Data-driven predictive maintenance and time-series applications

Kefalas, M.

Chapter 9

Conclusions and Outlook

In this chapter, we conclude the work done in this thesis, by presenting the research questions and their answers and discussing future research opportunities.

9.1 Conclusions

"Prediction is very difficult, especially if it's about the future!" We started this dissertation with these words from Nobel laureate in physics Niels Bohr, and we are ending it with the same quote. However, we state that, with this thesis, prediction is now a little bit easier.
This thesis discussed the importance of predictive maintenance (PdM) and the significance of prognostics and health management (PHM) in industry and society. We, specifically, investigated AI-based techniques for data-driven PdM and the remaining useful life (RUL) estimation - a key concept in this thesis - presented time-series applications, and discussed future directions. Next, we will refer back to our research questions and present their answers.

RQ1: What are the advancements, drawbacks, and opportunities in PHM and specifically of data-driven PdM in the aerospace industry? This research question was covered in **Chapter** 3. There has been significant work done in developing data-driven techniques for prognostics in the past decade. Our research showed that methods other than traditional time-series analysis are gaining popularity in the data-driven prognostics in aerospace, such as neural networks (NNs). Data-driven methods are emerging as they propel data-driven solutions by not requiring (a lot of) engineering knowledge making them, thus, available to a broader audience due to their domain-agnostic nature. Despite, however, the recent overall data-driven success in prognostics, in general, but also in aerospace there are still challenges and practical issues that need to be addressed, such as securely obtaining more data, real-time prognostics and prognostics on the edge (edge computing), explainability and interpretability of the methods and uncertainty management of the prognostics developed.

RQ2: Can automated machine learning (AutoML) methods be applied for the estimation of the RUL in data-driven PdM? This research question was covered in **Chapter** 4. The gathered results show that AutoML using classic machine learning algorithms can be effectively applied for the estimation of the RUL after the data has been appropriately pre-processed. We addressed this research question by estimating the RUL of the test instances from the widely used C-MAPSS dataset [198] using a specific AutoML method (TPOT [169]) and comparing the performance to that of state-of-the-art methods. Since the learning algorithms in the AutoML search space are classic machine learning methods, the time-series data of C-MAPSS need to be pre-processed and transformed into a regression problem. The data is pre-processed by extracting statistical features from expanding windows of the time-series to uncover the degradation that has been accumulating from the system's early life or after an overhaul. The experimental results indicate that such an approach can outperform or achieve comparable results compared to classic machine learning techniques for the estimation of the RUL. However, when compared to deep learning (DL) approaches, the performance is comparable or lower, suggesting that classic ML might not be able to uncover the highly non-linear relationship between the RUL and the observed data.

RQ3: Can we propose an automated framework for configuring RUL prediction models which are highly accurate and have less estimation uncertainty? This research question was covered in **Chapter** 5. The work of this chapter shows that this is indeed possible. In detail, we addressed this research question by developing an end-to-end pipeline for the RUL estimation using a recurrent neural network and task-specific pre-processing that are both optimized through a *bi-objective* hyperparameter optimization (HPO) method that *jointly* minimizes the pointwise RMSE and the uncertainty. This pipeline is an automated framework as it takes as input the raw data and returns RUL estimation models with low prediction error and prediction uncertainty. The method was validated on two subsets from the widely used C-MAPSS dataset [198] and was compared against a single-objective HPO variant.

RQ4: Can explainable AI facilitate the understanding of the data generating process of industrial processes? This research question was covered in **Chapter** 6. The research shows that explainable AI can add to the understanding of the data generating process and lend itself to PdM. We addressed this research question with the use of symbolic regression (SR) by means of genetic programming (GP) on real aircraft operational data with the aim of uncovering meaningful relationships between the exhaust gas temperature (EGT) - a standard industrial indicator of the health of an aircraft engine - and the rest of the monitored engine parameters, in the form of mathematical formulas. The experimental results show, apart from good model accuracy, explainability, and consistency from a physics/engineering perspective, which field experts validated. This indicated that the proposed method could uncover meaningful relationships in the data that the end-user can interpret. The resulting formulas, in turn, can assist in PdM by, e.g., asset monitoring through comparing the predicted values from the formulas to that of the observed measurements, as well as being used to create simulated data that can be used towards the development of (data-driven) PdM solutions.

RQ5: Can tabu search (TS) support the solution of the multi-objective flexible job-shop scheduling problem (FJSSP)? This research question was covered in **Chapter** 7. The research we performed shows that TS can assist in the solution of the multi-objective FJSSP. In detail, we addressed this research question by developing a memetic algorithm that uses TS as the local search method and a mutation operator to solve the multi-objective FJSSP. The three objectives to be minimized are the makespan, the total workload, and the maximum workload. The method is compared against the state-of-the-art algorithms by Yuan et al. [251] on the widely used Brandimarte datasets [32]. The experimental results show that the proposed method can dominate solutions from the baselines, identify new Pareto front solutions, and approximate the extended reference set created from the baseline and our solutions.

RQ6: Can time-series techniques from industry lend themselves to applications in the medical domain? This research question was covered in **Chapter** 8. The gathered results show that this is indeed possible. We addressed this research question using a case study from the field of Neurology, with the task being to distinguish between normal and abnormal electromyography (EMG) recordings from patients with neuromuscular disease. We attended to this by developing further a time-series classification pipeline, originally designed for the automotive industry, that extracts and learns statistical features from the EMG recordings to classify muscles and patients as either healthy or not. The experimental results show relatively high performance, indicating that the approach might uncover information from the EMG recordings that is not used in routine clinical assessment, which, in turn, could lead to the identification of new EMG biomarkers. Furthermore, the method can assist clinicians in diagnosing if a patient has a neuromuscular disease.

9.2 Outlook

Although considerable progress has been made in the context of this thesis, there is still work to be done in the field of data-driven prognostics. In the following, we briefly discuss recommended future work based on the work of this thesis and we will close this chapter with some general outlook and recommendations for data-driven prognostics.

Neural Architecture Search (NAS) and RUL In **Chapter** 4 we used AutoML with classic machine learning algorithms to estimate the RUL. There, the experimental results indicate that such an approach might not be able to uncover the highly non-linear relationship between the RUL and the observed data, as well as deep learning methods. Therefore, it would be beneficial for future work to include neural network architectures as well, by means of NAS, a challenging topic of AutoML [97]. NAS might be able to uncover architectures that can improve the state-of-the-art deep learning methods in prognostics while reducing the manual labor of tuning them.

Effective Time-Series Representations In **Chapter** 4 and in **Chapter** 8 we used methods that extract and learn from statistical features for a regression and a classification task, respectively. Even though this is a successful approach, it would be favorable to research effective time-series representations that are able to retain the necessary degradation information (e.g., for RUL estimation) and/or can capture the subtle patterns that allow for differential diagnosis of disease (e.g., in fields such as Neurology). Furthermore, these representations can reveal potential biomarkers for medical applications, as we saw in the work done in **Chapter** 8. In this view, we highly recommend further research on the discriminatory features discussed in the work of **Chapter** 8, as they could be of clinical significance in electrodiagnostic medicine.

RUL-target Label Creation In **Chapter** 4 and in **Chapter** 5 we used methods from literature to create the RUL-target labels in order to tackle the RUL estimation problem as a supervised regression problem. These methods include a linear and a piece-wise linear model to map to each time-step of the training data a real number that represents the RUL at that specific time-step. The decision regarding which of the two methods to use stems from the nature of the degradation. This means that for systems that can sustain high material stress, we will use the piece-wise linear method, whereas, for systems where degradation is more or less evident immediately, we will use the linear method. We encourage further research towards alternative RUL-target label creation models (e.g., quadratic curves) as this will broaden the tools researchers and end-users have to effectively pre-process the data by accurately representing different degradation profiles.

Multi-objective Hyperparameter Optimization (HPO) In **Chapter** 5 we designed an end-to-end pipeline for RUL estimation using a bi-objective HPO that jointly minimizes the prediction error and uncertainty. The reasoning behind this is that there can be conflicts between these two objectives in specific tasks. Future work should invest in multi-objective (or many-objective in the case of more than 2 objectives) HPO, as it is not the case that a single objective method can always capture the conflicting interests that exist in real-world problems.

Uncertainty Quantification (UQ) in Data-driven Prognostics A critical research topic that has received some attention in the DL community in recent years is that of uncertainty

quantification (UQ). Uncertainties arise from various sources, such as modeling and input data uncertainties. Quantifying uncertainty is crucial for any prognostic estimate, as discussed in **Chapter** 5, otherwise, it is of limited use and cannot be incorporated into safety-critical or operations-critical applications. By accounting for the uncertainties, the researcher or end-user can determine if, for example, the training data is not representative of the task or are too noisy (i.e., measurement uncertainties, operating environment uncertainties, future load uncertainties, input data uncertainties) or if the selected model is poorly selected (i.e., under-parameterized NN). Especially when it comes to NNs, UQ is a rising topic of interest, given that NNs are being industrially employed. Although, in recent years, there has been work done on UQ in DL and some on UQ in data-driven PdM, the field is still young, with no consensus on how to measure and use uncertainty in data-driven PdM. Furthermore, while the majority of model-based prognostic methods quantify the associated uncertainty through, for example, modeling the process and observation noise (e.g., Kalman filters), only a few studies in the data-driven domain address this matter, despite its importance [26]. Therefore, we recommend this as an essential and exciting direction for further research in data-driven PdM.
Next, we will briefly outline future research that will generally be useful for PHM and data-driven PdM.

General Recommendations for Future Research in Data-driven PdM As discussed in **Chapter** 3, the field of PHM requires methods, systems, and protocols to obtain more data securely. This means that future work should emphasize not only algorithmic performance but also data quality and the drawing up of certain conventions per industrial field that govern data quality. In this view, as suggested already in **Chapter** 3, there are opportunities in directions such as federated learning (FL) [6] which allows for data augmentation, as data is gathered from various parties securely, as well as data protection by training on the data of each data holder involved before aggregating and updating the parameters of the *global* model.
Additionally, there are opportunities for research in real-time prognostics for field applications. One promising direction to this, as suggested already in **Chapter** 3, is edge computing which allows for low latency since the data are processed closer to their source, thus, allowing accelerated insights compared to traditional offline methods. A natural extension to the previous is how we can perform data-intensive calculations on edge, where the hardware conditions are not as robust and computationally powerful as in the cloud or other data centers. For example, this would be extremely beneficial for vessels that operate in open seas for weeks at a time, without the ease of transmitting their massive operational data and that need real-time prognostics.
Furthermore, as we saw and discussed in the context **Chapter** 3 and **Chapter** 6, PdM methods have evolved to include neural networks (NNs) and similar deep learning (DL) tools in their arsenal. As we know, however, these tools are considered “black box”. This means that these models do not explain their predictions/outputs in a way that humans understand. As a result,

this lack of transparency and accountability can have severe consequences [191], especially in operations-critical or safety-critical systems, like in aerospace. Therefore, future work must emphasize interpretability of the results and explainability of the model workings to assist decision-makers in asserting the feasibility of the model logic, as well as in the troubleshooting of the developed methods, thus putting confidence in the overall prognostics process.

Finally, even though data-driven methods do not require a lot of domain knowledge, if that knowledge exists, it is beneficial to incorporate it into the learning process. The knowledge can be incorporated, for example, in the form of a specific loss function or a specific architecture. Knowledge-infused learning, as the process is called, is a promising future direction for research in data-driven PdM as it constrains the available options in the context of the specific problem and adds domain information that can better assist the learning process.