Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-19 (2022) 61–66

On the role of Data Quality in AI-based Prognostics and Health Management

L. Cattaneo*, A. Polenghi, M. Macchi**, V. Pesenti*****

**LIUC, Università Carlo Cattaneo, Castellanza, Italy (e-mail: lcattaneo@liuc.it)*

*** Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Milan, Italy (e-mail: adalberto.polenghi@polimi.it, marco.macchi@polimi.it)*

**** Consorzio Intellimech, via Stezzano 87, 24126, Bergamo, Italy (e-mail: valerio.pesenti@intellimech.it)*

Abstract: This paper aims at describing the key role of Data Quality along the entire development of a Prognostics and Health Management process based on Industrial Artificial Intelligence solutions and ready for industrial application. This is discussed through an industrial case in the textile sector where the importance of Data Quality emerges in different aspects. The industrial case leads to lessons learned useful for further research on a framework for Data Quality in AI-based maintenance systems.

Keywords: Industrial Data Analytics, Industrial Artificial Intelligence, Data Quality, PHM

1. INTRODUCTION

Industry is nowadays demanding an effective use of Artificial Intelligence (AI) to solve different problems across industrial processes such as production, maintenance, quality, logistics. Industry 4.0 has boosted this need, also leading to recognize Industrial AI as a powerful concept and method to be used for solving problems and for creating new values in such industrial processes (Lee, 2020). Therefore, problem-oriented applications are essential for the successful adoption of AI technology in the industrial field. Besides knowing the engineering tasks that are supported and the addressed problems, Industrial AI, as any other AI application domain, is based on data availability. In the context of Industry 4.0, lots of data are available from field thanks to the digitalization of industrial processes. Anyhow, companies often lack data featuring the proper quality level, as they miss the capability of data for providing useful information for decision support (Cattaneo et al. 2021). To fulfil the gap, this paper reflects on the need of a framework aimed at guaranteeing a proper level of Data Quality (DQ) in problem-oriented applications, having a special emphasis on the use of Industrial AI algorithms in advanced maintenance systems. Prognostics and Health Management (PHM) is considered as the body of knowledge providing the background for building advanced maintenance systems (Guillén et al. 2016). PHM is generally understood as the process of determining the current state of a production system or asset in view of reliability and prediction of its future state (Pellegrino et al. 2016), based on the control and analysis of Condition Monitoring (CM) data and degradation signals (Jardine et al. 2006). To be correctly implemented, PHM needs to follow a specific strategy, this can be formalized through the use of reference frameworks, as the case of the framework proposed in (Cattaneo et al. 2021) by extending the ISO 13374 – OSA-CBM standards. The specific objective of the paper is to use the development of an industrial application in the textile sector as a showcase to discuss different aspects to be considered in order to guarantee the DQ along the entire PHM solution development process. It allows to discuss the need of a framework to enable such a DQ for developing AI-based PHM applications. After a background on the concept of DQ in Section 2, Section 3 presents the industrial application in the textile sector, while Section 4 discusses the lessons learnt and concludes with some remarks for the future evolution of data quality in AI-based maintenance systems.

2. BACKGROUND

DQ is essential for a reliable decision-making process in the industrial environment (Yeganeh et al. 2014), where raw data must be transformed into useful information. However, particularly referring to the manufacturing context, DQ is a more general concept that encompasses more detailed dimensions. Namely, DQ dimensions include (Scannapieco et al. 2006):

- Accuracy, as the closeness between a value *v* and a value *v'*, which is thought as the correct representation of the real-life phenomenon that *v* aims to represent.
- Completeness, considered as *"the extent to which data are of sufficient breadth, depth, and scope for the task at hand"* (Wang et al. 1996).
- Consistency, related to *"the violation of semantic rules defined over (the set of) data items"* (Scannapieco, et al. 2006), e.g., the violation of integrity constraints or consistency rules.

Besides, (Tam et al. 2019) summarizes what presented by (Sarfi et al. 2012) as the 5C's of DQ, listing the following:

- Clean, intended as no error;
- Consistent, with no doubts about the correct version;
- Conformed, as data must be shareable for business use;
- Current, understood as always up to date;
- Comprehensive, when all needed data are available.

Also, in the (ISO 8000-8 2015) three categories are identified to measure DQ:

- Syntactic quality, as "*degree to which data conforms to its specified syntax"*;

10.1016/j.ifacol.2022.09.184

- Semantic quality, as "*degree to which data corresponds to what it represents*";
- Pragmatic quality, as "*degree to which data is found suitable and worthwhile for a particular purpose*".

More recently, (Lee, 2020) has introduced a characterization of the DQ according to "three B's", namely: Bad Quality, Broken and Background. This classification aims at managing the problems related to DQ in industrial data. In few words, the three B's primarily include a content similar to the 5C's of (Sarfi et al., 2012), whilst some additional remarks can be pointed out. Bad Quality explicitly refers to the capability of building a proper industrial IT architecture for data collection. Besides, Broken relates to the comprehensiveness of data to guarantee the capability of extracting key parameters and features in order to describe the object being investigated in accordance with the purpose of analysis; thus, when a proper quality is missed, the Broken can be interpreted as data with a low Pragmatic quality. Finally, the DQ concept is enriched by adding the Background part, meant as the capability of extracting hidden correlations among data, to finally label the data themselves. This task is often related to the availability of auxiliary data, such as work condition settings, maintenance records, task information, and so on. This is also a mandatory task when a supervised learning AI algorithm is implemented. The decomposition of DQ into several characteristics shows that there is not a unique way to look at it. Indeed, the issue is still open. This paper follows the 3B's classification to ensure DQ when developing an AI-based PHM application.

3. PHM DEPLOYMENT IN THE INDUSTRIAL CASE

The case refers to a manufacturer offering advanced weaving solutions. The rapier loom is the asset under study, selected from the asset portfolio of the company. This loom shoots the weft yarns across the warp one row at a time, working in a sequence to form the final textile fabric. A monitoring system is being developed, using data collected from optical sensors: the company aims at achieving a PHM application runnable in a real-time monitoring, so to enable technicians to reduce the downtime for maintenance interventions. The presentation of the case is streamlined according to the framework published in (Cattaneo et al. 2021). It provides guidelines for PHM implementation from a process viewpoint, focused on the interesting ISO 13374 – OSA-CBM levels and with the aim of extended them; Some of these levels (Lx), from 0 to 3, are illustrated in the next sections.

3.1 L0-Asset Analysis

Regarding the functioning of the asset, two rapiers on the two opposite sides of the loom reciprocally transfer the yarn. Each rapier is basically composed by a flat metal tape and by the so-called rapier head, see Fig. 1 for details. In turns, the rapier tape is a composite material made by three layers: Ceramic, Polyester and Carbon. The tape is in contact with the wheel throughout the operation cycle. Moreover, the tape is in charge of managing the head, finally allowing the creation of the textile fabric. The flat metal tape has been recognized as one of the most critical item for which it is worth dedicating a PHM application. Then, the tape has been investigated by means of an engineering analysis of the failures, to better set the problem. In particular, it is known that failure modes depend on the different parts of the surface of the tape, as the degradation process is effectively different. The external surface is in contact with pressure plates, used to maintain the tape's rigidity during the launching process: over time, the external surface is then subject to wear out, due to the continuous rubbing with the plates themselves; this wear can lead to a misalignment of the two rapier heads as failure mode. The internal surface is also subject to wear, from a continuous scrolling with the wheels, leading to the tape thinning as failure mode. About the lateral surfaces, with the rapier entering and exiting the shed at high speed, the yarn generates alternating compressive stress on the lateral surfaces, causing the peeling off from the external layer on; a difference in tape height is formed between substrate and the more external layer, leading to a cutting edge, which can finally cause the warp breakage as failure mode. The PHM application is focused on the avoidance of the warp breakage: considering their faster degradation, lateral surfaces of the tape are agreed with the experts of the rapier loom to be the critical part of the item to be monitored.

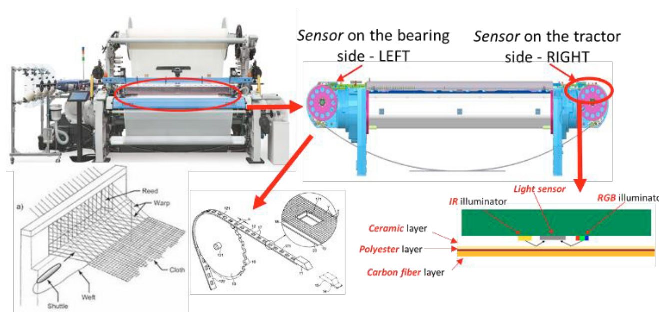

Figure 1. On the top left, the rapier loom. On the top right, zoom on the position of the sensors. Below, starting from the left, the weft and the warp, the tape and the wheel and a schema explaining the optical sensor functioning over the tape (and its layers).

3.2 L1-Data Acquisition

Two light sensors are used for the Data Acquisition (DA) phase. The sensors are mounted on the two opposite sides of the loom, to monitor both the two tapes (i.e. left and right sides of the loom, as reported in Fig. 1). Each sensor consists of two illuminators: one emits Infrared light beams, while the other emits visible light beams at different frequencies, i.e. Red, Green and Blue. The emitted beams hit the tape surface and are reflected, and the sensor is employed to detect these reflected light beams (Fig 1, bottom right). The choice of using this kind of sensor lies on the physical principle that different materials are expected to reflect different light beams. Thus, the capability of associating a different reflection to a different material in relationship to the different layers of the tape could be then employed to derive the state of wear. The difference in the reflection signal physically describes the process for which a new material compares on the surface tape due to the degradation and disappearance of the previous layer.

Data were collected from test benches starting from September 2020 to January 2021, with more or less a daily sample time (for about 130 days sampled). So, it means that each day the company has collected data for reflectivity for the left tape and the right tape, each of them spatially divided into 500 segments over the tape length, and considering all frequencies (Infrared,

Red, Green and Blue). Moreover, it is important to notice that data are collected starting from the moment of installation of a new tape till the moment in which the same tape has reached its end of life (EOL) due to an advanced state of degradation. Overall, it is worth remarking that the role of the company as the OEM (Original Equipment Manufacturer) of the rapier loom enabled to easily motivate, from a prior engineering knowledge, the use of the optical sensor to promote a high detectability of the evolution of the failure mode under study in the set problem. For the sake of simplicity, in the rest of the paper only the analysis for the left tape is presented.

3.3 L2-Data Manipulation

The Data Manipulation (DM) phase aims to process raw data so to obtain high-quality data, where quality is intended as described in the 3B's classification. Typically, it encompasses three steps: (i) Data pre-processing, aimed at checking integrity and consistency of the collected raw data, smoothing and eliminating the noise that might characterize them, coping with missing values or errors and making any transformation convenient for specific uses such as scaling; (ii) Feature extraction starting from the acquired signals and (iii) Feature selection, with the end result of obtaining the key parameters and features as indicators that properly describe the degradation process of the item/s under analysis. As implementation of step (i), an initial cleaning process leads to reduce the sample to about 100 days, after discarding data with evident errors. This enabled to avoid "Bad Quality" due to the data collected in the test benches. No specific detail will be provided to this regard in this paper.

Then, an assessment plan was particularly developed by defining a list of questions due to the problem-oriented application. These should guide the data analysis, primarily aimed at implementing steps (ii) and (iii) of this phase: (Q1) Is the reflectivity a proper indicator of the state of wear? (Q2) Are all frequencies proper indicators of the state of wear? (Q3) Is there any combination of the frequencies that describes the state of wear, so selecting a smaller number of features?

Concerning (Q1), analysis of variance with multiple responses (MANOVA) has been performed, taken as response variables the different frequencies. The analysis has used MATLAB®, in particular the function *manova1* (Krzanowski, 1988). Both time and space have been considered as factors that can influence the different frequencies of reflectivity. Starting from the time, groups of different levels of this factor have been defined in light of the different times registered in the dataset, ranging from the installation time of the tape, when the tape is new, to the last moment as the EOL, when the tape must be removed and changed. *manova1* returns a vector of p-values for testing the null hypothesis that the mean vectors of the groups lie on a space of various dimensions, i.e., starting from 0 dimension (that means equal averages) till at least 4 dimensions (if all the tests are rejected). The p-values lead to conclude that not only the averages in time are not the same, but also that they are in a space of more than 4 dimensions. In other words, reflectivity is confirmed as a good indicator of degradation until EOL, since it is able to catch the variability (so, degradation effect) over time. The analysis has been replicated considering the space factor. In this case, to "freeze" time and to analyse variability in space (i.e., along the length of the tape), only data belonging to the installation time and to the EOL of the tape were considered. Data have been grouped by *manova1* according to the segments over the length, i.e., from 1 to 500, in each time. Also in this case, *manova1* results in rejecting the null hypothesis of the test, allowing to assert that reflectivity at the different frequencies have different averages in space, i.e., they are able to express variability (so, degradation effect) also over the space dimension. Concerning (Q2), in order to understand if all the frequencies are able and needed to describe the state of wear, a correlation analysis has been performed, considering the entire dataset (all the segments, for all the times, for all the frequencies). Correlation has been computed using the Pearson's coefficient. Results are reported in Fig. 2, and they present a very high correlation degree between Green and Red (95.7%), between Green and Blue (94.4%) and between Red and Infrared (95.4%). These results allow to conclude that all the frequencies are correlated and therefore able to represent the wear process of the tape, as they are reporting similar piece of information.

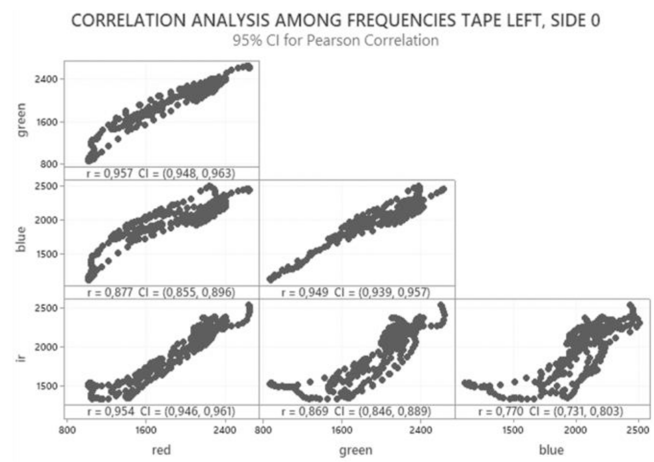

Figure 2. Correlation analysis among frequencies tape left.

Regarding (Q3), the answer corresponds to the extraction and selection of new features resulting from extant features combination. These new features should be able to describe the variability of the wear process and to simplify the analysis, both in time and space, by taking advantage of a lower dimensional space. A Principal Components Analysis (PCA) has been performed (Johnson et al. 2013). PCA is concerned with explaining the variance-covariance structure of a data set described by a certain number of variables through a lower number of components obtained as a linear combinations of the original variables. In this case, PCA was conducted using the whole dataset, so analysing data in terms of times, segments and frequencies. To take into account the segments position, a variable "tape" has been added to the data set, which takes values between 1 and 500, according to the position from which data are collected. From Table 1, it is possible to observe the eigenvalues of the correlation matrix for each principal component (PC), with the relative portion of expressed variability. Usually, the number of PC is chosen such that the percentage of variability expressed is the 80/90% of the original one. So, in this case, it is good to consider at least the first two PCs. Remembering that the magnitude of the elements of the eigenvectors measures the importance of the *p*-th variable to the *k*-th PC, irrespective of the other variables, it is possible to better analyse the results, looking at Table 2.

Table 1. Eigenvalues and Explained Variance

	PC1	PC2	PC3	PC4	PC5
Eigenvalues	3,6186	1,0463	0,2385	0,0653	0,0314
Proportion of variability	0,724	0,209	0,048	0,013	0,006
Cumulative	0,724	0,933	0,981	0,994	1,000

Table 2. Eigenvectors and Principal Components

Variable	PC1	PC2	PC3	PC4	PC5
Tape	-0,064	0,963	-0,234	-0,114	0,003
Red	0,513	0,040	0,269	-0,517	-0,629
Green	0,515	-0,061	-0,169	-0,446	0,710
Blue	0,487	-0,089	-0,710	0,426	-0,263
Infrared	0,480	0,242	0,583	0,583	0,177

In particular, to define the first PC, almost all the four frequencies are significant, even if Green is slightly more important (0.515). Then, about the second PC, the Infrared is definitively more important than all the other frequencies (0.242 as can be seen from Table 2, PC2). Considering, in addiction, that the correlation analysis highlighted a high correlation among frequencies as general trend (>89.6%, as average), it is possible to conclude that the variables Green and Infrared could be selected as main features, since they are able also to cover the variability explained by Red and Blue, that could be therefore discarded for the rest of the analysis.

The DM phase has been enriched by a further step of analysis, to complete the feature engineering (as a completion of step ii) and iii)). Considering the huge quantity of data to be analysed both in time and space, it has been explored the possibility to reduce the space complexity. Indeed, it has been noted that some portions of the tape length degrade in the same way. For this reason, a new question was defined: (Q4) Is it possible to reduce the control of the degradation to a few points in the tape length? To this end, a clustering algorithm has been applied; its objective is to group consecutive segments characterized by the same level of wear. Clustering has been developed respecting some constraints: 1. space constraint, as it is fundamental that the original spatial position of the segments is maintained, since the wear is not uniform along the tape (some segments are wearing, others do not wear out at all); 2. time assumption, for simplicity it was assumed that the degradation inside each cluster is uniform in time; 3. frequency assumption, as clustering is made taking into account the Green frequency as selected feature from the previous step. To cope with the space constraints a clustering algorithm was developed as a customization of K-means clustering method. A short outline of the customization logic is hereafter presented: (i) Initial number of clusters is selected through the computation of the Silhouette indexes as proxy of the goodness of the clustering outcome (for details on Silhouette indexes see (Johnson et al. 2013)); (ii) K-means clustering is applied after receiving as input the data and the number of clusters obtained at previous step; (iii) an Adjustment function adjusts the clusters obtained by K-means, based on the segments position, and in light of a tolerance set to find a good quality of the clustering strategy; it results in the final number of clusters. After the algorithm is run, a final configuration is obtained, leading to a number of clusters equal to 8, with an average Silhouette index equal to 0.3714. Thanks to this clustering algorithm, it is then possible to limit the time analysis to just 8 points, the centroids of each cluster. Moreover, the analysis of the degradation process of the clusters centroids allows to make a step forward through the assurance of DQ. In this case, DQ has to be intended as the capability of selecting only the data that are really necessary to perform the analysis. Indeed, from the clustering algorithm results, it has been possible to further recognize two different subgroups. The first contains clusters that do not change their trends over time in a significant way (clusters 1, 2, 5 and 8), while the second one contains clusters that show important changes in reflectivity (clusters 3 and 6). In Fig. 3, cluster 8 and cluster 3 are reported as example of elements belonging to the two different subgroups. This result implies that not all the segments are affected by degradation, *alias* not all the segments need to be controlled. Thus, the focus of the analysis has been moved on the second subgroup and cluster 3 and cluster 6 have been compared in terms of trends, to find out the cluster that registers the highest variation over time.

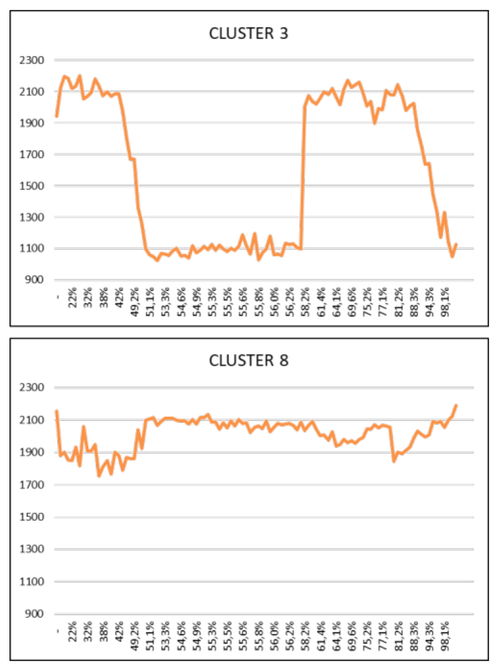

Figure 3. Clusters 8 and 3 are reported to compare the two subgroups recognized after the cluster analysis. On the x-axis time is reported in terms of percentage of number of strokes.

This cluster is also called the "bottleneck" cluster and, being the faster cluster in terms of degradation, it will be responsible for determining the EOL of the whole tape. The selection algorithm has been implemented checking the variation in terms of slope of the trend of the centroids. As result, cluster number 3 has been selected as "bottleneck". From now on, the analysis is performed by taking as reference cluster only the bottleneck one, which contains segments numbered from 202 to 235. Overall, going through Q1-Q4 enabled to increase the quality of the manipulated data, leading to what can be considered for the case the most comprehensive data model, even with limited dimensionality. On one hand, reflectivity at Green and Infrared frequencies may suffice as selected features, or eventually the first two PCs of the whole frequencies (if it is acceptable to lose the physical meaning). Besides, the trend analysis of the degradation process can be limited to analyse the clusters' centroids with special focus on the cluster showing the faster degradation, i.e., the "bottleneck". This high quality in the understanding of the problem is essential to be confident to have prepared the ground for the next steps with engineered features as indicators to establish a complete performance evaluation and prediction model. In other words, while making the data manipulation through Q1-Q4, "Broken" quality issues should be avoided. The last part of this DM phase is eventually concerned with another important issue, that is directly connected with the third B of the 3B's classification: Background. It is important to recall that Background refers to the capability of correlating the collected data with other sources of information to make possible the "labelling" process (Lee, 2020). In this specific case, the "labelling" process has been developed analysing further in details the available data and exploiting the engineering knowledge about the physical composition of the metal tape and its way of degrading over time. The objective is to discover how much and in which ways the tape changes its reflectivity. Looking at the bottleneck cluster, it is possible to notice the presence of four well distinct zones. Considering also that the tape is composed by three layers of different materials, it is assumed that degradation is strongly related to the structural composition of the tape. As the weft's wires repeatedly rub against the tape, the friction and the resulting surface stress cause a progressive wear of the lateral surface, until it is completely worn, and the underlying layer emerges. So, detection of the slope change is needed to opportunely label the different zones: the first plateau corresponds to the ceramic layer (new tape), the second plateau corresponds to the rising of the polyester layer, the third plateau corresponds to the rising of the carbon fibre layer and the last transient (last decreasing pattern) corresponds to a worsening of the conditions of the last layer (carbon fibre degradation). See Fig. 4 for a visual explanation. This analysis has allowed to understand that it is necessary to define a multi-state model (Lei et al. 2018), where the unhealthy state has to be further divided into different states, according to the different composition of the tape. Moreover, data have been labelled according to which state they belongs.

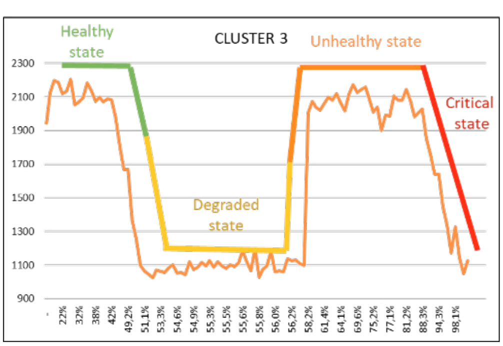

Figure 4. Multi state degradation model as seen from the bottleneck cluster. On the x-axis time is reported in terms of percentage of number of strokes performed by the loom.

3.4 L3- State Detection

The DM phase has led to the identification of relevant features for understanding the tape states. Further, the labelling process has allowed to define a multi-state model. Now, the State Detection (SD) phase should be able of building an alerting system. Exploiting again the engineering knowledge, it has been decided that an alert should be sent each time a new layer appears, since this means that the tape is getting thinner. Then an alarm, indicating a serious level of wear, should correspond to the degradation of the last layer, the carbon fibre. The labelling process has allowed to implement a supervised learning algorithm. The MATLAB® Classification Learner Toolbox has been used for the implementation of the SD algorithm. The model has been trained using data referred to the bottleneck cluster, organized in such a way.

(i) The response variable is a binary variable, that represents the state of the tape. Each state is indicated with 0 or 1 and corresponds to the appearance of a specific material layer: state 0 corresponds to ceramic layer, state 1 corresponds to polyester layer, state 0 corresponds to carbon fibre layer, state 1 corresponds to the last state that represents the degradation of the carbon fibre layer. It is worth noting that the same label appears twice (both 0 and 1) since the reflectivity values corresponding to the ceramic layer are similar to the ones representing the carbon fibre, and the same situation happens in the case of polyester and the final degraded condition, i.e., degraded carbon fibre layer.

(ii) The regressors of the model are the reflectivity values of the selected frequencies (Infrared and Green) as features measured on an identified segment over the tape length, that belongs to the bottleneck cluster. It has been selected segment 222 since it is the "median" segment of the bottleneck cluster. Among the alternatives provided by the Classification Learner Toolbox of MATLAB®, a Linear Discriminant classification model has been selected. The selection has been furtherly validated making two tests on two different datasets: the reflectivity of the same identified segment using the other two frequencies (Blue and Red) and the bottleneck centroid, using again Green and Infrared. For each test the values of Precision, Recall and F1 score are tabled in Table 3.

Table 3. Validation tests

	Precision	Recall	F1 score
Test 1	99,11%	98,98%	99,04%
Test 2	97,12%	97,27%	97,19%

High scores are obtained for each computed indicator, so it is possible to conclude that the model works properly and can be used as detector of the wear state. However, the model presents a limitation, as it is able to distinguish only two states: 0 and 1, which does not help to detect the appeared layer. To solve this issue, the classification is nested within an algorithm able to check for a difference in light of the state transition: i) if the transition is from 0 to 1, then the change of state regards the first layer, as it changed from ceramic (0) to polyester (1); ii) otherwise, the change of state regards the third layer, from polyester (1) to carbon layer (0). Joined with these simple rules, the algorithm allows to recognize which layers emerge during the degradation process.

4. DISCUSSION AND CONCLUSIONS

Checking the DQ in PHM is a mandatory requirement before implementing any kind of AI algorithm in the PHM process. The industrial case, showcased in the paper, is an exemplary proof of the need of a systematic approach that should consider the DQ with its role embedded along the development of the PHM processes. The industrial case permits to report some lessons learnt (LLx).

(LL1) DQ is inherently guaranteed when a proper framework is considered, from the data acquisition until the validation of the AI algorithm results: this framework is an aid to promote a consistent and holistic view.

(LL2) Within the same framework, the initial phases – DA and DM phases – are key to elaborate the raw data to achieve the need of high-quality data. During these phases, different DQ aspects may be considered: i) the selection of adequate data sources (sensors, controllers, etc.) enabling the detectability of the evolution of the failure mode; ii) the data pre-processing, aimed to provide cleaned data to the algorithm/s, thus avoiding a "Bad Quality" due to the collected raw data; iii) the feature engineering, to guarantee a high quality in the comprehensive capability to describe the object under study, and to avoid the occurrence of the "Broken" issues; iv) the "labelling" process, to correlate the patterns discovered in the analysed data to the extant sources of information representative of the operation of the object under study, to avoid the risk to loose knowledge of the "Background".

(LL3) Guaranteeing the DQ requires a mixture of engineering knowledge, rules of thumbs and mathematical methods, and lots of effort. Methods that can be used to assess the DQ are taken from multivariate statistical analysis (as data visualization, analysis of variance, hypothesis testing, PCA, clustering). They are aimed at specific tasks, such as data reduction and simplification, investigation of the dependence among variables, grouping, hypothesis testing. These are necessarily joined with an engineering knowledge of the object under study, which is relevant for i) providing initial hypothesis about the way in which the object is expected to response/behave, ii) driving some choices in the algorithm/s to comply with physical constraint or meaning, and/or iii) validating the results with respect to the engineering task expected by the algorithm/s and, thus, the decision support.

In the end, the authors look for future works to gain a synthesis of the wide knowledge required to guarantee DQ in AI-based PHM development. This inspires a next research step, aimed at systematizing the DQ theory inside the extant framework.
Last but not least, future works have to focus also on the development of the last phases of the framework, with particular attention to L5, i.e. the Prognostics one. Requirements in terms of DQ have to be defined and validated to guarantee robust results for AI-based algorithms for prognostics assessment. Indeed, even if the showcase here reported has a complete history of run to failure data, prognostics has to take into account different sources of variation, such as working conditions, age of the loom, maintenance interventions previously performed, etc. Therefore, a single run to failure data collection is however not sufficient to explore this last level of the framework.

ACKNOWLEDGEMENT

This paper is the result of a case study developed during the Smart Manufacturing Lab Course of Politecnico di Milano. We thank the students Martina G., Andrea C. and Mohamed H. D. S. for their passionate contribution in developing the work here presented.

REFERENCES

Cattaneo, L., Polenghi, A., and Macchi, M. (2021). A framework to integrate novelty detection and remaining useful life prediction in Industry 4.0-based manufacturing systems. *International Journal of Computer Integrated Manufacturing*, 1–21.

Guillén, A. J., Crespo, A., Macchi, M., and Gómez, J. (2016). On the role of Prognostics and Health Management in advanced maintenance systems. *Production Planning and Control*, 27(12), 991–1004.

ISO 13374-1. 2003. Condition Monitoring and Diagnostics of Machines — Data Processing, Communication and Presentation— Part 1: General Guidelines. *British Standards Institution* 3: 31.

ISO 8000-8. 2015. "Data Quality - Information and Data Quality: Concepts and Measuring." *BSI Standards Publication*.

Jardine, A.K.S., Lin, D. and Banjevic, D. (2006). A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance. *Mechanical Systems and Signal Processing*.

Johnson, R. A. and Wichern, D. W. (2013). *Applied Multivariate Statistical Analysis.* Pearson New International Edition PDF eBook. (6th Edition).

Krzanowski, W.J. (1988), *Principles of Multivariate Analysis*, Oxford University Press, Oxford.

Lee, J. (2020). *Industrial AI,* Chapter 2*,* Springer, Singapore.

Lei, Y., Li, L. N., Guo, N., Li, T. Y., and Lin, J. (2018). Machinery Health Prognostics: A Systematic Review from Data Acquisition to RUL Prediction. *Mechanical Systems and Signal Processing* 104: 799–834.

Pellegrino, J., Justiniano, M., Raghunathan, A. and Weiss, B.A. (2016). *Measurement Science Roadmap for Prognostics and Health Management for Smart Manufacturing Systems*.

Sarfi, R. J., Tao, M. K., Lyon, J. B., and Simmins, J. J. (2012). Data Quality as It Relates to Asset Management. *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*.

Scannapieco, M., Batini, C. (2006). *Data Quality: Concepts, Methodologies and Techniques. Data-Centric Systems and Applications*. Springer.

Tam, A. S. B., and Price J. W. H. (2008). A Generic Asset Management Framework for Optimising Maintenance Investment Decision. *Production Planning and Control*

Yeganeh, Naiem K, Shazia Sadiq, and Mohamed A Sharaf. (2014). A Framework for Data Quality Aware Query Systems. Information Systems 46: 24–44.

Wang, R.Y., and Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems,* 12 (4): 5–33.