DOI: https://doi.org/10.14201/0AQ03151124

A MACHINE LEARNING BASED FRAMEWORK FOR PDM

UN MARCO BASADO EN EL APRENDIZAJE AUTOMÁTICO PARA PDM

Jorge Meira[1], Luis Rodrigues[1], Marta Fernandes[1], Jonas Queiroz[2], Paulo Leitão[2], Carlos Ramos[1], and Goreti Marreiros[1]

[1] *GECAD – Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, Institute of Engineering, Polytechnic of Porto, 4200-072 Porto, Portugal*

{janme, 1180120, mmdaf, csr, mgt}@isep.ipp.pt

[2] *Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal*

{jqueiroz, pleitao}@ipb.pt

Abstract: The need for adaptation led the industry to evolve into a new revolution, where connectivity, amount of data, new devices, stock reduction, personalization and production control gave rise to Industry 4.0. Predictive maintenance is based on historical data, models and knowledge of the domain in order to predict trends, patterns of behavior and correlations by statistical models or Machine Learning to predict pending failures in advance. This paper presents a review of most applied machine learning techniques, comparing different authors' approaches used in predictive maintenance. Also, a conceptual machine learning framework is proposed to tackle various predictive maintenance challenges such as failure forecast, anomaly detection and Remaining Useful Life prediction.

Keywords: predictive maintenance; industrial maintenance; machine learning.

Resumen: La necesidad de adaptación llevó a la industria a evolucionar hacia una nueva revolución, donde la conectividad, la cantidad de datos, los nuevos dispositivos, la reducción del stock, la personalización y el control de la producción dieron lugar a la Industria 4.0. El mantenimiento predictivo se basa en datos históricos, modelos y conocimiento de el dominio con el fin de predecir tendencias, patrones de comportamiento y correlaciones mediante modelos estadísticos o aprendizaje automático para predecir con antelación los fallos pendientes. Este artículo presenta un estudio de las técnicas de aprendizaje automático más aplicadas, comparando los enfoques de diferentes autores utilizados en el mantenimiento predictivo. Además, se propone un marco conceptual de aprendizaje automático para abordar varios desafíos del mantenimiento predictivo, como la previsión de fallos, la detección de anomalías y la predicción de la vida útil restante.

Palabras Clave: mantenimiento predictivo, mantenimiento industrial, aprendizaje automático.

1 Introduction

The Industry 4.0 is the emergence of the «smart» factory, which means smart grids, mobility, flexibility of industrial operations and its interoperability, integration with customers and suppliers and the adoption of innovative business models [21]. The 4th industrial revolution focuses mainly on creating a digital representation of physical processes to gain better insights into what is happening in physical processes [10].

In addition, in the last decade the industry has adopted information and communication technologies in most of its activities, but especially in logistics and production operations. This evolution together with the Industrial Internet of Things (IIoT) has changed the way systems can interact, monitor, control and administer. Therefore, facilitating the integration of processes and systems between sectors and technologies contributes to better communication and cooperation with each other in a new intelligent way, revolutionizing production, logistics and resource planning more effectively and economically [5, 28].

According to Vanson Bourne's 2017 global survey, sponsored by ServiceMax, GE Digital, a provider of software and industrial services, identified that 82% of companies had at least one unplanned downtime in the three years prior to the study, lasting average of four hours. These outages cost about 260,000 dollars an hour across all businesses, with two episodes of downtime lasting 4 hours each equating to more than 2 million dollars. The cost, causes and repercussions of unplanned downtime are driving investments in digital tools, such as Machine Learning, mobility tools, IoT platform, Digital Twin [33, 32].

Machine downtime has a direct impact on production costs and is directly related to companies' ability to be competitive in terms of cost, quality and performance.

To this end, data is the key to this new generation of information that can anticipate or collaborate in making predictive decisions. Predictive maintenance (PdM) is based on the early detection of equipment problems, leading to maintenance being performed based on the actual condition of the machines. In this way, the repair or replacement of components is only carried out after detecting a certain level of deterioration, instead of being carried out after the failure occurs or at a predefined time. By preventing serious failures, predictive maintenance reduces unexpected failures and maximizes the mean time between failures (MTBF), in addition to reducing work accidents and their severity. In this way, it allows a reduction in the average time of repair (MTTR) and extends the useful life of the equipment. All of this results in increased revenue, lower maintenance and production costs, transformed into a competitive advantage for the company [21, 22, 26].

This work aims to review the most used machine learning techniques in *PdM*. We also present a conceptual framework for machine learning models used for forecasting and detecting failures and predict the remaining useful life of industrial equipment.

2 Machine Learning in predictive maintenance

In recent years, the fourth industrial revolution has attracted worldwide attention leading to a transformation of traditional production into factories equipped with intelligent sensors where technology is ubiquitous. Leading to produce huge amounts of data. One application that can benefit of this large amount of data is predictive maintenance, which consists of using historical data in predictive algorithms to identify trends in order to detect when the equipment will need repair, maintenance or replacement. This approach allows the industry to predict the degradation of machine performance and autonomously manage and optimize service needs on the equipment. Thus, predictive maintenance in production environments brings several benefits that are extremely strategic [18].

In this way, Machine Learning is the main asset capable of detecting and predicting failures, such as estimating the remaining life of equipment [25], in addition to being used to diagnose failures [1]. Second, Machine Learning allows to be adjusted to new changes in the factory plant, providing stakeholders with visualization of the results of the changes made. Finally, Machine Learning algorithms can identify which variables are important in the performance, deterioration and life cycle of the equipment, which is really powerful to support decision making. In this way, it can lead to major cost reductions, greater predictability and availability of systems.

Due to all these aspects, in the last years machine learning has achieved a great growth in the publications of studies related to the application of techniques for predictive maintenance in the most varied sectors. At the same time, several researches and literary reviews appear with explicit methodologies to classify and present the studies carried out [37, 36, 8].

ML techniques have been increasingly applied and have shown improved performance over conventional approaches. In practice, the application of ML techniques is not easy due to the lack of efficient procedures for obtaining training data and specific knowledge necessary to train the models. However, a factor worth mentioning is that most articles use real data instead of synthetic data, thus bringing the results closer to the real application [8].

In an overview, the techniques most frequently used for PdM are Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machines (SVM) and K means with a focus on applications of fault diagnosis and RUL prediction tasks. For this purpose, are used signals that: include acoustic emission, electrical signature parameters (current and voltage), temperature, pressure, rotation speed and vibration [36, 8].

In the Table 2, several applications of Machine Learning in predictive maintenance over the past few years are presented, providing an overview of the existing literature.

2.1 Random Forest-based approaches

In Wu et al. 2016 [34], the authors obtained an excellent performance in indus- trial PdM using the parallel Random Forest (PRF) technique. In this experiment, they collected signals from three sensors (ie, cutting force, vibration and acoustic emission) to create a seven-channel data acquisition system in order to extract 4 resources from each channel. In this way they were able to predict tool wear with an accuracy of 99.20% through the use of 28 dimensions.

Following an approach based on Random Forest to generate predictive models dynamically, Canizo et al. 2017[7] proposes an improvement of the paper from Kusiak & Verma 2011[19], where wind turbines are monitored. These models are designed using status data (alarms activated and deactivated) and operational data on the performance of wind turbines. In this way, the authors obtained an accuracy of 82.04%, achieving an improvement of 5.54% compared to the previous work.

More recently, Ayvaz & Alpay 2021[4] carried out a study with several ML techniques in order to forecast and detect failures using data generated from IoT sensors in a production line. The results of the comparative evaluations of the ML algorithms indicated that RF models obtained the best score achieving a value of 0.982 for R2, followed by the XGBoost method with 0.979.

Table 1. Studies related to PdM using Accuracy for classification and MSE for regression

References	Techniques	Results	Data
(Wu et al. 2016)[34]	PRF	99.20%	DR
(Durbhaka & Selvaraj 2016)[12]	K-means	81.80%	DR
	K-NN	87.00%	
	SVM	78.80%	
	CRA	93.00%	

References	Techniques	Results	Data
(Canizo et al. 2017)[7]	RF	82.04%	DR
(Kanawaday & Sane 2017)[15]	NB	96.61%	DR
	SVM	95.52%	
	CART	94.46%	
	Deep Neural Network	98.69%	
(Aydin & Guldamlasioglu)[3]	LSTM	85.00%	DS
(Zhang et al. 2017)[35]	DNN	100%	DL
(Mathew et al. 2017)[24]	SVR proposto	0.5522	DS
	SVR padrão	0.7322	
(Eke et al. 2017)[13]	K-means	-	DR
(Amihai et al. 2018)[2]	RF	-	DR
(Kolokas et al. 2018)[16]	RF	99.25	DR
	NB-G	98.00%	
	NB-B	98.50%	
	MLP	99.10%	
(von Birgelen et al. 2018)[31]	SOM	-	DL
(De Benedetti et al. 2018)[11]	ANN	90%	DR
(Huuhtanen & Jung 2018)[14]	CNN	-	DR
(Lasisi & Attoh- Okine 2018)[20]	SVM	97.62%	DR
	RF	92.86%	
	LDA	90.48%	
(Uhlmann et al. 2018)[30]	K-means	-	DR
(Bruneo & Vita 2019)[6]	LSTM	99.88%	DL
	SVM	97.82%	
	DNN	98.57%	
(Kumar et al. 2019)[17]	HMM + Regressão Linear	0.7261	DR
	HMM + Regressão quadrática	pura 0.7561	

References	Techniques	Results	Data
(Wo Jae Lee et al. 2019)[23]	SVM	89%, 80%, 91%	DR
	RNN	97%, 89%, 93%	
	CNN (Time history data)	78%, 90%, 85%	
	CNN (Spectrum Data)	99%, 95%, 99%	
(Cheng et al. 2020)[9]	ANN	96.422%	DR
	SVM	96.547%	
(Ayvaz & Alpay 2021)[4]RF	RF	0.9821	DR
	XGBoost	0.9791	
	Gradient Boosting	0.7761	
	MLP Regressor	0.6751	
	SVR	0.3471	
	AdaBoost	0.3381	

1. The values presented refer to R2
2. The values presented refer to the root mean square error of reconstruction (RMSE)
DR – Real data
DL – Literature data
DS – Synthetic data

2.2 Approaches based on Neural Networks

Neural networks are one of the most common and applied ML algorithms, and have been proposed in several industrial applications, including soft sensing [29] and predictive control [27].

In these related projects, one of the selected articles is Aydin & Guldamlasioglu 2017 [3] that implemented LSTM networks to predict the current condition of an engine using Apache Spark's large-scale data processing framework. The data was obtained through sensors collecting temperature data, engine pressure, fuel and coolant bleeding. In this way, the authors were able to obtain an accuracy of 85.00% in the forecast of the estimated remaining service life (RUL) of the engine. Using a simpler ANN structure, Kolokas et al. 2018 [16] compared the use of an MLP with other ML algorithms for the detection of failures in equipment in the aluminum industry, using data from nine months of operation. In this experiment, the author managed to obtain an accuracy of 99.10% in the MLP, surpassing the accuracy obtained in the comparative algorithms of Gaussian Naive Bayes (NB-G) with 98.00% and Bernoulli Naive Bayes (NB-B) with 98.50%. On the other hand, the DT and RF algorithms used for comparison stood out slightly better than the neuronal network, achieving accuracy of 99.25% and 99.25%, respectively.

Using a data set from the literature provided by NASA, Bruneo & Vita 2019 [6] presented an ML approach based on LSTM to demonstrate that these structures can be considered viable techniques for the analysis of historical data in order to predict RUL. The authors, to prove their approach, compared the intended technique with other algorithms such as SVM and DNN (Deep Learning Network). In the experiments carried out, the LSTM managed to excel, obtaining the highest precision with 99.88%, thus surpassing the DNN with 98.57% and the SVM with 97.82%.

2.3 Support Vector Machine based approaches

SVM is another ML method widely used and known to perform classification and regression tasks, due to its high precision. In Mathew et al. 2017 [24], the authors used a type of SVM for regression purposes called Support Vector Regression (SVR). In this work, a modified regression kernel is proposed for forecasting problems in order to determine the remaining useful life (RUL). The tests are performed with a simulated set of time series, and have shown that the proposed SVR has surpassed the standard SVR model, improving the RMSE (mean square error of the reconstruction root) from 0.732 to 0.552.

Lasisi & Attoh-Okine 2018 [20] compared the technique of SVM, RF and Linear Discriminant Analysis (LDA) to detect defects in rail geometry. This study used real data from a first-class railroad in the USA. Of the models used, the SVM managed to stand out by ranking with a 97.62% accuracy rate, followed by the RF with 92.86% and finally, the LDA with 90.48%.

2.4 K-means based approaches

The k-means model is a popular clustering algorithm that uses an unsupervised strategy to determine a set of clusters. In Durbhaka & Selvaraj 2016

[12] they analyzed the behavior of wind turbines using vibration signal analysis. In this work, the kNN and SVM algorithms are compared with K-means to classify types of failures in wind turbines. In conjunction with the comparison, the authors proposed a collaborative recommendation approach (CRA) method to analyze the similarity of all ML algorithm results in predicting the replacement and correction of deteriorating turbines to avoid sudden breakdowns. Individually the ML models were able to obtain an accuracy of 81.80% for K means, 97.00% for kNN and 78.80% for SVM. However, the proposed CRA approach was able to obtain an accuracy of 93.00% when analyzing the similarity of all the results of the models.

2.5 Other approaches

In Kanawaday & Sane 2017[15] they proposed a two-phase approach to predict low-quality production cycles, in order to enable the necessary measures to be taken to avoid the low-quality cycle. The first phase of the proposed architecture was the use of ARIMA models to predict the values. of the parameters for the rest of the production cycle, later these values are supplied to the supervised models for the classification of the production cycles. In this study, four classification techniques were evaluated (Naive Bayes, Support Vector Machine, CART and Deep Neural Network), with DNN obtaining the best result achieving an accuracy of 98.69%, followed by NB with 96.61%. The rest, SVM and CART, achieved 95.52% and 94.46%, respectively. The authors concluded the study with the indication that machine learning techniques are a fundamental part of predictive maintenance.

3 Machine Learning Framework Architecture for PdM

We propose a conceptual framework to be developed as an API service in future work. This framework will allow to apply several machine learning models presented in the literature for different predictive maintenance approaches.

The component diagram represented in Figure 1 shows the architecture of what the framework needs to have in order to offer fault prediction and detection. In this way, the components called «Data Repository», «Data Extraction» and «Preprocessing» are specific components related to the storage and data processing used to feed the machine learning models. The «Data Segregation» component is responsible for dividing the data into two subsets in order to enable training and evaluation of models. Regarding the «Model Training» component, it has several subcomponents representing the different approaches to predictive maintenance that can be implemented. The ”RUL forecast” uses regression models in order to estimate the remaining useful life of the equipment. The ”Failure Prediction” and ”Failure Detection” components uses anomaly/outlier detection techniques for classification and forecasting techniques for regression. The «Model Repository» component stores the models created by the «Model Training» component with the respective evaluation produced by the «Model Evaluation» component. The component responsible for forecasting and failures, called the ”Failure Prediction Service” receives the data through HTTPS and uses the existing models in the ”Model Repository” to infer the arriving data.

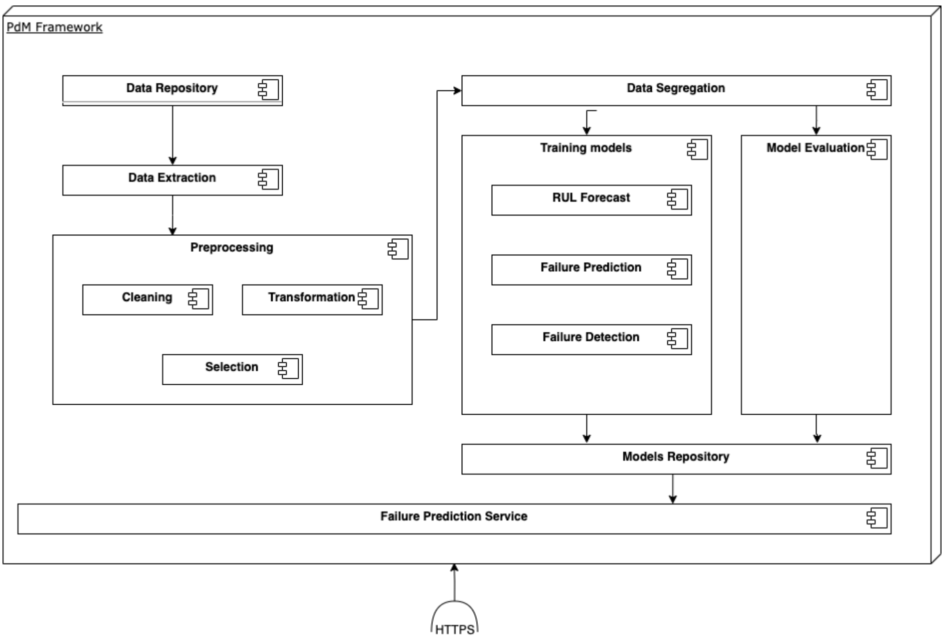

Fig. 1. Component diagram representing our PdM Machine Learning proposal framework.

4 Conclusions

The growing digitalization of companies marks the beginning of a new era for industrial maintenance. A new generation of smart sensors appeals to an increasing number of manufacturers who wish to improve their maintenance methods. Ongoing research into predictive maintenance techniques discussed in this paper promises to deliver technologies that may improve equipment reliability. Integrating our proposed framework for predictive maintenance will allow to predict equipment's RUL and failures before they occur, enabling plants to avoid unnecessary equipment replacement, save costs, and improve process safety, availability, and efficiency.

Acknowledgments

This work has been partially supported by the European Regional Development Fund (ERDF) through the Interreg Spain-Portugal V-A Program (POCTEP)under grant 0677DISRUPTIVE2E (Intensifying the activity of Digital Inno-vation Hubs within the PocTep region to boost the development of disruptiveand last generation ICTs through cross-border cooperation).

References

1. Adhikari, P., Rao, H., Buderath, M.: Machine learning based data driven diagnostics & prognostics framework for aircraft predictive maintenance. In: 10th International Symposium on NDT in Aerospace Dresden, Germany (2018)
2. Amihai, I., Gitzel, R., Kotriwala, A.M., Pareschi, D., Subbiah, S., Sosale, G.: An industrial case study using vibration data and machine learning to predict asset health. In: 2018 IEEE 20th Conference on Business Informatics (CBI). vol. 1, pp. 178–185. IEEE (2018)
3. Aydin, O., Guldamlasioglu, S.: Using lstm networks to predict engine condition on large scale data processing framework. In: 2017 4th International Conference on Electrical and Electronic Engineering (ICEEE). pp. 281–285. IEEE (2017)
4. Ayvaz, S., Alpay, K.: Predictive maintenance system for production lines in manufacturing: A machine learning approach using iot data in real-time. Expert Systems with Applications **173** (2021)

5. Boyes, H., Hallaq, B., Cunningham, J., Watson, T.: The industrial internet of things (iiot): An analysis framework. Computers in industry **101**, 1–12 (2018)
6. Bruneo, D., De Vita, F.: On the use of lstm networks for predictive maintenance in smart industries. In: 2019 IEEE International Conference on Smart Computing (SMARTCOMP). pp. 241–248. IEEE (2019)
7. Canizo, M., Onieva, E., Conde, A., Charramendieta, S., Trujillo, S.: Real-time predictive maintenance for wind turbines using big data frameworks. In: 2017 IEEE International Conference on Prognostics and Health Management (ICPHM). pp. 70–77. IEEE (2017)
8. Carvalho, T.P., Soares, F.A., Vita, R., Francisco, R.d.P., Basto, J.P., Alcalá, S.G.: A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering **137**, 106024 (2019)
9. Cheng, J.C., Chen, W., Chen, K., Wang, Q.: Data-driven predictive maintenance planning framework for mep components based on bim and iot using machine learning algorithms. Automation in Construction **112**, 103087 (2020)
10. Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., Barbosa, J.: Machine learning and reasoning for predictive maintenance in industry 4.0: Current status and challenges. Computers in Industry **123**, 103298 (2020)
11. De Benedetti, M., Leonardi, F., Messina, F., Santoro, C., Vasilakos, A.: Anomaly detection and predictive maintenance for photovoltaic systems. Neurocomputing **310**, 59–68 (2018)
12. Durbhaka, G.K., Selvaraj, B.: Predictive maintenance for wind turbine diagnostics using vibration signal analysis based on collaborative recommendation approach. In: 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). pp. 1839–1842. IEEE (2016)
13. Eke, S., Aka-Ngnui, T., Clerc, G., Fofana, I.: Characterization of the operating periods of a power transformer by clustering the dissolved gas data. In: 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED). pp. 298–303. IEEE (2017)
14. Huuhtanen, T., Jung, A.: Predictive maintenance of photovoltaic panels via deep learning. In: 2018 IEEE Data Science Workshop (DSW). pp. 66–70. IEEE (2018)
15. Kanawaday, A., Sane, A.: Machine learning for predictive maintenance of industrial machines using iot sensor data. In: 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS). pp. 87–90. IEEE (2017)
16. Kolokas, N., Vafeiadis, T., Ioannidis, D., Tzovaras, D.: Forecasting faults of industrial equipment using machine learning classifiers. In: 2018 Innovations in Intelligent Systems and Applications (INISTA). pp. 1–6. IEEE (2018)

17. Kumar, A., Chinnam, R.B., Tseng, F.: An hmm and polynomial regression based approach for remaining useful life and health state estimation of cutting tools. Computers & Industrial Engineering **128**, 1008–1014 (2019)
18. Kunst, R., Avila, L., Binotto, A., Pignaton, E., Bampi, S., Rochol, J.: Improving devices communication in industry 4.0 wireless networks. Engineering Applications of Artificial Intelligence **83**, 1–12 (2019)
19. Kusiak, A., Verma, A.: Prediction of status patterns of wind turbines: A data-mining approach. Journal of Solar Energy Engineering **133**(1) (2011)
20. Lasisi, A., Attoh-Okine, N.: Principal components analysis and track quality index: A machine learning approach. Transportation Research Part C: Emerging Technologies **91**, 230–248 (2018)
21. Lee, J., Bagheri, B., Kao, H.A.: Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics. In: International proceeding of int conference on industrial informatics (INDIN). pp. 1–6 (2014)
22. Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., Liao, H.: Intelligent prognostics tools and e-maintenance. Computers in industry **57**(6), 476–489 (2006)
23. Lee, W.J., Wu, H., Yun, H., Kim, H., Jun, M.B., Sutherland, J.W.: Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. Procedia Cirp **80**, 506–511 (2019)
24. Mathew, J., Luo, M., Pang, C.K.: Regression kernel for prognostics with support vector machines. In: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). pp. 1–5. IEEE (2017)
25. Schmidt, B., Wang, L.: Predictive maintenance of machine tool linear axes: A case from manufacturing industry. Procedia manufacturing **17**, 118–125 (2018)
26. Sezer, E., Romero, D., Guedea, F., Macchi, M., Emmanouilidis, C.: An industry 4.0-enabled low cost predictive maintenance approach for smes. In: 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). pp. 1–8. IEEE (2018)
27. Shin, J.H., Jun, H.B., Kim, J.G.: Dynamic control of intelligent parking guidance using neural network predictive control. Computers & Industrial Engineering **120**, 15–30 (2018)
28. Sisinni, E., Saifullah, A., Han, S., Jennehag, U., Gidlund, M.: Industrial internet of things: Challenges, opportunities, and directions. IEEE Transactions on Industrial Informatics **14**(11), 4724–4734 (2018)
29. Soares, S.G., Araújo, R.: An on-line weighted ensemble of regressor models to handle concept drifts. Engineering Applications of Artificial Intelligence **37**, 392– 406 (2015)

30. Uhlmann, E., Pontes, R.P., Geisert, C., Hohwieler, E.: Cluster identification of sensor data for predictive maintenance in a selective laser melting machine tool. Procedia manufacturing **24**, 60–65 (2018)
31. Von Birgelen, A., Buratti, D., Mager, J., Niggemann, O.: Self-organizing maps for anomaly localization and predictive maintenance in cyber-physical production systems. Procedia CIRP **72**, 480–485 (2018)
32. Williamson, J.: Unplanned downtime affecting 82% of businesses (2019)
33. Word, M.: Microsoft Word - After The Fall whitepaper. https://bit.ly/2TCuuOb (2017), [Online; accessed may. 29, 2021]
34. Wu, D., Jennings, C., Terpenny, J., Kumara, S.: Cloud-based machine learning for predictive analytics: Tool wear prediction in milling. In: 2016 IEEE International Conference on Big Data (Big Data). pp. 2062–2069. IEEE (2016)
35. Zhang, R., Peng, Z., Wu, L., Yao, B., Guan, Y.: Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence. Sensors **17**(3), 549 (2017)
36. Zhang, W., Yang, D., Wang, H.: Data-driven methods for predictive maintenance of industrial equipment: a survey. IEEE Systems Journal **13**(3), 2213–2227 (2019)
37. Zonta, T., da Costa, C.A., da Rosa Righi, R., de Lima, M.J., da Trindade, E.S., Li, G.P.: Predictive maintenance in the industry 4.0: A systematic literature review. Computers & Industrial Engineering p. 106889 (2020)