**Appendix D**

This appendix details the challenges hindering the adoption of Machine Learning (ML) for Predictive Maintenance (PdM) identified from the literature. This appendix is composed of two main parts: i) description of the challenges identified through the SLR and ii) summary table (Table C1) that links each challenge to the documents that mention it.

As described, the SLR led to the identification of 29 pertinent documents. Through the analysis of these documents, 15 relevant challenges were identified. For each challenge, the name and the related proposition are reported below. Furthermore, relevant references that were exploited to define the challenges are mentioned. In this context, relevant quotation from the papers are mentioned as well. A given challenge could be mentioned by different documents. Table C1 provides an overview of which challenges are mentioned by which documents.

***Data Heterogeneity***

Real data used for PdM comes from several sources (e.g., different sensors) and they are often non-structured, representing an obstacle for ML algorithms (Rezaeianjouybari and Shang, 2020). Indeed, data collected through different sources may vary in terms of data format, structures, quality level, sampling rates, etc. (Kamm et al., 2023). Such heterogeneity of the input data limits the applicability of ML algorithms (Chen et al., 2023). Indeed, as stated by Erhan et al. (2021), “*analysing those data poses challenges as in how to combine and interpret the amount of (different) information provided*”.

Based on these considerations, the subsequent proposition was defined:

*C1: The adoption of Machine Learning is hindered by data coming from several sources (e.g., sensors, IoT devices etc…), which may be characterized by different formats, structures, quality level, sampling rates, time stamps, etc..*

***Data Scarcity***

To train ML algorithms, enough labelled data are required (Theissler et al., 2021). However, “*access to labelled failure data is scarce due to the rarity of failures*” (Ferreira and Gonçalves, 2022). Indeed, as stated by (Dalzochio et al., 2023), “*it is not possible to let the component evolve to failure, thus, making it impossible to collect failure data to train ML model*”. Thus, this leads to difficulties in training ML algorithms.

Based on these considerations, the subsequent proposition was defined:

*C2: An effective training phase of Machine Learning requires a large amount of failure data, which are often difficult to acquire in practical applications.*

***Data Storage***

As just described, a correct adoption of ML highly relies on data (Theissler et al., 2021). However, following the high number of sensors installed and the rarity of failures (Ferreira and Gonçalves, 2022), the number of data required to be stored is high (Dalzochio et al., 2023). This, however, represents a challenge for the adoption of ML algorithms as it leads to “*huge requirements of … storage*” (Wang et al., 2022). Similarly, Kamm et al., (2023) reported that “*to enable the application of the developed algorithms in an industrial setting, … high storage capacity is necessary*”.

Based on these considerations, the subsequent proposition was defined.

*C3: High requirements in terms of storage capacity are needed to store the data required by Machine Learning.*

***Training complexity***

The pursue of a high level of ML model accuracy can lead to overfitting problems (Dalzochio et al., 2023), especially when insufficient data are available. Indeed, as stated by (Fink et al., 2020), “*If insufficient data are available, … there is an increasing risk of overfitting*” This (i.e. the overfitting), in turn, makes so that *“ML models tend not to generalize well”* (Fink et al., 2020). In other words, in case of overfitting, ML models lack generalization capabilities. For this reason it is required to develop methods that can adapt to the data without overfitting (Hurtado et al., 2023).

Based on these considerations, the subsequent proposition was defined:

*C4: There is a high risk of overfitting when training a Machine Learning model.*

***Machine Learning Model Selection***

In the context of PdM, several different ML techniques can be used (Dalzochio et al., 2020). Each technique is characterized by a set of advantages and disadvantages, thus, “*the selection and implementation of appropriate algorithm or the approach is also very challenging*” (Gawde et al., 2023). Indeed, based on the considered scenario, the performance of different ML techniques may vary. It follows that “*a proper selection of existing methods is essential to achieve maximum performance in the health management*” (Nguyen et al., 2023).

Based on these considerations, the subsequent proposition was defined:

*C5: The selection of an appropriate Machine Learning algorithm is challenging due to the several available Machine Learning models.*

***Computational Complexity***

“*An effective training phase of AI algorithms requires a large amount of data samples*” (Nguyen et al., 2023). This may lead to long training times and high requirements for computational resources (Serradilla et al., 2022). Accordingly, “*the computational cost for training these ML models is also a challenge to be addressed*” (Dalzochio et al., 2020).

Based on these considerations, the subsequent proposition was defined:

*C6: Training a Machine Learning model requires a massive amount of data. Consequently, this results in extensive consumption of high-performance computing resources and long training times.*

***Feature Selection***

When considering ML for PdM, there could be applications characterized by a high number of features. Thus, “*in these cases, selecting the most critical features is necessary to avoid the problems generated by data excess*” (Dalzochio et al., 2023). However, “*feature selection and efficient feature comprehension relies heavily on labor and requires extensive domain knowledge*” (Ferreira and Gonçalves, 2022). This is also confirmed by Zhang et al., (2019), who mention that “*feature engineering requires expert knowledge and prior experience*”.

Based on these considerations, the subsequent proposition was defined:

*C7: Selecting the most relevant features is necessary to avoid the problems generated by data excess, but it is not an easy task since it relies heavily on labour and requires extensive domain knowledge.*

***Data Privacy***

ML for PdM requires acquiring data. In this context, sensors can be used. Sensors are not usually integrated with in-built protection mechanisms due to limited resources, thus, they are subjected to cyber-threats (Erhan et al., 2021). This could also be related to data storage. Indeed, when discussing the requirements for a successful adoption of ML for PdM, Gawde et al. (2023) stated that “*in the world of high security systems, data privacy is another important task*”. Similarly, also (Tran et al., 2023) reported that the use of ML for PdM is subject to the challenge related to the vulnerability against cyber breaches.

Based on these considerations, the subsequent proposition was defined:

*C8: The use of Machine Learning is accompanied by issues related to data breaches and cyber-attacks.*

***Infrastructure Selection***

As discussed above, PdM maintenance relies on vast amounts of data, especially when using ML. Hence, the infrastructure needs to efficiently store, manage, and process these data. However, the question emerges on whether it is better to do so on-premises or on a cloud infrastructure (Dalzochio et al., 2020). Similarly, when analysing these data and training the ML algorithms, high computation power is needed. But should the computational power be from edge computing, centralized servers, or a mix of these (Tran et al., 2023)? As an example, Erhan et al., (2021) stated that “*an important challenge is to identify techniques for computations*”.

Based on these considerations, the subsequent proposition was defined.

*C9:* *It is difficult to decide the best infrastructure for the intended application due to the variety and heterogeneity of available software and hardware.*

***High Infrastructure Cost***

As stated by Zhang et al., (2019), when discussing about the adoption of ML for PdM, “*issues such as economic costs need to be considered first in the implementation process*”. Indeed, as suggested by Dalzochio et al., (2023), “*high data density and high data collection rates may lead to network resource and storage consumption, demanding infrastructure-related investments*”. Another reason behind the high infrastructure costs is the high costs of sensors. As an example, Tran et al., (2023) reported that, in addition to inconsistent results of signals collected from different operation conditions, “*high costs are the shortcomings of the sensor*”.

Based on these considerations, the subsequent proposition was defined.

*10: High infrastructure-related investments are required to have an appropriate network and storage capabilities for Machine Learning.*

***Low Quality/Noisy Data***

Industrial environments are usually characterized by processes affected by noise (Fernandes et al., 2022). The presence of noisy data and other disturbance has progressively become relevant (Ren et al., 2023). This may lower the performance of some ML algorithms that are not able to properly deal with noise (Wen et al., 2022). In this context, the level of noise strongly depends on the source, which may also lead to errors or missing values (Chen et al., 2023), generating low quality data.

Based on these considerations, the subsequent proposition was defined:

*C11: The acquired data are affected by noise.*

***One Model for One Machine***

As previously mentioned, different ML techniques can be used for PdM purposes. The effectiveness and accuracy of a technique is strongly related to the considered equipment. Thus, its performance is case sensitive, and its application is not generalizable to all the systems (Nguyen et al., 2023). Moreover, “*previously validated predictive models for intelligent analysis of industrial data may not be applicable to new working conditions*” (Ren et al., 2023). In other words, the ML models are usually specific for a given asset and a given working condition due to difficulties in generalization

Based on these considerations, the subsequent proposition was defined:

*C12: Poor generalization ability of developed models: many developed models show poor performance when applied to different types or even the same type of equipment under different operational environments.*

***Machine Learning Model Interpretability***

ML and especially DL models may rely on features that are difficult to comprehend since they are not directly related to physical properties of the monitored asset (Nguyen et al., 2023). Thus, the interpretation of the model represents a challenge (Theissler et al., 2021). This is particularly true for complex models (e.g., generative models), which are likely to lack interpretability (Serradilla et al., 2022), leading to scarce trust of the industrial users.

Based on these considerations, the subsequent proposition was defined:

*C13: Features used by Machine Learning models are difficult to understand and interpret physically; consequently, industry experts do not trust the models.*

***Machine Learning Result Interpretability***

According to Ferreira and Gonçalves (2022), “*ML methods fall in the black-box category, such that their forecast loses the necessary interpretability*”, which “*refers to the transparent decision logic of the model itself*”. This issue is especially associated with DL, since “*the decisions of DL models lack explainability and interpretability*” (Serradilla et al., 2022), possibly due to the complexity of the model. It follows that the low interpretability of the ML results is a concern for domain experts (Fink et al., 2020), hindering the adoption of ML for PdM.

Based on these considerations, the subsequent proposition was defined:

*C14:* *The decisions of Machine Learning models lack explainability and interpretability.*

***Sensor Selection***

As previously mentioned, ML for PdM requires the installation of sensors to monitor process parameters based on which identifying the condition of the asset. The market offers a wide variety of sensors, which, “*despite the same purpose, generate different degrees of errors and the possibility of a sensor malfunction, leading to incorrect data*” (Dalzochio et al., 2023). Accordingly, the sensor selection is simultaneously a challenge and a vital activity that must be properly carried out (Gawde et al., 2023). This may help obtaining an accurate diagnosis process.

Based on these considerations, the subsequent proposition was defined:

*C15: There are a lot of types of sensors, and it is hard to decide which type to adopt based on the requirements in terms of reliability, accuracy etc.*

Table C1 below reports the links between papers and challenges. If a document mentions a given challenge, Table C1 reports an “X” in the cell identifying the intersection between the document and the challenge.

**Table C1**: Summary of the challenges mentioned by each of the 29 pertinent documents found through the SLR.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Paper** | **C1** | **C2** | **C3** | **C4** | **C5** | **C6** | **C7** | **C8** | **C9** | **C10** | **C11** | **C12** | **C13** | **C14** | **C15** |
| Chen et al., (2023) | X | X |  |  |  |  |  |  |  |  | X |  | X |  |  |
| Kamm et al., (2023) | X | X | X |  |  | X |  |  |  |  | X |  |  |  |  |
| Hurtado et al., (2023) |  | X |  | X |  |  |  |  |  |  | X |  |  |  |  |
| Dalzochio et al., (2023) | X | X | X | X | X | X | X | X |  | X | X | X |  |  | X |
| Nguyen et al., (2023) | X | X |  |  | X | X |  |  |  |  | X |  | X |  |  |
| Wang et al., (2022) | X | X | X | X |  | X |  |  | X |  |  | X |  | X | X |
| Gawde et al., (2023) | X |  | X |  | X |  | X | X |  |  |  |  |  |  | X |
| Tran et al., (2023) |  |  |  |  | X | X |  | X | X | X |  |  |  |  | X |
| Ren et al., (2023) | X |  |  |  |  |  |  |  |  |  | X |  |  |  |  |
| Yazdani-Asrami et al., (2022) |  |  |  |  | X | X |  | X |  | X | X |  |  |  |  |
| Serradilla et al., (2022) | X |  | X | X | X | X | X |  |  |  | X |  | X | X |  |
| Ferreira and Gonçalves, (2022) | X | X |  |  | X |  | X |  |  |  | X | X |  | X |  |
| Fernandes et al., (2022) | X |  |  |  |  | X |  |  |  |  | X |  |  |  |  |
| Pan et al., (2022) | X |  |  | X |  |  | X |  |  |  | X |  |  |  |  |
| Zhou et al., (2022) | X |  |  |  |  |  |  |  |  |  |  |  | X |  | X |
| Esteban et al., (2022) |  | X |  |  |  |  | X |  |  |  |  |  |  |  |  |
| Wen et al., (2022) | X | X |  |  |  |  | X |  |  |  | X | X |  |  |  |
| Theissler et al., (2021) |  | X |  |  |  |  |  | X |  |  |  |  | X | X |  |
| Kumar and Hati, (2021) |  | X |  | X |  | X |  |  | X |  |  |  |  | X | X |
| Erhan et al., (2021) | X |  | X |  |  | X |  | X | X |  |  |  |  |  | X |
| Zonta et al., (2020) | X |  |  |  |  |  |  |  |  |  |  | X |  |  |  |
| Dalzochio et al., (2020) | X |  |  |  | X | X |  |  | X |  | X | X |  |  | X |
| Rezaeianjouybari and Shang, (2020) | X | X |  | X | X |  |  |  |  |  |  |  |  |  |  |
| Fink et al., (2020) |  |  |  | X |  |  |  |  |  |  |  | X | X | X |  |
| Carvalho et al., (2019) |  |  |  |  |  |  | X |  |  |  |  |  |  |  |  |
| Zhang et al., (2019) | X |  |  | X | X |  | X |  |  | X | X |  |  |  |  |
| Ogunfowora and Najjaran, (2023) |  | X |  |  |  |  |  |  |  |  |  | X | X |  |  |
| Tama et al., (2023) |  | X |  |  |  |  |  |  |  |  | X |  | X | X |  |
| Zhao et al., (2020) |  | X |  |  | X | X |  | X |  |  |  |  | X | X |  |

**References**

Carvalho, T.P., Soares, F.A., Vita, R., Francisco, R. da P., Basto, J.P., Alcalá, S.G., 2019. A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering 137, 106024.

Chen, C., Fu, H., Zheng, Y., Tao, F., Liu, Y., 2023. The advance of digital twin for predictive maintenance: The role and function of machine learning. Journal of Manufacturing Systems 71, 581–594.

Dalzochio, J., Kunst, R., Barbosa, J.L.V., Neto, P.C.D.S., Pignaton, E., Ten Caten, C.S., Da Penha, A.D.L.T., 2023. Predictive Maintenance in the Military Domain: A Systematic Review of the Literature. ACM Comput. Surv. 55, 1–30. https://doi.org/10.1145/3586100

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., Barbosa, J., 2020. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. Computers in Industry 123, 103298.

Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., Liotta, A., 2021. Smart anomaly detection in sensor systems: A multi-perspective review. Information Fusion 67, 64–79.

Esteban, A., Zafra, A., Ventura, S., 2022. Data mining in predictive maintenance systems: A taxonomy and systematic review. WIREs Data Min & Knowl 12, e1471. https://doi.org/10.1002/widm.1471

Fernandes, M., Corchado, J.M., Marreiros, G., 2022. Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: a systematic literature review. Applied Intelligence 52, 14246–14280.

Ferreira, C., Gonçalves, G., 2022. Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods. Journal of Manufacturing Systems 63, 550–562.

Fink, O., Wang, Q., Svensen, M., Dersin, P., Lee, W.-J., Ducoffe, M., 2020. Potential, challenges and future directions for deep learning in prognostics and health management applications. Engineering Applications of Artificial Intelligence 92, 103678.

Gawde, S., Patil, S., Kumar, S., Kotecha, K., 2023. A scoping review on multi-fault diagnosis of industrial rotating machines using multi-sensor data fusion. Artif Intell Rev 56, 4711–4764. https://doi.org/10.1007/s10462-022-10243-z

Hurtado, J., Salvati, D., Semola, R., Bosio, M., Lomonaco, V., 2023. Continual learning for predictive maintenance: Overview and challenges. Intelligent Systems with Applications 200251.

Kamm, S., Veekati, S.S., Müller, T., Jazdi, N., Weyrich, M., 2023. A survey on machine learning based analysis of heterogeneous data in industrial automation. Computers in Industry 149, 103930.

Kumar, P., Hati, A.S., 2021. Review on Machine Learning Algorithm Based Fault Detection in Induction Motors. Arch Computat Methods Eng 28, 1929–1940. https://doi.org/10.1007/s11831-020-09446-w

Nguyen, K.T.P., Medjaher, K., Tran, D.T., 2023. A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines. Artif Intell Rev 56, 3659–3709. https://doi.org/10.1007/s10462-022-10260-y

Ogunfowora, O., Najjaran, H., 2023. Reinforcement and deep reinforcement learning-based solutions for machine maintenance planning, scheduling policies, and optimization. Journal of Manufacturing Systems 70, 244–263.

Pan, T., Chen, J., Zhang, T., Liu, S., He, S., Lv, H., 2022. Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives. ISA transactions 128, 1–10.

Ren, L., Jia, Z., Laili, Y., Huang, D., 2023. Deep learning for time-series prediction in IIoT: progress, challenges, and prospects. IEEE transactions on neural networks and learning systems.

Rezaeianjouybari, B., Shang, Y., 2020. Deep learning for prognostics and health management: State of the art, challenges, and opportunities. Measurement 163, 107929.

Serradilla, O., Zugasti, E., Rodriguez, J., Zurutuza, U., 2022. Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. Appl Intell 52, 10934–10964. https://doi.org/10.1007/s10489-021-03004-y

Tama, B.A., Vania, M., Lee, S., Lim, S., 2023. Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. Artif Intell Rev 56, 4667–4709. https://doi.org/10.1007/s10462-022-10293-3

Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., Elger, G., 2021. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. Reliability engineering & system safety 215, 107864.

Tran, M.-Q., Doan, H.-P., Vu, V.Q., Vu, L.T., 2023. Machine learning and IoT-based approach for tool condition monitoring: A review and future prospects. Measurement 207, 112351.

Wang, H., Zhang, W., Yang, D., Xiang, Y., 2022. Deep-learning-enabled predictive maintenance in industrial internet of things: methods, applications, and challenges. IEEE Systems Journal 17, 2602–2615.

Wen, Y., Rahman, M.F., Xu, H., Tseng, T.-L.B., 2022. Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective. Measurement 187, 110276.

Yazdani-Asrami, M., Sadeghi, A., Song, W., Madureira, A., Murta-Pina, J., Morandi, A., Parizh, M., 2022. Artificial intelligence methods for applied superconductivity: material, design, manufacturing, testing, operation, and condition monitoring. Superconductor Science and Technology 35, 123001.

Zhang, W., Yang, D., Wang, H., 2019. Data-driven methods for predictive maintenance of industrial equipment: A survey. IEEE systems journal 13, 2213–2227.

Zhao, S., Blaabjerg, F., Wang, H., 2020. An overview of artificial intelligence applications for power electronics. IEEE Transactions on Power Electronics 36, 4633–4658.

Zhou, H., Huang, X., Wen, G., Lei, Z., Dong, S., Zhang, P., Chen, X., 2022. Construction of health indicators for condition monitoring of rotating machinery: A review of the research. Expert Systems with Applications 203, 117297.

Zonta, T., Da Costa, C.A., da Rosa Righi, R., de Lima, M.J., da Trindade, E.S., Li, G.P., 2020. Predictive maintenance in the Industry 4.0: A systematic literature review. Computers & Industrial Engineering 150, 106889.