Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 93 (2020) 909–914

www.elsevier.com/locate/procedia

53rd CIRP Conference on Manufacturing Systems

A Method for Implementation of Machine Learning Solutions for Predictive Maintenance in Small and Medium Sized Enterprises

Rebecca Welte[a], Manfred Estler[a], Dominik Lucke[a,b]*

[a]Hochschule Reutlingen, ESB Business School, Alteburgstraße 150, 72762 Reutlingen, Germany
[b]Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany

* Corresponding author. Tel.: +49-711-970-1897; fax: +49-711-970-3606. *E-mail address:* dominik.lucke@ipa.fraunhofer.de

Abstract

In recent years, machine learning algorithms have made a huge development in performance and applicability in industry and especially maintenance. Their application enables predictive maintenance and thus offers significant efficiency increases. However, a successful implementation of such solutions still requires high effort in data preparation to obtain the right information, interdisciplinarity in teams as well as a good communication to employees. Here, small and medium sized enterprises (SME) often lack in experience, competence and capacity. This paper presents a systematic and practice-oriented method for an implementation of machine learning solutions for predictive maintenance in SME, which has already been validated.

Peer-review under responsibility of the scientific committee of the 53rd CIRP Conference on Manufacturing Systems

Keywords: Predicitve Maintenance; Machine Learning; Implementation Method.

1. Introduction

Today, many companies see Artificial Intelligence (AI), and there in particular the field of Machine Learning (ML) as an important strategic component with which they want to achieve competitive advantages [1–3]. With these technologies there is a large potential of ML techniques in manufacturing applications [4]. Especially methods used for predictive maintenance, which allow the reduction of unforeseen failures and improve the availability of machines and equipment, offer large opportunities for process optimization also in small and medium sized enterprises (SME). However, in order to fully exploit this potential of AI, a clear and company-wide strategy is required, even if the planned application only covers certain areas of use such as maintenance. Ideally, such an AI strategy covers all areas reaching from the company's computing center down to the IT based systems on the production shop floor. Experience shows that SMEs in particular find it very difficult to develop and implement such strategy processes. This is especially the case in connection with the introduction of new technologies such as AI or rather ML.

General domain agnostic process models already are developed for data mining projects in companies such as CRISP-DM [5] or ASUM [6,7]. These do not address all areas of great importance for manufacturing SMEs to a sufficient extent, in particular as they focus on business and technical aspects and general data mining projects [6,8,9]. Organizational aspects and the required staff development are hardly considered. Especially the latter is of particular importance for SMEs, as they usually do not have the necessary prior knowledge and related competences in the field of AI and ML under consideration [10,11]. Therefore, there is a substantial need for utilizing knowledge and experiences from outside the company. This can comprise the acquisition of consulting services as well as the selection and purchasing of adequate software tools [10]. Based on experiences of industrial projects in SME, applications there require a more specific guide or process model containing also

2212-8271 © 2020 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 53rd CIRP Conference on Manufacturing Systems
10.1016/j.procir.2020.04.052

recommendations regarding the organization, the required competencies and their development, the investment and operational cost as well as the data privacy issues. Another identified requirement from literature [10–12] as well as own experiences was the high demand for AI solutions in the improvement of maintenance and technical processes. Here, approaches based on statistical analysis and ML are important categories to predict machine and component deterioration and remaining useful lifetime. Especially the setup of ML projects and models is a challenging task for SME, which have mostly only a small budget and staff capacities. Therefore, this paper develops a holistic and generally valid process model for the development of ML solutions in maintenance especially focused on SME. Firstly, the evaluation of the previous related work shows some gaps, which were not sufficiently addressed so far in research. This work shows the most important fields of action from which specific ML projects can be derived. A process model is subsequently developed, especially taking into the account the requirements of SMEs. After the presentation of the process model, the validation is carried out on the basis of a real use case for ML-based maintenance of a medical technology company. It shows that the integration of ML into maintenance activities can offer an improved forecasting view.

2. Related Work

For the development of the method for the standardized introduction and value-added use of ML projects, a literature research was carried out, as well as different solutions of numerous consulting providers in the market were examined. [13] and [14] provide a broad overview of data mining process models. Considering the relevance, SME focus and applicability the following models have been selected and will be presented briefly.

CRISP-DM (Cross-Industry Standard Process for Data Mining), a general process model that has been developed and tested for a long time, is used to examine historical data stocks. CRISP-DM splits the data mining process into six main phases: business understanding, data understanding, data preparation, modeling, evaluation, deployment. [5] In the last few years the ASUM (Analytics Solutions Unified Method for Data Mining/Predictive Analytics) model has been developed by IBM. The ASUM model consists of five phases: analyze, design, configure and build, deploy, operate and optimize. The phases are controlled by a project management team. [7,9] Whereas CRISP-DM as well as ASUM are domain agnostic, the Data Mining Methodology for Engineering Applications (DMME) propose an enhanced CRISP-DM in order to address better the specifics of data analytics projects in a manufacturing environment. [8] The authors of [14] developed a conceptual framework for introducing big data analytics into manufacturing systems. The proposed framework addresses not only technical aspects but also required knowledge and competences in the different phases of a data analytics project. [14] The VDI guideline 3633 also provides a general process model for simulation models in materials handling, logistics and production, which also be applied for ML projects in production. It defines on a detailed level the phases of a typical simulation project from the simulation task definition and system analysis, data collection, model formalisation and implementation and the experiments. [15] Also, a market research offers various solutions from a cooperative development up to a complete plug and play solution or the simulation software only. In most cases of the existing approaches no difference was made according to the size of the company.

Most approaches were mainly developed by specialized data scientists of big companies und give clear guidance for the technical aspects of such a project. However, SMEs are typically in a completely different situation. They lack a clear digitalization strategy as well as skills and competencies of their employees in AI and have access to only very limited monetary resources for such projects [11]. The much more critical component for a sustainable implementation within SMEs, the consideration of affected employees and their qualification, as well as the organizational integration of such projects in companies are not taken into focus in existing frameworks. Therefore, an extended process model is required for project management in SMEs.

3. Concept

Addressing the identified requirements of SME and to overcome the drawbacks of other methods this chapter gives an overview of the proposed method for implementation of ML solutions in maintenance. The developed method consists of four main phases, each containing several steps. Here, the modelling of the main phases orientates on the typical big management decision milestones of such a project. It is enhanced on a process level and with two further components addressing the organisational aspects (Fig.1). The ML team serves the central control and coordination of the individual process steps as well as a competence centre for bundling and transfer of once acquired knowledge to the entire company. It is an interdisciplinary team consisting of employees from different fields of competence. The second important component is the qualification of employees, which also accompanies all four phases. A ML project is only successful, if affected employees recognize the objectives, reasons and advantages of the ML project. Moreover, all affected employees have to be given a basic understanding and specific training of AI with a focus on ML and the potential advantages in order to enable communication and discussion between the ML team and employees. In the end, it is the employee who decides whether an ML project is successful or not. In this work the general framework for implementing an ML application is exemplarily introduced and illustrated using a use case in predictive maintenance. But the framework is not restricted to the application in this particular field, it can rather be applied in a general way in many different use cases, such as quality impacting processes. The orientation should be fundamentally defined in the AI strategy.

3.1. Preparation phase

The intention of the preparation phase is to define the ML project and clarify the organization as well as the technical infrastructures. It comprises the definition of the business case, the selection of the implementation approach and ML infrastructure as well as the role assignment of the project team.

3.1.1. Maintenance ML use case

The starting point is the problem analysis and definition of objectives, in order to describe the maintenance ML use case.

ML should not be introduced because it is currently a trend technology with a high potential. Therefore, the first step is to define strategic goals for maintenance improvement that can only be reached with data mining and ML techniques economically. Strategic goals are independent from the technology, examples are to improve the availability or technical processes for a better quality in parallel with a reduction of maintenance cost. At the end clear and measurable goals have to be defined. Typically, possible machines and components either frequently fail unplanned and have high failure cost or are installed several times in the identical setup and thus offer a scaling effect. Then, it has to be checked if ML techniques are the most economical way to reach the defined maintenance goals on the selected machines and components. Methods for that are an extended root cause analysis and the determination of failure cost [16]. Based on experiences from previous industrial projects, it is recommended to begin with a smaller pilot project. In this way a SME can approach the topic iteratively with a reduced risk. Also, at the beginning of this phase the focus is less on the economic benefit than on the goal of quickly making progress and thus gaining experience and knowledge for scaling the developed solution on other components or machines and future larger projects. Even after the first pilot project, it is recommended to start with simple, minimally invasive projects with a short amortization time and clear measurable benefits. In order to get clear requirements, an extensive dialogue with the affected departments and employees such as the production planning, maintenance or logistics is indispensable. Here, a standardized procedure for description of the maintenance ML use case is recommended such as user stories, which is worked out in interdisciplinary workshops. At this point a training in AI and ML technologies comprising the definition, possibilities, benefits, risks and limits for all affected employees is recommended. Some employees, e.g. workers in the manufacturing, perceive AI application as a threat. This must be countered by an open and transparent information policy. The purpose here is that all stakeholders have a common understanding and ideas of AI and ML technologies. Also, this simplifies the discussion on the potential use cases. A typical ML use case description template contains the use case title, the problem to be solved, the desired state, the benefits and key performance indicators (KPI) as well as the ideas of the applied technology solution. The decision made here can have an immense influence on the further progress of the ML development in the company. Here, the wrong choice of initial applications increases the risk of early failure, provoking "scorched" earth for further AI and ML projects, making it difficult to obtain approval for further projects using this technology and drastically reducing the enthusiasm or willingness of the workforce.

3.1.2. Machine learning solution approach

After the definition of the maintenance ML use case, a suitable project approach has to be selected, depending in the form external partners are involved. For ML projects the following cases are suitable:

- application of external consultants and experts,
- cooperative projects with external service providers,
- body leasing of AI experts,
- cooperation with universities and research organizations
- application of plug and play ML systems including hardware and software as well as
- development of internal competencies.

The most popular approaches are the development of internal competencies as well as the use of external consultants and experts. Often a dual-skill strategy is pursued in combination, whereby skills are developed and expanded within the SME and professional support is obtained from external sources. In the long term however, it is more efficient to manage such projects purely internally and thus to secure a competitive advantage in accordance with the first mover principle. The choice of the project approach and, if applicable, the partner determines the degree of internal or external

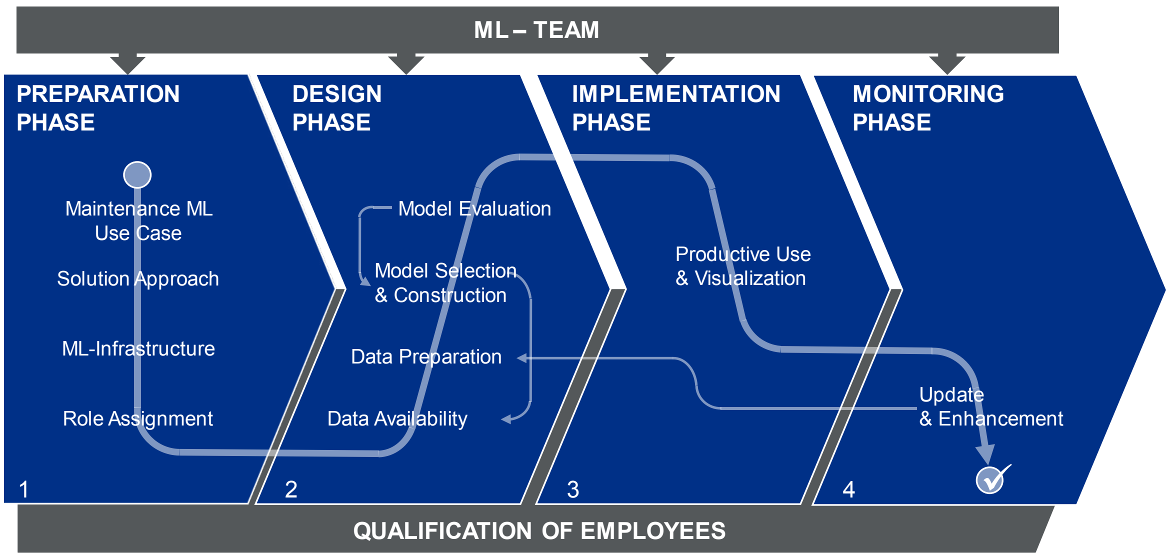

Fig. 1. Method for Implementation of Machine Learning Solutions

involvement, the distribution of roles in the ML project team itself and the scope of services.

3.1.3. ML Infrastructure

Depending on the selected project approach the type of ML infrastructure is influenced. It can be distinguished between the following three basic types:

- on premise solution with complete local IT hardware and software,
- cloud based solution or
- hybrid solution with varying part of local and cloud-based IT hardware and software.

The decision should be made in congruence the company-wide IT strategy. Further aspects that have to be considered in this context are the investment and operating cost, IT security and data privacy of the solution as well as the required competencies for development, implementation and operation. Often companies start ML projects with an on-premise solution, as critical data are kept in the private local environment. Changing from a local solution to cloud services is usually not a problem. A widely used tool for this purpose is the cost benefit analysis including pairwise comparison.

3.1.4. Role assignment

After the maintenance ML use case has been defined and the choice of a ML project approach and, if necessary, partners has been made, the employees required for the implementation of the project have to be selected. Maintenance ML projects should not be prematurely only assigned to the IT department, but should also be positioned in the business departments. The following five typical roles can be distinguished:

- The domain experts know the business process of the use case, are based in the specialist department, define the requirements for the ML project and are responsible for adjustments to the system if necessary.
- The data domain experts know what data is collected, where it is stored and enable access to further data sources.
- The data scientists are IT experts and statisticians with ML expertise. They are responsible for model design and data analysis.
- The software engineers are responsible for integrating the solution into the existing IT environment, to enable the productive use.
- The project manager is responsible for the coordination of the project and members included.

The involvement of technical experts from the domain is just as important as that of ML experts, who know how to deal with data. Which role is to be filled internally or externally by whom depends on the chosen solution approach and use case. The workforce of SMEs often shows deficits when it comes to sufficient resources of data scientist, as well as data domain experts and software engineers. Therefore, they have to integrate external experts, which leaves them in the unpleasant situation of disclosing their sensitive data. That’s why a carefully considered decision is essential.

In addition to these roles, other bodies such as the workers council should also be involved. A stakeholder analysis carried out in advance is helpful to maintain an overview of the many different parties involved.

3.2. Design phase

The design phase follows mainly standard processes for simulation projects [15], firstly the available data is sifted and further data recorded and processed. Then the ML model is built and evaluated. This takes up most of the time during the implementation of ML projects. This is to be understood as a major challenge for SMEs, which is also strongly related to employee qualification. Often such companies do not know how to start. Therefore, trained employees are needed in this area. Furthermore, in these cases there is usually no perfect starting point where the required data is already available online.

3.2.1. Checking data availability

Following the preparation phase, it must be checked which data is required for the design of the ML model.

This includes

- the definition of the first data set required in terms of measurands (AS-SHOULD-BE),
- checking available data in terms of measurands, frequency, resolution and reliability (AS-IS),
- the comparison the two data sets and derivation which required data have to be collected additionally in terms of measurands, frequency, resolution and reliability and
- the derivation of required sensors and data collection systems.

Most important step here is the first assessment if these data are suitable in terms of consistency, frequency and resolution. This technical question can be answered primarily by the domain experts, i.e. the maintenance staff and machine operators. Typical relevant measured variables in production include machines and components parameters such as vibration, temperature and acoustics, but also process parameters and the maintenance history.

Based on these findings, the data domain experts must clarify whether this data is already being collected, transmitted and stored in sufficient quantity and quality. The data quality depends on the consistency as well as on the sampling frequency. This determines the resilience and performance of the ML model to be developed in order to obtain reliable statements later on. Suitable communication channels must be used or set up for transmission. The storage and provision of the data should take place in a central IT system and not in a disorderly manner in different databases. The form of this central IT system depends on the chosen solution approach and the associated ML infrastructure. Increasingly, IoT platforms are being used on whose level not only the storage of data but also its analysis is possible. If the access to this data is not yet available, the required data sources must be tapped. Depending on the age and condition of the machines or components, there are different possibilities. In some cases, it is possible to work with already existing signals from the machines or larger distributed control systems by means of interfaces or existing sensor technology. Often however, the machines are not capable of this due to their age so a retrofitting project integrating new or additional sensors is required and triggered. Here, often minimal invasive sensor solutions are preferred in order to keep existing machine and process certifications.

3.2.2. Data preparation

Once the required data is collected and stored, it must be checked whether this data needs to be prepared, also known as data cleaning. This means that they are put into a form that can be used as a basis for data analysis, since the raw data are often not available in the correct format or have outliers. This is usually due to transmission or formatting errors that lead to incorrect, inconsistent, redundant or false data. Also, the test data set is split into training data and test data.

3.2.3. Selection and building of the maintenance ML model

After preparing the test dataset, the appropriate ML algorithm has to be selected and then the model has to be built. ML algorithms can be distinguished in reinforcement learning, supervised, unsupervised and deep learning algorithms [4]. Therefore, several algorithms and model parameters have to be tested in order to find the best matching algorithms. Here, a strong collaboration between the domain experts and data scientists is required, to evaluate the results and switch if required to another class of algorithms, e.g. based on regression or hidden-Markov models.

Major Problems with ML algorithms are that they require large data sets to create valid results. Approaches to reduce this problem is to create larger data set for example synthetic enriched data sets and co training [17]. Newer approaches using physics models to generate on less data more realistic values [18,19].

3.2.4. Evaluation of the model

After building the ML model, the evaluation step checks whether it meets the objective of the maintenance ML use case defined at the beginning. The ML model built is tested for its accuracy by applying to the unknown part of the test data set. Here the learned output is compared with the actual output. If the accuracy of the algorithm is not sufficient, model or data adjustments have to be made. It is difficult to quantify a sufficient accuracy, since this depends on the ML algorithms used. In unsupervised learning this evaluation is not possible, therefore, it is difficult to evaluate the quality of the model. A possible and at the same time risky way to do this is to integrate the model in the real environment and wait and see what results it delivers. Therefore, unsupervised algorithms are often used primarily for data exploration in order to gain a better insight and build up supervised models with detailed information content.

***3.3.** Implementation phase*

Once a stable and reliable ML model exists for the maintenance ML use case, it must be embedded in the factory environment and put into operation. On the one hand, the model is applied to live data in order to detect deteriorations in the condition or to predict failures in near real time. How the integration designed here depends on the chosen solution approach and the associated ML infrastructure. On the other hand, this analysis only generates added value when the information obtained is visualized or reported. Accordingly, software and end devices must be introduced that are easy to understand, user-friendly and mobile. Additional stress factors, be they visual by flashing or acoustic by loud alarm messages, have to be avoided. This aspect must be aligned with employee needs, since the full potential of ML systems is only activated when they are used correctly by employees. In the long term, they will only accept innovations if they are integrated into the planning and recognize a clear added value for their work. A first pilot project should be implemented up to this phase to test the integration possibilities and to receive direct feedback from the employees before a further rollout.

3.4. Monitoring phase

Periodic checks of the ML model are required in order to adjust and for extensions or refinements. Here, a differentiation is to be made on the basis of the training phase. If the training phase is limited in time to the model structure, additional offline training is periodically required to extend or refine the model. This training method is used much more frequently because it is easier to monitor. Alternatively, there are models whose training phase does not have a defined end and therefore develops continuously. Consequently, the model also learns during the analysis of the live data and adapts accordingly, whereby an overfitting could result. The check then serves to compare the model with the defined KPI.

4. Validation

The developed process model for the introduction of ML projects in maintenance has been tested successfully at a mid-sized medical device company. The company was at the beginning of the application of ML methods so that the experience level was low. A common understanding of the implementation method as well as the introduction of a ML application within the maintenance framework of the company could be achieved. For this purpose, a ML demonstrator was set up for a clear demonstration of the ML model, which was defined in such a general way that it is also applicable to other SMEs, as the requirements are comparable. Here, the focus of the presented approach is on the overall implementation process of a ML solution so that the ML model is only described briefly. The ML demonstrator focused to a failure prediction of a rotating drive unit based on electrical current data. As no previous experiences exist in ML in the company widely used unsupervised learning algorithms for data exploration have been used for the ML model. Based on the results more detailed prediction models using supervised learning methods are in development. A big advantage for the development of the ML demonstrator was an already existing reliable source of electronically stored real-time data of this particular machine. Since no further data from other sources were needed in this case, the demonstrator could be developed rather quickly. Particularly outstanding in the developed model are the data-independent components: the preparation phase at the beginning of the process and the employee support by the local innovation team in parallel. All employees at the site were given the opportunity to receive further training in digitization and to learn about current topics on site. The employees affected by the ML project were specially picked up. The success of the project depends on the coordination and exchange of information between the domain experts, who know the assets, and the data scientists, who know where what kind of data is stored and how to work with it. The ML model was applied in the reference company in maintenance, as it is one of the biggest levers for cost reduction within the fully automated production lines, which offers a huge amount of machine data or even the possibility to make the data accessible. Even unplanned downtimes of individual plant components can cause complete production downtimes and correspondingly high losses. The economic added value of the business case, which relates to the units of rotating drives that are very often installed in the production, can thus be easily calculated. Furthermore, the transferability to another site with rotating drives has been tested too.

Through this work, the awareness of employees in the production and the maintenance department was positively developed. Ideas for further applications of ML methods for predictive maintenance were created by the workers. Thus, the implementation method proposed increased the motivation of the workers towards this new technology.

5. Conclusions

This paper presents a systematic and pragmatic method for implementation of machine learning solutions for predictive maintenance in SME. It comprises the whole process from the initial idea, the organizational and technical setting of the project as well as the design, implementation and monitoring of the ML solution. Using structured phases enriched with recommendations it is suitable for SME with little or no AI or ML competencies. The developed method has been tested and validated in a medical technology company. The validation showed that the mainly the capabilities of SME regarding competencies in data science and ML methods currently restrict possible applications to simple use cases. Thus, especially for SME the access to external ML experts from consulting companies or software vendors build an important success factor when starting the first piloting projects. This enables a step-by-step transformation process for a dual skill strategy for the internal competence development, which is essential for a sustainable and irreversible introduction of ML as a strategic component for the future development of a company. Many service providers having long experience with ML application, e.g. in the field of predictive maintenance, are available on the market and can assist with the introduction of this technology. Nevertheless, companies are well-advised to not completely rely on this external support. Otherwise the full potential of ML technology cannot be achieved for the company.

References

[1] Bundesministerium für Wirtschaft und Energie. Potenziale der Künstlichen Intelligenz im produzierenden Gewerbe in Deutschland. https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.html. Accessed on 27.01.2020.

[2] Hecker D, Döbel I, Petersen U, Rauschert A, Schmitz V, Voss A Zukunftsmarkt Künstliche Intelligenz. Potenziale und Anwendungen. St. Augustin: Fraunhofer-Allianz Big Data; 2017.

[3] Purdy M, Daugherty P. Why artificial intelligence is the future of growth. https://www.accenture.com/t20170524t055435__w__/ca-en/_acnmedia/pdf-52/accenture-why-ai-is-the-future-of-growth.pdf. Accessed on 31.01.2020.

[4] Wuest T, Weimer D, Irgens C, Thoben K-D. Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research; 2016;4(1):23–45.

[5] 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.

[6] Angée S, Lozano-Argel SI, Montoya-Munera EN, Ospina-Arango JD, Tabares-Betancur MS. Towards an Improved ASUM-DM Process Methodology for Cross-Disciplinary Multi-organization Big Data & Analytics Projects. In: Uden L, Hadzima B, Ting IH, editors. Knowledge management in organizations: 13th International Conference, KMO 2018, Žilina, Slovakia, August 6–10, 2018 proceedings. Cham: Springer; 2018. p. 613–624.

[7] IBM Corporation. Analytics Solutions Unified Method - IBM. ftp://ftp.software.ibm.com/software/data/sw-library/services/ASUM.pdf. Accessed on 30.01.2020.

[8] 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.

[9] Ponsard C, Touzani M, Majchrowski A. Combining Process Guidance and Industrial Feedback for Successfully Deploying Big Data Projects. OJBD; 2017;3:26–41.

[10] Reder B Studie Machine Learning / Deep Learning 2019. München: IDG Business Media; 2019.

[11] Märkel C. Künstliche Intelligenz im Mittelstand: Relevanz, Anwendungen, Transfer. https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/kuenstliche-intelligenz-im-mittelstand.pdf?__blob=publicationFile&v=6. Accessed on 05.04.2020.

[12] Hermann D, Sieber JV, Lämmle J, Berger M, Prünte M, Friese P. Technologieanwendung im industriellen Mittelstand in Baden-Württemberg: Welche innovativen Technologien kommen in Produktion und Logistik zum Einsatz? https://www.ipa.fraunhofer.de/de/Publikationen/studien/Bestellung_der_Studie_RZI40.html. Accessed on 03.04.2020.

[13] Kurgan LA, Musilek P. A survey of Knowledge Discovery and Data Mining process models. The Knowledge Engineering Review; 2006;21(1):1–24.

[14] 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.

[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] Lucke D, Adolf T, Le TT, Bérenguer C, CHRISTIEN J, Sieg-Zieba S, HAUG R. Integrated dynamic predictive maintenance planning with advanced deterioration and remaining useful lifetime estimation models. In: Podofillini et al, editor. 25th European Safety and Reliability Conference (ESREL 2015). Zürich, Switzerland: CRC Press; 2015. p. 2261–2269.

[17] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training. In: Bartlett P, editor. Proceedings of the eleventh annual conference on Computational learning theory. New York, NY: ACM; 1998. p. 92–100.

[18] Karpatne A, Atluri G, Faghmous JH, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova N, Kumar V. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. IEEE Trans. Knowl. Data Eng; 2017;29(10):2318–31.

[19] Pillai P, Kaushik A, Bhavikatti S, Roy A, Kumar V. A hybrid approach for fusing physics and data for failure prediction. International Journal of Prognostics and Health Management; 2016;7(25):1–12.