A Hybrid Machine Learning Approach for Predictive Maintenance in Smart Factories of the Future

Sangje Cho[1(✉)], Gökan May[1], Ioannis Tourkogiorgis[1], Roberto Perez[2], Oscar Lazaro[3], Borja de la Maza[4], and Dimitris Kiritsis[1]

[1] EPFL, ICT for Sustainable Manufacturing, EPFL SCI-STI-DK, Station 9, 1015 Lausanne, Switzerland
Sangje.cho@epfl.ch
[2] GF Machining Solutions, Meyrin, Switzerland
[3] Innovalia, Bilbao, Spain
[4] TRIMEK, Zuia, Álava, Spain

Abstract. Advanced technologies based on Internet of Things (IOT) are blazing a trail to effective and efficient management of an overall plant. In this context, manufacturing companies require an innovative strategy to survive in a competitive business environment, utilizing those technologies. Guided by these requirements, the so-called predictive maintenance is of paramount importance and offers a significant potential for innovation to overcome the limitations of traditional maintenance policies. However, real shop-floors often have obstacles in providing insights to facilitate the effective management of assets in smart factories. Even if a significant amount of machine and process data is available, one of the common problems of these data is the lack of annotations describing the machine status or maintenance history. For this reason, companies have limited options to analyse manufacturing data, despite the capability of advanced machine learning techniques in supporting the identification of failure symptoms in order to optimize scheduling of maintenance operations. Moreover, each machine generates highly heterogeneous data, making it difficult to integrate all the information to provide data-driven decision support for predictive maintenance. Inspired by these challenges, this research provides a hybrid machine learning approach combining unsupervised learning and semi-supervised learning. The approach and result in this article are based on the development and implementation in a large collaborative EU-funded H2020 research project entitled BOOST 4.0 i.e. Big Data Value Spaces for COmpetitiveness of European COnnected Smart FacTories.

Keywords: Industry 4.0 · Predictive maintenance · Machine learning
Big data · Asset management · Smart factories · Sustainable manufacturing

I. Moon et al. (Eds.): APMS 2018, IFIP AICT 536, pp. 311–317, 2018.
https://doi.org/10.1007/978-3-319-99707-0_39

1 Introduction

As of today, modern industries require efficiency and convenience for management of the entire Product Life Cycle(PLC), in order to overcome intensely competitive business environment. Advances of IoT opens an efficient way to innovative predictive maintenance strategies in smart manufacturing environments, and these advanced technologies generate industrial big data. To exploit this big data, trend-oriented predictive maintenance tasks are carried out based on actual condition of machinery to avoid occurrence of failures, since advanced enabling technologies and wireless technologies open innovative capability to monitor very details of machines' status and behaviour. In that regard, even if predictive maintenance allows increasing business values and smart machining services, big data from various technologies should be effectively managed for interoperability via standards for merging and transformation of data. According to Lee, a successful shift toward more intelligent machines can be addressed considering five distinct issues as follows: Manager and Operator Interaction, Machine Feet, Product and Process Quality, Big Data and Cloud, and Sensor and Controller Network [1]. To address these issues, predictive maintenance pilots in the EU-funded H2020 research project entitled BOOST 4.0 [2] provides product innovation through a data-driven approach. These pilots are elaborated to integrate digital platforms and industrial things so as to foster collaboration considering the features of Industry 4.0 including (i) horizontal integration through value networks to facilitate inter-corporation collaboration, (ii) vertical integration of hierarchical subsystems inside a factory to create flexible and reconfigurable manufacturing system, and (iii) end-to-end engineering integration across the entire value chain to support product customization [3].

Meanwhile, data analytics through advanced machine learning techniques has been improved with the development of strong hardware and useful algorithms, supporting engineers to find trends and symptoms of failures in order to carry out maintenance tasks optimally. However, types and formats of data vary significantly depending on data sources. Most companies do not have the competence for management of such big data, and often record data without tags describing machine status and/or maintenance history. These constraints limit the application of machine learning algorithms and thus supervised learning and semi-supervised learning cannot be performed for data analytics. For this reason, this research provides a hybrid machine learning approach combining unsupervised learning and semi-supervised learning based on the development and implementation in the BOOST 4.0 project.

2 A Main Architecture for the Predictive Maintenance Pilot

Thanks to recent scientific and technological developments, most industrial practices try to employ a predictive maintenance policy instead of conventional Maintenance (i.e. corrective and/or preventive maintenance). According to Sullivan et al., independent surveys indicate that this predictive maintenance policy can lead to high return on investment, reduction in maintenance cost, elimination of breakdown, reduction in downtime and increase in production [4] since conventional maintenance policies incur low reliability of machines or needless maintenance tasks. However, the main constraints for application of predictive maintenance in BOOST 4.0 can be summarized as no maintenance history and heterogeneous data. Inspired by these constraints, the predictive maintenance pilot in BOOST 4.0 deals with a business case as follows:

- The target product is a milling machine in the shop floor
- Products produced by a milling machine can be measured by a Coordinate Measuring Machine (CMM) in the same shop floor
- Providers of milling machines and CMM machines are different, and machine data is collected by each machine provider
- The results of measurement are useful for the milling machine, but currently there is no intersection between two kinds of data
- The Data formats of the milling machine and the CMM machine are heterogeneous
- Milling machine data does not have maintenance indicator.

This pilot study consists of the following steps: (i) the acquisition and storage of data, (ii) data analytics on operating data and monitoring, (iii) continuous evaluation and prediction of the health status of the equipment as cyber-physical system (i.e. transform descriptors extracted previously in relevant behaviour models, allowing to represent the ways of functioning of the machine and the evolution of the equipment condition over time for detection and prognosis of failures), (iv) decision-making support by considering the context of use of the equipment.

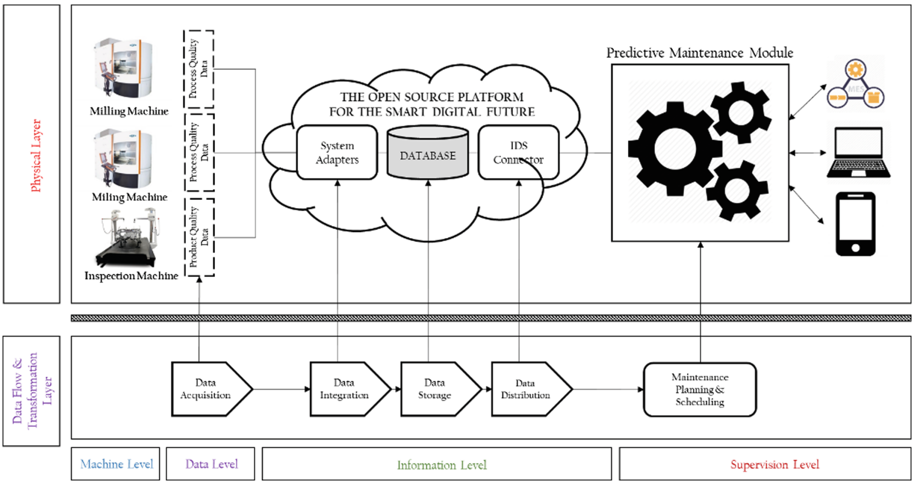

Fig. 1. An architecture for the predictive maintenance pilot

To address the concern of the acquisition and storage of data, application of the open source platform for the smart digital future considers data management (See Fig. 1). This open platform has the capacity of context management to merge heterogeneous data from milling machines and CMM machines. A system adapter in the open source platform will read all the data from milling and CMM machines.

Tracking ID of a product produced by milling machine, it will recognize relevant measurement data. Afterwards, it will update attribute values of defined entities which have the context of the shop floor. Context data will be managed by Context Blocker within IDS connectors. Context data from the shop floor will be accessible through IDS connect which is a back-end processor communicating with the milling machine predictive maintenance system. Depending on data security and requirements, IDS connector will send relevant data to the predictive maintenance system. The Predictive maintenance system will exploit distributed data for maintenance planning & scheduling. This system is in charge of data analytics on operating data and monitoring, evaluation and prediction of the health status, and decision-making support. On the other hand, the open source platform allows exploitation of a specific part of milling machine data for the CMM precision management system which is a part of an operation management pilot. In this context, this study focuses on the predictive maintenance application.

3 A Hybrid Approach of the Predictive Maintenance Pilot

This chapter describes the details of the predictive maintenance system for the pilot. As mentioned above, milling machine data have no maintenance indicator of any events. Therefore, available data limits application of machine learning algorithms and thus supervised learning and semi-supervised learning are not available for data analytics. For this reason, this research provides a hybrid machine learning approach combining unsupervised learning and semi-supervised learning (Fig. 2).

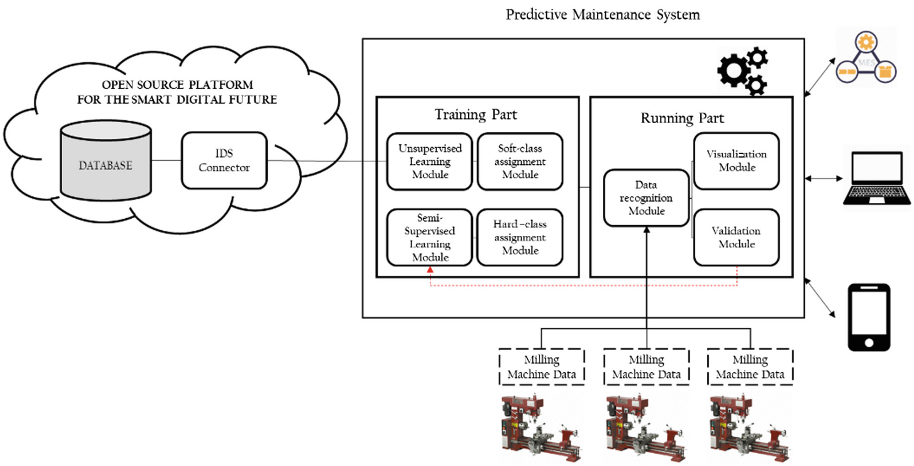

Fig. 2. A hybrid approach for the predictive maintenance pilot

The hybrid approach is comprised of two parts; a training part and a running part. The main role of the training part is to provide a probabilistic model to identify the status of machines. As a hybrid approach, this part has unsupervised learning module linked to soft-class assignment and semi-supervised learning module linked to hard-class assignment. Soft-class assignment initiates all the classes as inputs of an unsupervised learning module to overcome no maintenance indicator whereas hard-class assignment provides consistency of classes.

Based on the assignment modules, new data from milling machines could be used to identify if the machine status is normal or not. To provide a detailed explanation of procedures of data analytics, roles of each module are described below.

Receiving unlabelled training data sets (D_{1u}) which means they have no maintenance indicators from the open source platform, the training part will classify all the data set through an unsupervised learning module. Where the class indexes are k and $\pi_k = N_k/N$ (N_k: a number Dataset of class k, N is a total number of datasets), $\pi_k, \mu_k, \sum_k$ for each class k and $\mu_{D_{1u}}, \sum_{D_{1u}}$ are estimated. And then, where $x_n \in D_{1U}$, each of $\pi_k, \mu_k, \sum_k$ will be updated through the Expectation-Maximization (EM) [5] algorithm as follows:

$$\text{E step}\,:\,\gamma_{nk} = \frac{\pi_k N(x_n|\mu_k, \sum_k)}{\sum_{j=1}^{K} \pi_j N(x_n|\mu_j, \sum_j)} \tag{1}$$

$$\text{M step}\,:\,\mu_k^{new} = \frac{1}{N_k}\sum_{n=1}^{N} \gamma_{nk} x_n \tag{2}$$

$$\sum\nolimits_k^{new} = \frac{1}{N_k}\sum_{n=1}^{N} \gamma_{nk}(\mathrm{x}_n - \mu_k^{new})(\mathrm{x}_n - \mu_k^{new})^T \tag{3}$$

$$\pi_k = \frac{N_k}{N} \tag{4}$$

The log likelihood can be estimated as follows:

$$\ln p(\mathrm{X}|\mu, \sum, \pi) = \sum_{n=1}^{N} \ln\left\{\sum_{k=1}^{K} \pi_k N(x_n|\mu_k, \sum_k)\right\} \tag{5}$$

In the case of $P[\mu_k \in D_{1u}] \geq w$ (where, w is a predefined acceptance parameter), the k class will be soft-assigned to a normal class because normal data set is overwhelming in all the data sets. Otherwise, it will be soft-assigned to an abnormal class. When new data sets (D_2) from active milling machines are delivered to the running part, the running part will estimate $\arg\max_k P[x_t \in C_k]$ $(x_t \in D_2)$. If max $P[x_t \in C_k] \geq w$, d_2 will be assigned to Class k, and π_k, μ_k, $\sum_k$ will be updated. Otherwise, d_2 will be reported as unidentified data set. The visualization module will display all the graphs with significant values to deliver results of analytics (See Fig. 3). Maintenance engineers will validate these results through dashboard and will give maintenance annotations

(D_{IL}) to abnormal/unidentified data sets. These maintenance annotations will go to the semi-supervised learning module this module will initiate classes based on D_{IL}. To update $\pi_k, \mu_k, \sum_k$ of each class k, where $x_n \in D_{1U}(D_1 = D_{1U} \cup D_{1L})$, E step of EM algorithm is as follows:

$$\text{E step} : \gamma_{nk} = \begin{cases} \dfrac{\pi_k N(x_n|\mu_k, \sum_k)}{\sum\limits_{j=1}^{K} \pi_j N(x_n|\mu_j, \sum_j)} & \text{if } d_n \in D_{2U} \\ 1\ (k = y(n)) & \text{if } d_n \in D_{2L} \end{cases} \tag{6}$$

M step and estimation of the log likelihood can be estimated following Eqs. (2), (3), (4) and (5).

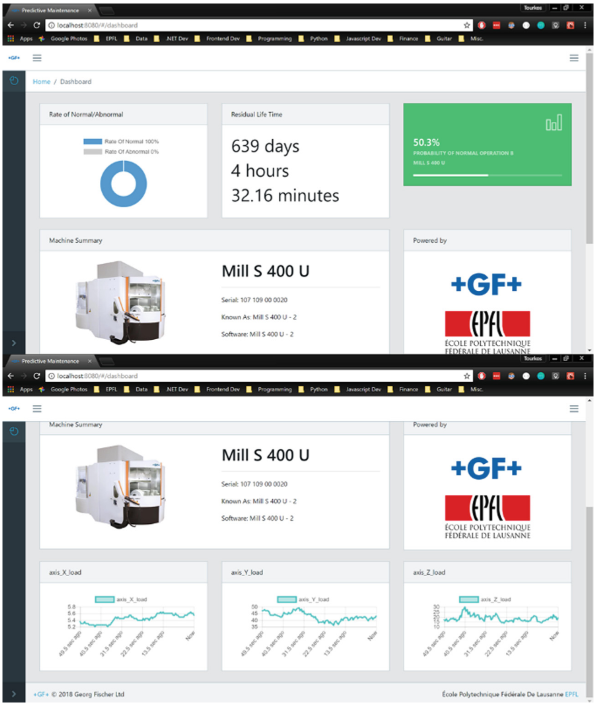

Fig. 3. A dashboard of predictive maintenance system

Return values from the hybrid approach are presented in a dashboard (See Fig. 3). Considering how to show results of the proposed approach, the dashboard is comprised of three kinds of aspects, i.e. a summary of the machines, machines` KPI, and machine details. A summary of the machines shows a graph representing a rate of available machines, and abnormal machines with probability of failures. Moreover, machines` KPI indicates meaningful information such as failure probability, failure events, aging, warning events, and so on. Machine details display status of each machine. The hybrid approach allows extension of visualisation depending on validation process.

The resulting predictive maintenance system will be demonstrated in the milling machine scenario, and then, its scope will be extended with illustration of all the predictive maintenance pilots in BOOST 4.0 for a wide range of application.

4 Discussion and Concluding Remarks

The main purpose of this research was to provide a novel predictive maintenance approach for the predictive maintenance pilot of BOOST 4.0. This study addressed the problems caused by no maintenance annotations and heterogeneous data sources. To resolve the issue in an efficient way, this paper included an architecture exploiting the open source platform for the smart digital future for merging heterogeneous data from milling machines and CMM machines through context management. In addition, the hybrid predictive maintenance approach was proposed to overcome constraints of no maintenance annotations.

Accordingly, the implications for knowledge and practice could be summarized as follows: (i) the shop floor enables a high-value service for users of the equipment by avoiding downtimes through predictive knowledge. This service brings results of minimizing the total cost and offers optimization of material usage, and (ii) through increased equipment availability, as well as the manufacturing technology as a whole, will help customers to gain a competitive advantage where much unforeseen downtime reduces the profitable production time.

As for future work, the proposed approach will be implemented and validated on not only the milling machine case but also other predictive maintenance pilots within tasks of BOOST 4.0, demonstrating its capacity and potential to support maintenance engineers and machine operators.

Acknowledgements. This work has been carried out in the framework of BOOST 4.0 Project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 780732.

References

1. Lee, J., Kao, H.A., Yang, S.: Service innovation and smart analytics for industry 4.0 and big data environment. Procedia Cirp **16**, 3–8 (2014)
2. BOOST4.0 Project. http://boost40.eu/. Accessed 06 Apr 2018
3. Wang, S., Wan, J., Li, D., Zhang, C.: Implementing smart factory of industrie 4.0: an outlook. Int. J. Distrib. Sens. Netw. **12**(1), 3159805 (2016)
4. Sullivan, G., Pugh, R., Melendez, A.P., Hunt, W.D.: Operations & Maintenance Best Practices-A Guide to Achieving Operational Efficiency (Release 3) (No. PNNL-19634). Pacific Northwest National Laboratory (PNNL), Richland (2010)
5. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)