Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 130 (2024) 1537–1542

www.elsevier.com/locate/procedia

57th CIRP Conference on Manufacturing Systems 2024 (CMS 2024)

A Reference Model for Predictive Maintenance Model Development

Sielaff, Lennard[a]; Lucke, Dominik[a,b,*]; Wolf, Yannic[c]

[a] *Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany*
[b] *Hochschule Reutlingen, ESB Business School, Alteburgstraße 150, 72762 Reutlingen, Germany*
[c] *Chair of Transportation Systems Engineering, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany*

* Corresponding author. Tel.: +49 711 970-1897. *E-mail address:* dominik.lucke@ipa.fraunhofer.de

Abstract

Deterioration modeling plays a pivotal role in various industries, enabling predictive maintenance strategies and cost-effective resource allocation. Furthermore, with an escalating influx of data across multiple domains, opportunities for predictive analysis continue to expand. The development of wear models becomes a more and more complex process especially when acquiring and handling large amounts of customer data of different products and stakeholders since additional topics such as data privacy have to be included. Therefore, the development process causes high efforts in time, capacity and coordination between the different disciplines such as domain experts, data scientists and IT-administration. This paper presents a reference model for predictive maintenance model development. Based on adopted best practice approaches from the industrial production context, the reference model is structured around the four phases of the CRISP-DM model. Altogether it encompasses 48 defined steps. Covering the whole life cycle, beginning with component identification, use case description and culminating in model deployment and maintenance, each step is meticulously crafted to ensure quality and speed up the predictive maintenance model development. By adhering to this systematic approach, wear models for components can be developed with confidence, with lower effort and costs and mitigating uncertainties. The reference model is validated in the automotive industry since there already exist large amounts of fleet data of different products and customer. By providing a reliable and systematic approach to predictive maintenance model development, this reference model empowers stakeholders to optimize maintenance schedules, reduce downtime, and enhance overall operational efficiency.

Keywords: smart factories; predictive maintenance; big data analytics

1. Introduction

Predictive maintenance has gained increasing importance in recent years due to several factors. Firstly, the availability of vast amounts of data has made it possible to monitor equipment and identify potential failures before they occur. This shift from reactive to proactive maintenance allows companies to minimize downtime, optimize maintenance schedules, and reduce costs [1]. Secondly, predictive maintenance enables a more efficient use of components and resources. By predicting when a machine or component is likely to fail, companies can plan maintenance activities accordingly, ensuring that maintenance is performed only when required. This reduces the risk of unnecessary maintenance and extends the lifespan of equipment, leading to cost savings. Furthermore, predictive maintenance helps to minimize unexpected production stops. By detecting and addressing potential issues before they cause a breakdown, companies can avoid unplanned downtime and maintain a smooth production flow. This improves overall productivity and – from the view of an OEM – customer satisfaction [2]. However, developing predictive maintenance solutions can be challenging. It requires collaboration among various disciplines such as engineering, production, data scientists, product development, maintenance, and software engineers. Each discipline has its own perspective and requirements, making it difficult to align everyone's goals and

10.1016/j.procir.2024.10.279

expectations [3]. Moreover, there is often a lack of experience within a single company when it comes to developing predictive maintenance solutions. This lack of experience can lead to inefficiencies, delays, and suboptimal solutions. Companies may struggle to navigate the complexities of implementing predictive maintenance and may face difficulties in integrating various technologies and systems [4]. Also, often during development of predictive maintenance solutions same or similar components are developed once again, instead to use and adapt existing ones [5]. To summarize, in the development processes of predictive maintenance solutions there are still improvement potentials. To address these challenges, a reference model is essential. A reference model provides a standardized framework and guidelines for developing predictive maintenance solutions. It helps align different disciplines, streamlines the development process, and ensures that key considerations are not overlooked. By following a reference model, companies can benefit from the collective knowledge and best practices of the industry, accelerating the development and implementation of predictive maintenance solutions.

2. Related work

The CRISP-DM (Cross-Industry Standard Process for Data Mining) model [6], the ASUM-DM (Analytics Solutions Unified Method for Data Mining/Predictive Analytics) model [7,8] and Team Data Science Process (TDSP) [9,10] are domain-agnostic, proven process models. CRISP-DM divides the data mining process into the business understanding, data understanding, data preparation, modelling, evaluation, deployment phases [6]. Developed more than 20 years ago, it is still often in use in data mining projects. The ASUM model, developed by IBM several years ago, consists of the phases analyse, design, configure and build, deploy, operate and optimize [7,8]. The TSDP model, developed by Microsoft, comprises the phases business understanding, data acquisition and understanding, modelling and deployment [9–11]. In contrast to CRISP-DM, ASUM and TDSP already contain information related to the project organization and management. In order to better meet specific requirements of domains, several specific models have been developed. Examples for the engineering and manufacturing area are the Data Mining Methodology for Engineering Applications (DMME) [12] proposing an enhanced CRISP-DM or the methods of Kozjek, et al. [13]. The proposed framework of Kozjek, et al. includes technical aspects and required knowledge and competences in the different phases of a data analytics project [13]. The model of Welte et al. focus in the process model on applicability in SME [14]. For simulation models in material handling, logistics and production, the VDI guideline 3633 specifies a general process model at a detailed level [15]. The guideline distinguishes the phases of a typical simulation project into the simulation task definition, system analysis, data collection, model formalisation and implementation and the experiments [15]. Grambau et al. [16] proposes a reference architecture for predictive maintenance product service systems.

3. The reference model for predictive maintenance model development

The reference model for predictive maintenance (PredM) model development consists of two sub models: the architectural model and the function and process model. Moreover, in contrast to other models, also a standardized use case description is included. The use case description presented in [17] accompanies predictive maintenance model over the whole life cycle. This enables a much easier reuse and adaption of components in other models. The architectural model contains the different required functional blocks and data storages and the relations. The process model describes the required steps for the whole development and operation process of a predictive maintenance model.

3.1. Architectural model

The architecture model (Fig.1) developed includes a description of how and with the help of which systems the data that can be used to develop and implement a predictive maintenance model can be collected and processed. Starting with data acquisition on the asset, through data pre-processing, the use of a central data warehouse and development environments, to the use of the models for prediction. The architecture model contains both the data flows and event occurrences with which various systems are either synchronised or trigger events in other systems. The underlying processes are mapped in the respective process models and are based on the architecture shown here. In order to make the architecture model clearer, it has been divided into five areas, each of which is connected to each other via data flows or event occurrences. The five areas are: Data Collection, Data Management, Predictive Maintenance Development Environment, Quality and Visualisation and, as a final area, Model Execution. In addition, the areas of data collection and part of data management are subdivided into "test asset" and "fleet asset", as the possible data flow for these two options differs. In the model, the corresponding systems are highlighted in colour. The five areas are briefly presented in the following sections.

3.1.1. Data collection

Selected sensor values are recorded in the fleet asset using the existing sensors and the control unit already installed in the asset and transmitted to the data management system. In test assets, there is also the option of installing additional sensors that can also be analysed. This additional sensor data must be synchronised with the on-board sensors to enable joint data processing. The additional recorded values can be pre-processed and visualised in the asset before the data is transmitted to the data management system.

3.1.2. Data management

The recorded data is pre-processed and, if it comes from customer assets, pseudonymised. At this point, e.g., limit values for event detection from modelling can be specified. The data recorded in this way and any labelled data are transferred to different databases after pre-processing. A database for the test

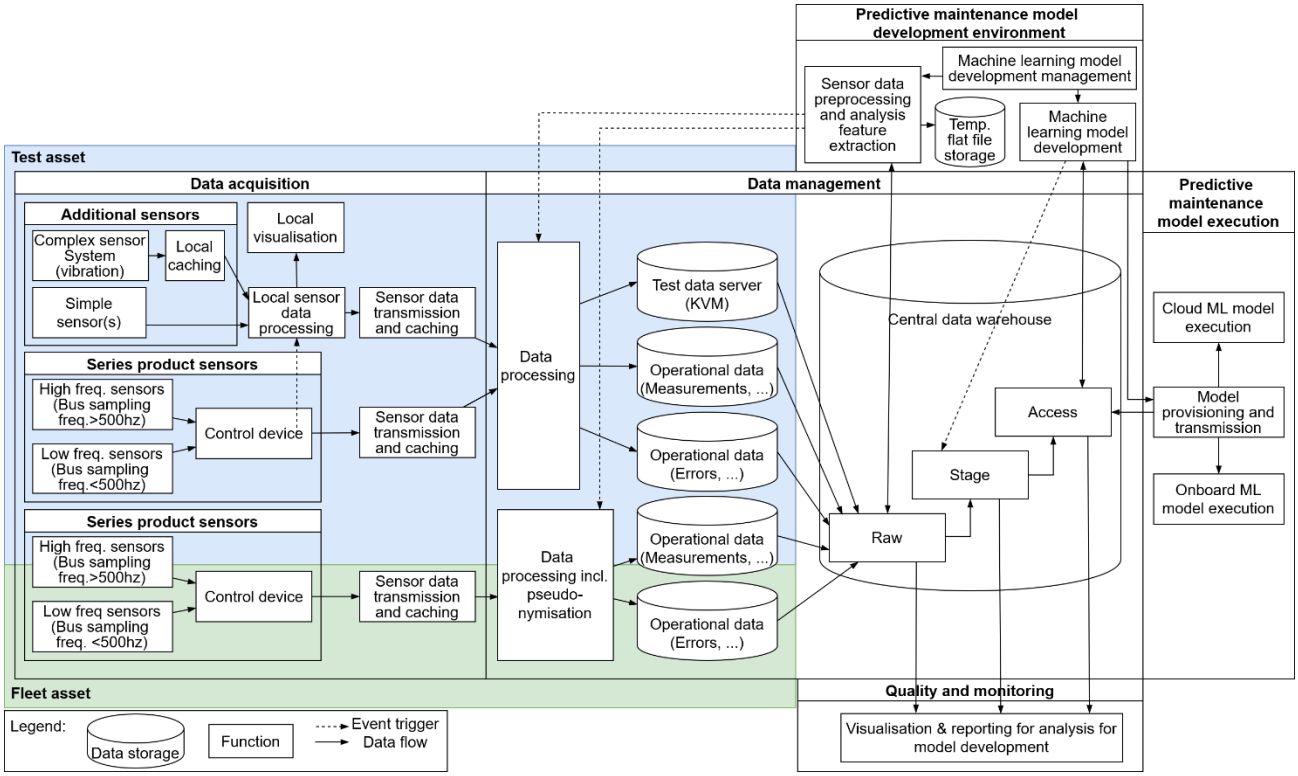

Fig. 1. Architectural model to develop a predictive maintenance model

data and the operational data of the assets (test assets and fleet assets) is divided into sensor values from the assets and management data that has already been pre-processed in the assets (e.g. error coding). This subdivision into different databases is justified both technically and, in part, historically. The data and information collected in this way is transferred to a common central data warehouse, which is classically divided into three levels: Raw, Stage and Access. At the raw level, the raw data is collected in a simplified manner and then cleansed and linked at the stage level. The access level is generally used to make the data available to other functions.

3.1.3. Predictive maintenance development environment

The raw data is first prepared in the development environment. Either raw data from the data warehouse is used, enriched and fed back, or rules are defined for handling the raw data, which are then implemented in other systems. The data from the access area of the data warehouse is used to refine the machine learning models and carry out model development. Both the rule creation and the handling of the raw data, as well as the model development, are typically supervised by a development management team.

3.1.4. Quality and monitoring

These functions are based on the central data warehouse and make it possible to visualise the data and check the quality. For this purpose, the functions have interfaces to all levels of the warehouse in order to check the raw data, the processed data and the results. Not only the data, but also the model results are checked and evaluated.

3.1.5. Predictive maintenance model execution

The model execution utilises the predictive maintenance models developed and either in the cloud or directly in the assets. To do this, the respective environment must first be set up accordingly. If necessary, the results of the model execution are also transferred back to the data warehouse to enable further developments.

3.2. Process model

As for the architecture model, an idealised predictive maintenance model development was first run through with all stakeholders for the process model. Using the CRIP-DM process as a template, the development was divided into four phases (development and validation phases are merged). Here, too, care was taken to ensure that the processes described in the reference model are generally valid in order to describe the implementation for predictive maintenance modelling in general. Stakeholder specific processes were either included in generalisations or incorporated into the processes using an "OR" relationship. For the phases: Business Understanding, Data Preparation, Modelling and Evaluation and Deployment phases each have their own process diagram, most of which can be displayed as a self-contained process. There are only a few connections between the phases in the process, which are clearly labelled as such. The Business Process Modelling Notation (BPMN) was chosen for the visualisation and the preferred direction is from top to bottom. Possible loops are included and can be recognised by the corresponding arrows. The process starts in an empty circle at the top and runs according to the arrows via "X (OR)" or "Cross (AND)" connections to the end of the process, the empty circle at the bottom. Dashed arrows describe information that is required for the respective process step. The model for the first phase “Business Understanding” is already published in [17]. It consists of first analysing the asset’s components and evaluating for which component a predictive maintenance

model is needed. Once the relevant components have been identified, the next step is to discuss the objectives of the wear and the prediction model, as well as the technical and organizational skills required to implement the predictive maintenance model. The expected algorithms that might be used for the model and the data likely to be required are also identified in this phase. All the information collected this way is entered into a standardized overview, which is updated throughout the process. The first process step presented in this paper, "Data Preparation", begins by reading this use case description in order to provide all participants with the required information. [17]

3.2.1. Data Preparation

At the beginning of the next phase, data preparation and data pre-processing (Fig. 2), the use case description is used as a basis. In addition to the system to be analysed, this also describes the objectives and gives all participants an overview of what data needs to be collected and how it should be prepared. If further data needs to be collected for the implementation of the use case, it is first checked whether other data sources need to be integrated in addition to the control units in the assets. If this is not the case, a distinction is made as to whether the data is from series or fleet assets, as data from series assets must first be pseudonymised. Once all data sources have been recorded, it is necessary to check whether the data still needs to be processed. This is usually the case unless the data has already been prepared for a previous analysis. The processed data is then evaluated by technical experts before it is stored in the central data warehouse. Once the various data sources are all stored in the data warehouse, they are synchronised and checked for their overall quality and data distribution. If the data quality does not meet expectations, the data acquisition process steps must be repeated. If the quality is sufficient, the synchronised data is filtered over time in order to extract the relevant data sections. These are checked again for correctness and data quality in order to obtain the complete and cleansed data set. The final process stage in the data pre-processing area deals with data labelling. For this, the special places within the data set must be recognised and marked as such. If this is not possible, the process steps for data collection must be run through again. The process ends with a complete and cleaned data set with labelled events on the basis of which a predictive maintenance model can be created with the help of machine learning.

3.2.2. Modelling and Evaluation

At the beginning of the modelling and evaluation phase (Fig. 3), the use case description is also used first to give everyone involved an overview of the system, the objectives and the work done so far for the use case. Once everyone involved has been informed, the cleansed and labelled data set is divided into training and validation data. If a suitable model for data processing and prediction already exists, the training data can be used directly to train the existing model. If no model is yet known that can be used for the predictive maintenance approach, a model must first be developed. To do this, different approaches and algorithms for creating a model are tested using part of the training data set. The results of the different

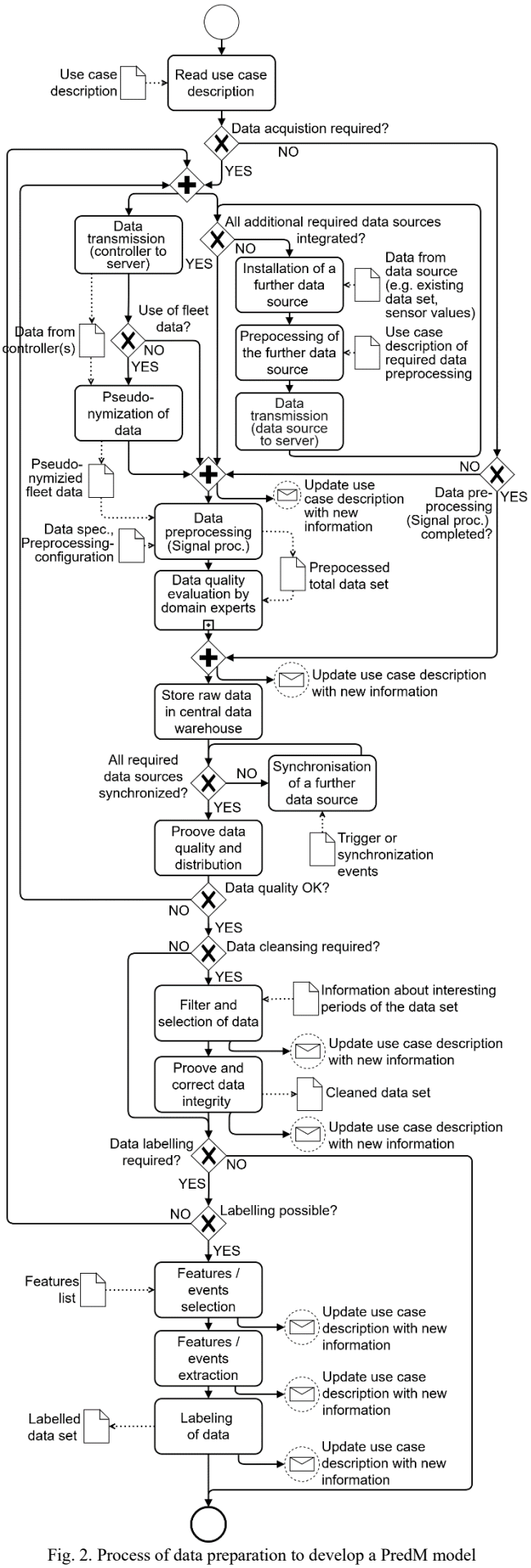

Fig. 2. Process of data preparation to develop a PredM model

approaches are evaluated on the basis of previously defined criteria in order to decide whether one of the existing approaches is good enough for the model development or whether a new approach or algorithm needs to be developed. The resulting model, either based on existing approaches or on a new development, must then be trained with the complete training data set. At this point, the processes merge again. Known models, known approaches or even completely new algorithms are trained on the basis of the entire training data set. The trained models are validated using the previously separated validation data and evaluated by experts. If the model results are found to be good, the use case description is updated, and the model can be used. If the results do not meet expectations, other approaches or algorithms for modelling must be tested.

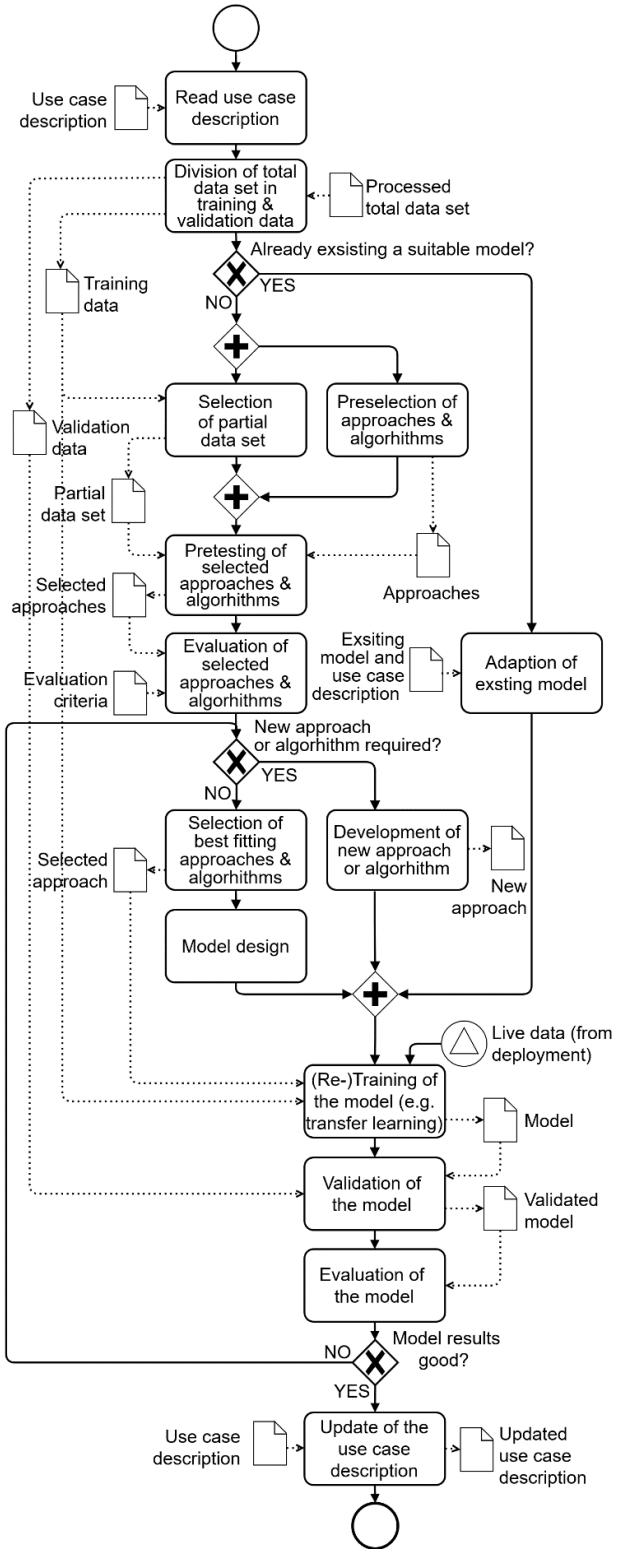

Fig. 3. Process of modelling and evaluation to develop a PredM model

3.2.3. Deployment

The distribution and application of predictive maintenance models differs directly in the first step between a centralised, cloud-based application and that in the individual assets (Fig. 4). In both cases, the models must first be loaded into the application environment and configured for the specific model, but the implementation differs greatly between a centralised and many local installations. Once the model is installed in the respective runtime environment and it is configured, the model is executed. Based on the model and the objectives defined in the Business Understanding process section, the model results are generated and made available. This can also include the import of model results back into the central data warehouse, which can then be used to adjust future model parameters. In any case, model execution and model behaviour are monitored during the application.

4. Validation of the Reference Model

In a collaborative research project, several companies with different areas of expertise came together to develop and

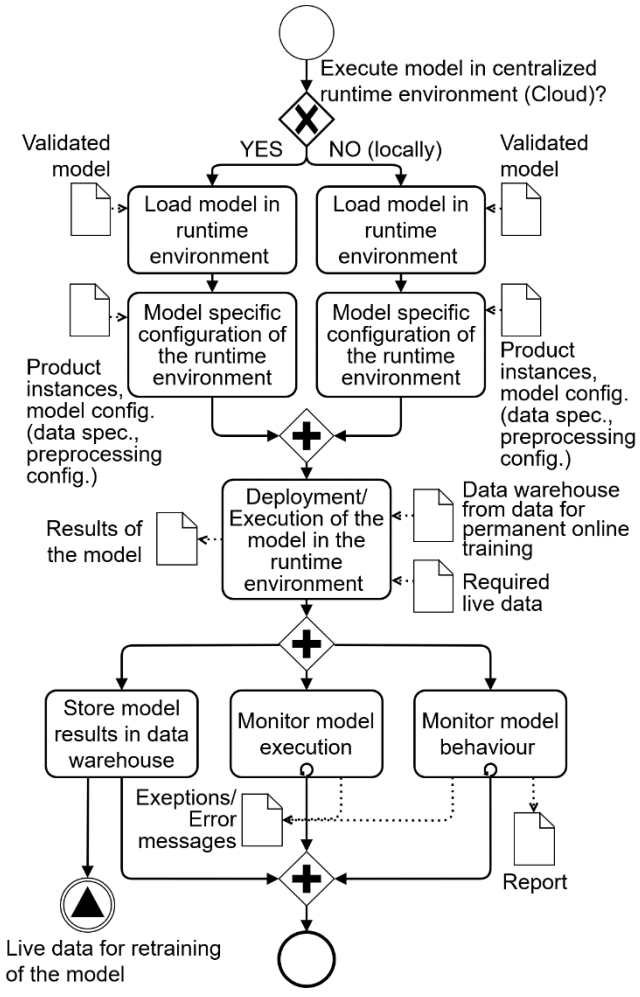

Fig. 4. Process of deployment to develop a PredM model

validate this reference model for developing predictive maintenance models. Throughout the project, experts from four companies and one research institute used and evaluated the reference model at different stages and with different use cases. By using the reference model to develop unique components and individual goals for the specific use cases, the developed reference model was carefully tested to understand if it could handle this wide range of applications and accommodate different developer perspectives. For validation, participants were surveyed about their experience working with the reference model compared to working without the model, and about the overall effectiveness of the reference model. The results of the survey show that the practical implementation of the reference model demonstrated its ability to increase development speed and streamline communication and coordination efforts. The use of the reference model greatly improved collaboration between individual companies and developers by establishing clear guidelines and defining expected outcomes for each work step. This facilitated smoother coordination between steps and highlighted areas where similar processes and applications could be leveraged across different use cases. As a result, previous work could often be reused, resulting in more efficient workflows. These findings underscore the significant benefits that the use of the reference model can bring, fostering improved outcomes both at the intercompany level and within individual organizations.

5. Conclusion

The presented reference model for the development of predictive maintenance models comprises four phases. The first phase, Business Understanding, has been described in a previous publication [17]. Building on that, this paper presents the overall (software-) architecture model and the processes in the Data Preparation, Modelling and Evaluation, and Deployment phases. The advantages of using this reference model were shown in a validation process involving different stakeholder and various components for which predictive maintenance models were developed. Adherence to this reference model during the development of predictive maintenance models can reduce processing time and effort, as well as communication and coordination overhead. In addition, the standardised approach allows certain processing steps to be reused across different models, thereby reducing redundant work. The validation with different components has also shown that the process in the Deployment phase is highly component- and use-case specific, making it difficult for the reference model to provide detailed predefined steps. This highlights the need for further research to better standardise this phase. There is also a need for research to extend the application of the reference model to other industries. While the reference model was developed for predictive maintenance models in the automotive context, it can be applied to other industries. However, it is necessary to investigate whether the model, individual processes or the involvement of other stakeholders are required in such cases.

Acknowledgements

The authors thank all the partners of the project who contribute to this research work. The PANAMERA project is funded by the German Federal Ministry for Economic Affairs and Climate Action in the program „Neue Fahrzeug- und Systemtechnologien“ (19I21016).

References

[1] Achouch M, Dimitrova M, Ziane K, Sattarpanah Karganroudi S, Dhouib R, Ibrahim H, Adda M. On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges. Applied Sciences; 2022;12(16):8081.
[2] Prasetyawan Y, Rachmayanti I. Proposing predictive maintenance strategy to increase OEE through system upgrade scenarios and AHP. IOP Conf. Ser.: Mater. Sci. Eng; 2021;1072(1):12031.
[3] Theissler A, Pérez-Velázquez J, Kettelgerdes M, Elger G. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. Reliability Engineering & System Safety; 2021;215:107864.
[4] Dalzochio J, Kunst R, Pignaton E, Binotto A, Sanyal S, Favilla J, Barbosa J. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. Computers in Industry; 2020;123:103298.
[5] Serradilla O, Zugasti E, Rodriguez J, Zurutuza U. Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. Appl Intell; 2022;52(10):10934–64.
[6] Wirth R, Hipp J. CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining; 2000.
[7] IBM Corporation. Analytics Solutions Unified Method - IBM. https://public.dhe.ibm.com/software/data/sw-library/services/ASUM.pdf. Accessed on 11.04.2024.
[8] Ponsard C, Touzani M, Majchrowski A. Combining Process Guidance and Industrial Feedback for Successfully Deploying Big Data Projects. OJBD; 2017;3:26–41.
[9] Mark Tabladillo. What is the Team Data Science Process? https://learn.microsoft.com/en-us/azure/architecture/data-science-process/overview?source=recommendations. Accessed on 11.04.2024.
[10] Haakman M, Cruz L, Huijgens H, van Deursen A. AI lifecycle models need to be revised. Empir Software Eng; 2021;26(5).
[11] Díaz Álvarez Y, Hidalgo Reyes MÁ, Lagunes Barradas V, Pichardo Lagunas O, Martínez Seis B. A Hybrid Methodology Based on CRISP-DM and TDSP for the Execution of Preprocessing Tasks in Mexican Environmental Laws. In: Pichardo Lagunas O, Martínez-Miranda J, Martínez Seis B, editors. Advances in Computational Intelligence. Cham: Springer Nature Switzerland; 2022. p. 68–82.
[12] Huber S, Wiemer H, Schneider D, Ihlenfeldt S. DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model. Procedia CIRP; 2019;79:403–8.
[13] Kozjek D, Vrabič R, Rihtaršič B, Lavrač N, Butala P. Advancing manufacturing systems with big-data analytics: A conceptual framework. International Journal of Computer Integrated Manufacturing; 2020;33(2):169–88.
[14] Welte R, Estler M, Lucke D. A Method for Implementation of Machine Learning Solutions for Predictive Maintenance in Small and Medium Sized Enterprises. Procedia CIRP; 2020;93:909–14.
[15] Verein Deutscher Ingenieure e.V; 2014. Simulation of systems in materials handling, logistics and production: Fundamentals;ICS 01.040.03, 03.100.10(3633-1): Beuth. Berlin.
[16] Grambau J, Hitzges A, Otto B. Reference Architecture framework for enhanced social media data analytics for Predictive Maintenance models. 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC): IEEE; 2019. p. 1–8.
[17] Wolf Y, Sielaff L, Lucke D. A Standardized Description Model for Predictive Maintenance Use Cases. Procedia CIRP; 2023;118:122–7.