maintenance: Systems, Purposes and Approaches". In: *CoRR* abs/1912.07383 (2019). arXiv: 1912.07383. URL: http://arxiv.org/abs/1912.07383.

[2] Thyago Peres Carvalho et al. "A systematic literature review of machine learning methods applied to predictive maintenance". In: *Comput. Ind. Eng.* 137 (2019). DOI: 10.1016/j.cie.2019.106024. URL: https://doi.org/10.1016/j.cie.2019.106024.

## Calendar Week: 45 06 November, 2020

### Completed Tasks

1. Followed and supported the team, on a lengthy meeting, with the setting up of the working tools we are going to use.

2. Read the survey titled *Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A survey.* Firstly, the survey presented an overview on maintenance methods. Then the Data-Driven methods are presented as the most important methods for achieving effective maintenance. The increase of importance of these methods comes as a result of many big industrial projects that are happening around the world, ex. Industry 4.0 in Germany. I also learned the structure of a data-driven PdM system and its challenges for implementation. Lastly the paper presented Machine Learning and Deep Learning methods that are used in various research papers for implementing such a system. An analysis of these methods is given in terms of accuracy and the challenges faced. [1]

### Challenges

1. Get used to the workflow of the project and improve the knowledge on the tools we are going to use.

2. Help the team further improve our work environment.

### Tasks in-progress

1. Read the other surveys.

2. Research on how are some of the PdM problems are handled in practice.

## References

[1] Weiting Zhang, Dong Yang, and Hongchao Wang. "Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey". In: *IEEE Syst. J.* 13.3 (2019), pp. 2213–2227. DOI: 10.1109/JSYST.2019.2905565. URL: https://doi.org/10.1109/JSYST.2019.2905565.