

# Data-Driven Predictive Maintenance

Subjects: Computer Science, Artificial Intelligence

Contributor: Bruno Miguel Veloso

Cyber-physical systems in Industry 4.0 are reforming conventional decision-making processes, mainly through integrating entities and functionalities via telecommunication systems and intelligent data processing approaches. This reformulation brings new challenges and increases complexity. Nevertheless, these advancements might provide new solutions for typical problems, such as system failures, and thus, for maintenance approaches. Predictive Maintenance (PdM) is a data-based approach that emerged as a prominent field of research among many existing maintenance approaches. We have three main categories in PdM: model-based prognosis, knowledge-based prognosis, and data-driven prognosis. Data-driven PdM strategies appeared with great prominence and importance both in industry and academia.

Keywords: condition-based maintenance ; predictive maintenance ; machine learning ; deep learning ; artificial intelligence ; railway industry

---

## 1. Introduction

Maintenance corresponds to the process that deals with equipment or system components to ensure their normal functioning under any circumstances. Over the years, several different maintenance approaches have been developed, each representing a different generation over time due to technological advances. Three main maintenance approaches can be classified as below<sup>[1]</sup>:

- **Corrective maintenance:** It means run-to-failure, which is the simplest and the oldest method. The idea is to act only after a machine or equipment fails. It would almost always lead to high (unexpected) downtime, besides having maintenance staff expenditure. This method usually generates a critical situation that will demand a great cost for companies.
- **Preventive maintenance:** It provides planning of regular replacement of components and/or equipment. Considering historical failure data and/or the data provided by the equipment manufacturer, *Mean Time To Failure* (MTTF) is calculated, which in turn is used by the maintenance team to propose a preventive action plan. Although this approach prevents unexpected shutdown, it usually needs additional costs and an increased unexploited lifetime.
- **Predictive Maintenance (PdM):** It needs direct monitoring of the mechanical condition and other parameters to determine the operating conditions over time. Indeed, due to technological advances, existing tools can process real-time data acquired from different equipment parts to predict any sign of failure.

In the last few years, many works have addressed data-driven Predictive Maintenance (PdM) by the use of Machine Learning (ML) and Deep Learning (DL) solutions, especially the latter. The monitoring and logging of industrial equipment events, like temporal behaviour and fault events — anomaly detection in time-series — can be obtained from records generated by sensors installed in different parts of an industrial plant. However, such progress is incipient because we still have many challenges, and the performance of applications depends on the appropriate choice of the best features and methods to capture the system behaviour.

## 2. Data-driven Predictive Maintenance

Predictive maintenance attempts to predict failures and avoid system shut down proactively, which differs from traditional maintenance techniques (e.g., corrective and preventive). Detecting and preventing failures in industries with high operational risk (e.g., the railway industry) is ultimately essential to improve not only the system efficiency (e.g., equipment utilisation) but also its effectiveness (e.g., the integrity of the environment and human safety). Industries seek to minimise the number of operational failures, minimise operating costs, and increase productivity, making maintenance management crucial. Consequently, planning and analysis strategies are necessary to assess the equipment's operating status and useful life.

Due to the complexity involved in an industrial process, several automated solutions have been developed to support decision making by performing future projections about equipment state using signal processing techniques. Modern transportation, for example, is highly dependent on these automated solutions to move cargo and passengers. The global increase in production and logistics needs higher use of the railway industry. Thus, common damages will occur in the overall structure and components due to factors such as weather and degradation. These could potentially lead to accidents of different proportions, which can even cause fatalities<sup>[2]</sup>.

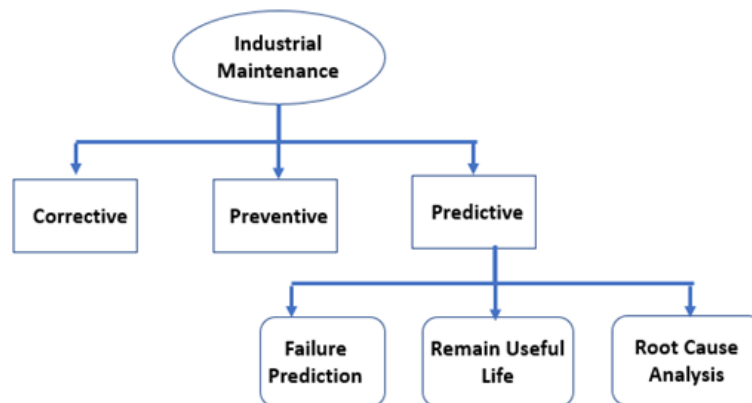


Figure 1 - Classification of automatic industrial maintenance approaches.

Over the years, PdM practices have been developed from several perspectives: **Failure prediction**, to predict equipment failure over time interval; **Remain Useful Life** estimation, to estimate the remaining useful lifetime of equipment; and **Root Cause Analysis**, identification of the causes of the failure. These two perspectives are illustrated in Figure 1 and are detailed next.

- **Failure Prediction** is the most generic and direct perspective for the Predictive Maintenance practices for which the main goal is to predict the approximate moment where some failure could occur.
- **Remain useful life** is strongly related to prognostics, which provides the amount of time equipment will be operational before it requires any repair or replacement. Prognostic is directly related to Mean Time to Failure (MTTF) estimation and the likelihood of system failure. It can be regarded as a forecasting process given the current machine conditions and its historical record
- **Root Cause analysis** is related to diagnosis. The identification of the most probable causes of the failure.

In the past decade, many works addressed data-driven PdM using ML/DL approaches, but mainly the latter. The monitoring and logging of industrial equipment events, like temporal behaviour and fault events, can be obtained from data and records generated by various sensors installed on the equipment. Specifically, sensors can be implemented to PdM to decrease the failure rate and enhance the system reliability<sup>[3]</sup>. Such sensors can monitor and generate alerts for equipment with the need for attention. The progressive development of industrial (wireless) sensor networks and emerging technologies, e.g., IoT<sup>[3][4][5]</sup>, brings about generating a massive amount of data with scale and higher reliability. In this perspective, ML/DL algorithms are particularly relevant to create advanced mining methods for the PdM.

Recent advances in sensors and computing technology have given rise to PdM, which maximises system utilisation, minimises maintenance costs, and improves safety, reliability, and efficiency. In particular, with recent technology advances in cloud storage, communication, and sensing, for the railway industry, we can monitor any part of the system more precisely and in real-time. Thus, more complex solutions are necessary to analyse data with more scalability, precision, and efficiency.

Research in PdM practices for the railway industry is progressively receiving more attention from the industry and academia. A recent literature review regarding Big Data Analytics in the railway industry can be found in<sup>[6]</sup>, where the level and the types of big data models are reviewed and summarised for operations, maintenance, and safety applications. Most of the works focus on solutions that assess the infrastructure health state like railway points (switches) and interlocking systems. Although, in the case of trains, there are many other challenges related to internal conditions, like the general functioning of wagons (e.g., wheels, air compressed units, brakes), and external conditions, like weather, geographical position, in addition to other variables.

The dynamic context of the railway system is exceptionally challenging and these areas, by themselves, require the study of many combinations of analysis. In this sense, we define a taxonomy specific to the context of the railway industry. Differently, from<sup>[6]</sup>, our taxonomy classifies the related works in three areas: infrastructure, scheduling policies, and

vehicles. We also organise the works based on the type of data analysis method used to address PdM practices. We also employed a classification grounded on ML and DL algorithms, following the work in<sup>[4]</sup>. In practice, PdM needs a timely decision-making process that requires models to process data and adjust themselves on time.

### 3. Conclusion and Future Research Directions

Although the data-driven PdM is gaining more research attention, specifically in the past few years, the number of works specifically designed for the railway industry is quite limited.

Considering the research trends reviewed, we can observe some significant gaps in the literature. As noted, only a few works have faced the problem of using data as time series. Sensors typically gather data in the time-series format. Thus, we can envision this scenario as a task of anomaly detection in time series. Anomaly detection is the problem that identifies specific patterns or events in data that are pretty different from the rest and can arise in the data for many reasons.

In manufacturing systems, reducing downtime is critical, and anomaly detection enables PdM for downtime reduction. Recent works have addressed anomaly detection for PdM supported by learning strategies on sequential data <sup>[1][2][3][4][5][6][7]</sup>. In the last few years, we can find several papers published approaching Anomaly Detection with Time-Series data applied to the most different domains <sup>[8][9][10][11][12][13][14][15][16][17][18][19][20][21][22][23][24][25][26][27][28][29][30][31][32][33]</sup>.

The major challenge is dealing with models with a high volume of time series in real-time to perform anomaly prediction. Moreover, currently used metrics are not feasible in this context. It will be indispensable to look for new alternatives that can efficiently evaluate models.

The other essential line of action is to look for different DL algorithms and architectures like RNN, GAN, TL and RL. Recent works have proposed approaches based on DL to resolve anomaly detection in time series <sup>[34][18][20][32][35][36]</sup>. Nevertheless, new proposals in this research line will be necessary.

The last challenge would be to achieve the desired synergy between ML/DL methods and RCA by gaining automatic reasoning power to explain causality.

---

### References

1. Ribeiro, R.P.; Pereira, P.M.; Gama, J. Sequential anomalies: A study in the Railway Industry. *Mach. Learn.* 2016, 105, 127–153.
2. Baptista, M.; Sankararaman, S.; de Medeiros, I.P.; Nascimento, C.L.; Prendinger, H.; Henriques, E.M.P. Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Comput. Ind. Eng.* 2018, 115, 41–53.
3. Rabatel, J.; Bringay, S.; Poncelet, P. Anomaly detection in monitoring sensor data for preventive maintenance. *Expert Syst. Appl.* 2011, 38, 7003–7015.
4. Liu, J.; Guo, J.; Orlik, P.V.; Shibata, M.; Nakahara, D.; Mii, S.; Takác, M. Anomaly Detection in Manufacturing Systems Using Structured Neural Networks. In *Proceedings of the 2018 13th World Congress on Intelligent Control and Automation (WCICA)*, Changsha, China, 4–8 July 2018; pp. 175–180.
5. Zare, S. Fault Detection and Diagnosis of Electric Drives Using Intelligent Machine Learning Approaches. Master's Thesis, University of Windsor, Windsor, ON, Canada, 2018.
6. Yolacan, E.N. Learning from Sequential Data for Anomaly Detection. Master's Thesis, Northeastern University, Boston, MA, USA, 2014.
7. Andrade, T.; Gama, J.; Ribeiro, R.P.; Sousa, W.; Carvalho, A. Anomaly Detection in Sequential Data: Principles and Case Studies. In *Wiley Encyclopedia of Electrical and Electronics Engineering*; American Cancer Society: Atlanta, GA, USA, 2019; pp. 1–14.
8. Zhang, W.; Yang, D.; Wang, H. Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Syst. J.* 2019, 13, 2213–2227.
9. Toledano, M.; Cohen, I.; Ben-Simhon, Y.; Tadeski, I. Real-time anomaly detection system for time series at scale. In *Proceedings of the KDD 2017: Workshop on Anomaly Detection in Finance*; Anandakrishnan, A., Kumar, S., Statnikov, A., Faruquie, T., Xu, D., Eds.; PMLR: Stockholm, Sweden, 2018; Volume 71, pp. 56–65.

10. Lu, Y.; Kumar, J.; Collier, N.; Krishna, B.; Langston, M.A. Detecting Outliers in Streaming Time Series Data from ARM Distributed Sensors. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 779–786.
11. Calikus, E.; Nowaczyk, S.; Sant'Anna, A.P.; Dikmen, O. No Free Lunch But A Cheaper Supper: A General Framework for Streaming Anomaly Detection. *arXiv* 2019, arXiv:1909.06927.
12. Malhotra, P.; Vig, L.; Shroff, G.; Agarwal, P. Long Short Term Memory Networks for Anomaly Detection in Time Series. In Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Bruges, Belgium, 22–24 April 2015.
13. Thi, N.N.; Cao, V.L.; Le-Khac, N.A. One-Class Collective Anomaly Detection Based on LSTM-RNNs. In Transactions on Large-Scale Data-and Knowledge-Centered Systems XXXVI; Springer: Berlin/Heidelberg, Germany, 2017; Volume 36, pp. 73–85.
14. Gamboa, J.C.B. Deep Learning for Time-Series Analysis. *arXiv* 2017, arXiv:1701.01887.
15. Shipmon, D.T.; Gurevitch, J.M.; Piselli, P.M.; Edwards, S.T. Time Series Anomaly Detection: Detection of anomalous drops with limited features and sparse examples in noisy highly periodic data. *arXiv* 2017, arXiv:1708.03665.
16. Giannoni, F.; Mancini, M.; Marinelli, F. Anomaly Detection Models for IoT Time Series Data. *arXiv* 2018, arXiv:1812.00890.
17. Zhang, C.; Song, D.; Chen, Y.; Feng, X.; Lumezanu, C.; Cheng, W.; Ni, J.; Zong, B.; Chen, H.; Chawla, N.V. A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data. In Proceedings of the AAAI, New Orleans, LA, USA, 2–7 February 2018.
18. Pereira, J.; Silveira, M. Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Auto encoders with Attention. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 1275–1282.
19. Lamrini, B.; Gjini, A.; Daudin, S.; Armando, F.; Pratmarty, P.; Travé-Massuyès, L. Anomaly Detection Using Similarity-based One-Class SVM for Network Traffic Characterization. In Proceedings of the 29th International Workshop on Principles of Diagnosis, Warsaw, Poland, 27–30 August 2018.
20. Maya, S.; Ueno, K.; Nishikawa, T. dLSTM: A new approach for anomaly detection using deep learning with delayed prediction. *Int. J. Data Sci. Anal.* 2019, 8, 137–164.
21. Lindemann, B.; Fesenmayr, F.; Jazdi, N.; Weyrich, M. Anomaly detection in discrete manufacturing using self-learning approaches. *Procedia CIRP* 2019, 79, 313–318.
22. Su, Y.; Zhao, Y.; Niu, C.; Liu, R.; Sun, W.; Pei, D. Robust Anomaly Detection for Multivariate Time Series through Stochastic Recurrent Neural Network. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; KDD '19. ACM: New York, NY, USA, 2019; pp. 2828–2837.
23. Nguyen, L.H.; Goulet, J.A. Real-time anomaly detection with Bayesian dynamic linear models. *Struct. Control. Health Monit.* 2019, 26, e2404.
24. Feremans, L.; Vercruyssen, V.; Meert, W.; Cule, B.; Goethals, B. A framework for pattern mining and anomaly detection in multi-dimensional time series and event logs. In Proceedings of the International Workshop on New Frontiers in Mining Complex Patterns, held in Conjunction with ECML-PKDD 2019, Würzburg, Germany, 16–20 September 2019.
25. Feremans, L.; Vercruyssen, V.; Cule, B.; Meert, W.; Goethals, B. Pattern-Based Anomaly Detection in Mixed-Type Time Series. In Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, Würzburg, Germany, 16–20 September 2019.
26. Munir, M.; Siddiqui, S.; Chattha, M.; Dengel, A.; Ahmed, S. FuseAD: Unsupervised Anomaly Detection in Streaming Sensors Data by Fusing Statistical and Deep Learning Models. *Sensors* 2019, 19, 2451.
27. Munir, M.; Siddiqui, S.; Dengel, A.; Ahmed, S. DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. *IEEE Access* 2018, 7, 1991–2005.
28. Zhang, X.; Lin, Q.; Xu, Y.; Qin, S.; Zhang, H.; Qiao, B.; Dang, Y.; Yang, X.; Cheng, Q.; Chintalapati, M.; et al. Cross-dataset Time Series Anomaly Detection for Cloud Systems. In Proceedings of the 2019 USENIX Conference on Usenix Annual Technical Conference, USENIX ATC '19, Renton, WA, USA, 10–12 July 2019; pp. 1063–1076.
29. Elsner, D.; Khosroshahi, P.A.; MacCormack, A.D.; Lagerström, R. Multivariate Unsupervised Machine Learning for Anomaly Detection in Enterprise Applications. In Proceedings of the 52nd Hawaii International Conference on System Sciences, Maui, HI, USA, 8–11 January 2019.

30. Brandsæter, A.; Vanem, E.; Glad, I.K. Efficient on-line anomaly detection for ship systems in operation. *Expert Syst. Appl.* 2019, 121, 418–437.
31. Tran, L.; Fan, L.; Shahabi, C. Outlier Detection in Non-stationary Data Streams. In *Proceedings of the 31st International Conference on Scientific and Statistical Database Management, SSDBM '19*, Santa Cruz, CA, USA, 23–25 July 2019; ACM: New York, NY, USA, 2019; pp. 25–36.
32. Yeh, Y.C.; Hsu, C.Y. Application of Auto-Encoder for Time Series Classification with Class Imbalance. In *Proceedings of the Asia Pacific Industrial Engineering & Management Science Conference, APIEMS 2019*, Kanazawa, Japan, 2–5 December 2019; pp. 14–17.
33. Graß, A.; Beecks, C.; Soto, J.A.C. Unsupervised Anomaly Detection in Production Lines. *Machine Learning for Cyber Physical Systems*; Beyerer, J., Kühnert, C., Niggemann, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 18–25.
34. Fan, Y.; Nowaczyk, S.; Rögnvaldsson, T. Transfer learning for Remaining Useful Life Prediction Based on Consensus Self-Organizing Models. *arXiv* 2019, arXiv:1909.07053.
35. Vercruyssen, V.; Meert, W.; Davis, J. Transfer Learning for Time Series Anomaly Detection. In *Proceedings of the Workshop and Tutorial on Interactive Adaptive Learning Co-Located with European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2017)*, Skopje, Macedonia, 18–22 September 2017; pp. 27–36.
36. Oh, M.H.; Iyengar, G. Sequential Anomaly Detection Using Inverse Reinforcement Learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19*, Anchorage, AK, USA, 4–8 August 2019; ACM: New York, NY, USA, 2019; pp. 1480–1490.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/34588>