Data-driven based Fault Prognosis for Systems: a Review ⋆

José L. Guadiana Álvarez, **Luisa F. Montoya Herrera**, **Adriana Vargas Martínez**, **Jorge de J. Lozoya Santos**, **Ricardo A. Ramírez Mendoza**, **Ruben Morales Menéndez**

Tecnológico de Monterrey, School of Engineering and Sciences, Monterrey, NL 64849 Mexico
{A00828118, A00828117}@itesm.mx, {adriana.vargas.mtz, jorge.lozoya, ricardo.ramirez, rmm}@tec.mx

Abstract: Condition-based maintenance (CBM), is a maintenance strategy that is based on the continuous monitoring of systems. With Fault Prognosis, prediction such as at what time an element of a system will fail, or its Remaining Useful Life can be made. A review on Fault Prognosis and its most used algorithms to apply it, evaluating the obtained results is made. While Machine Learning is the preferred methodology, a hybrid algorithm using data-driven and knowledge-based methods can sometimes offer a better solution, depending on the system and the available data from it.

Keywords: Fault Detection and Diagnosis, Fault Prognosis, Condition Based Maintenance, Machine Learning.

1. INTRODUCTION

The demands for faster and more efficient manufacturing processes has increased exponentially in recent years, as conventional maintenance strategies are insufficient for major manufacturers to stay competitive against companies using emerging technologies. To meet these arising needs, the concept of *Condition-Based Maintenance* (*CBM*) was created, which focuses on scheduling maintenance activities by periodically or continuously monitoring specific indicators that are related to the health and performance of the maintained systems according to Ma et al. (2018). An important objective of *CBM* is to determine the optimal time for replacement or overhaul of a machine or a tool, Kim et al. (2012a). Since this represents minimizing production costs and maximizing the *Remaining Useful Lifetime* (*RUL*) of the systems. *CBM* must have the ability to not only identify the anomalies in a machine or an element, but also to classify and predict them to act efficiently to prevent breakdowns.

One of the key components of *CBM* is fault prognosis, which can be defined as the prediction of the *RUL* of components based on real-time available data. The main goal of prognosis is to assess how long the faulty components can continue working with acceptable performance, Kordestani et al. (2018). Given a reliable and accurate prognosis method, faults in critical components could be avoided by programming a maintenance window. This will ultimately minimize manufacturing costs caused by downtime, while maximizing the *RUL* of said components.

As shown in Figure 1, Advanced Anomaly Detection and Deep Learning are on the down-side of the ”Peak of Inflated Expectations” in the Hype Cycle for Data Science and Machine Learning graph from Gartner, Inc. This means that a lot of research about it is being made right now. Meanwhile, predictive analytics is already on the ”Slope of Enlightenment” with an expected time to real implementation of 2 to 5 years. All of theses topics are in the scope of Fault Prognosis and havea lot of research focus.

The rest of this paper is organized as follows: section 2 will discuss the state of the art. In section 3, an study of the different possible applications for these methods will be shown. Finally, section 4 will mention opportunity areas and future work to be done with these technologies in the automotive and aerospace industries.

2. STATE OF THE ART

Over the years, many different fault prognosis methods have been introduced. However, Sankavaram et al. (2009) states that they are all framed within the same three categories which are: model-based, knowledge-based, and

⋆ Special thanks to CONACYT and ITESM Campus Monterrey for financial and technical support of this investigation.

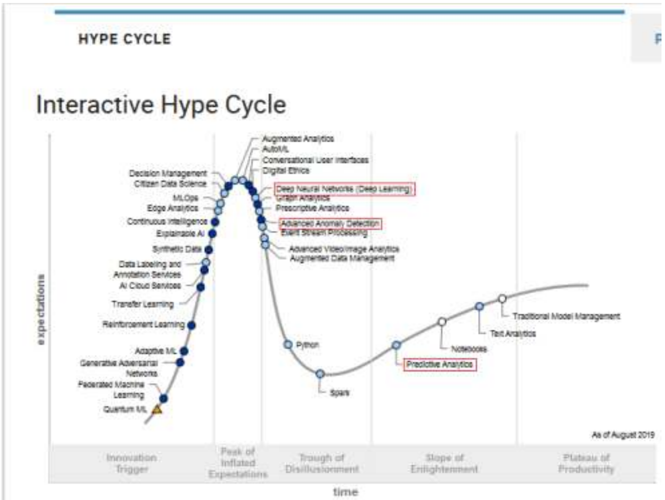

Fig. 1. Hype Cycle for Fault Prognosis and Deep Learning. Gartner, Inc Shubhangi Vashisth (2019)

data-driven methods. Model-based approaches require an accurate mathematical model, to be developed and use residuals as features, where residuals are the outcomes of consistency checks between the sensed measurements of a real system and the outputs of a mathematical model Vachtsevanos G. and Wu (2006). The knowledge-based approach uses symbolic representations of human knowledge to solve problems. However, this can be very difficult when dealing with complex systems. Some of the most known knowledge-based methods are Petri nets, flow charts of multiple signals, and *Bayesian Networks* (*BN*), which are largely used to fault diagnosis and reliability analysis, Luo et al. (2005). Data-driven approaches are used when system models are not available, unknown, or are too complex, but instead, system monitoring input and output data are available, Sankavaram et al. (2009).

Another approach that we could find in literature is a hybrid-based approach that results from combining two of the aforementioned. An example of this approach would be *Bayesian Neural Networks* (*BNN*) which use Bayesian learning to train a common *BNN*. By doing this, *BNN* incorporates a measure of uncertainty in the prediction which is missing from the current neural networks architectures, Shridhar (2019).

This paper will focus on data-driven methods, mainly in aerospace and automotive industries because of the complexity of these systems, but also in the medicine field which is and interesting area for develop and to take advantage of this approaches. Most common data-driven approach are artificial intelligence approaches which refers to techniques that fits into Machine Learning (*ML*), like Support Vector Machine (*SVM*) Kim et al. (2012b), Fuzzy Logic Ramasso and Gouriveau (2010) and Deep Learning (*DL*), Wu et al. (2018), Kiakojoori and Khorasani (2014). The most widely used category in recent years is (*ML*), which is a type of artificial intelligence whereby an algorithm or method will extract patterns out of data and it could be done using supervised learning (labeled examples as training data are given and then, predictions for all the unseen points are made) or unsupervised learning (only unlabeled training data are received and predictions are made for all the unseen points); using different *DL* techniques as Deep Learning Networks (*DLN*).

2.1 Deep Learning architectures

DLN methods could be seen as a cascade of many layers of processing units that combine the predictor features to approximate the target feature, in a similar way to which it is done in ANN, Diez Oliván (2017). ANN are a tool that has been developed in recent years to classify, detect, diagnose and predict in different applications of knowledge. It is a complex network formed by a large number of simple processing units through the way of connection, Hong (2011). The neural network receives as its input the set of input signals and gives an appropriate response (output signals), which are the solution to a specific task, Komyakov et al. (2016).

Some examples of DLN architectures are:

- *Feed Forward Neural Networks* (*FFNN*): They forward the input data unidirectionally from one layer to another.
- *Recurrent Neural Networks* (*RNN*): They forward the input data bidirectionally from one layer to another.
- *Convolutional Neural Networks* (*CNN*): It differs from other networks because it has convolutional, max pooling, and fully-connected layers.
- *Bayesian Neural Networks* (*BNN*): Is a *NN* with a prior distribution on its weights.

A comparison between these different structures can be seen in Table 1, as well as other commonly used algorithms for fault prognosis.

The deep learning architectures shown above have worked correctly in different areas, obtaining good results in most of their applications. The medicine field, which is an emerging area in the use of prognosis, have used *FFNN*, *CNN* (which work very well for small databases) and *RNN* Gómez-Ruiz et al. (2004), Lehman et al. (2015) in many investigations obtaining result with high accuracy. The feasibility and precision to work with each of these structures varies depending on the application and the characteristics of the data that are had but in general, with any structure good results are obtained.

2.2 Prognosis and health management in a process

Regardless of the type of network structure used to carry out the failure analysis and the estimation of the *RUL*

Table 1. Comparison of Machine Learning approaches for Fault Prognosis

Name	Advantages	Disadvantages	Precision	Feasibility
Feed-Forward Network (FNN)	Faster to train comparing to the other networks.	Because of its simpler structure, they do not handle well sequential and complex problems.	High	Low
Recurrent Neural Network (RNN)	Widely used for sequential problems, particularly the short-long term memory (LSTM).	Difficult training because of gradient vanishing and exploding issues.	High	High
Convolutional Neural Network (CNN)	Extract features and characteristics of a system. Mostly used in image classification and speech problems.	High computational costs and they require a bid training data set.	High	Medium
Bayesian Neural Network (BNN)	Considers the uncertainties for the predictions of future states. Because of their learning process, they are less susceptible to overfitting.	Complex learning algorithm as it uses Bayes Theorem.	High	High

of the system or element, it is necessary to define an appropriate methodology to be able to give the inputs that best fits in the network with the data.

Figure 2 shows how the process of *PHM* is made. The data collected from a process is normally used to make a correct prognosis and diagnosis and to be able to carry out a *CBM*. First it is necessary to monitor the signals by means of sensors in order to collect the necessary data to make the analysis, then the data and signals are preprocessed and a selection of the features that are going to be relevant in the analysis of the data is made, finally; by means of the different existing *ML* techniques, a diagnosis, prognosis and a *CBM* from the date could be made.

3. APPLICATIONS

Table 2 provides a summary of the main documented methods for fault prognosis.

Hu et al. (2016) have researched the field of prognosis oriented to the Lithium-ion battery health used in electric vehicles. They use a combination of sample entropy and Bayesian inference, a sparse Bayesian predictive modelling (SBPM), to devise a data-driven State of Health (SOH) model to predict the RUL of a battery system. The data they used was obtained by independently testing eight battery cells over three different temperatures.

Soualhi et al. (2014) used unlabeled data of bearings run to failure from the Prognostics Center of Excellence of

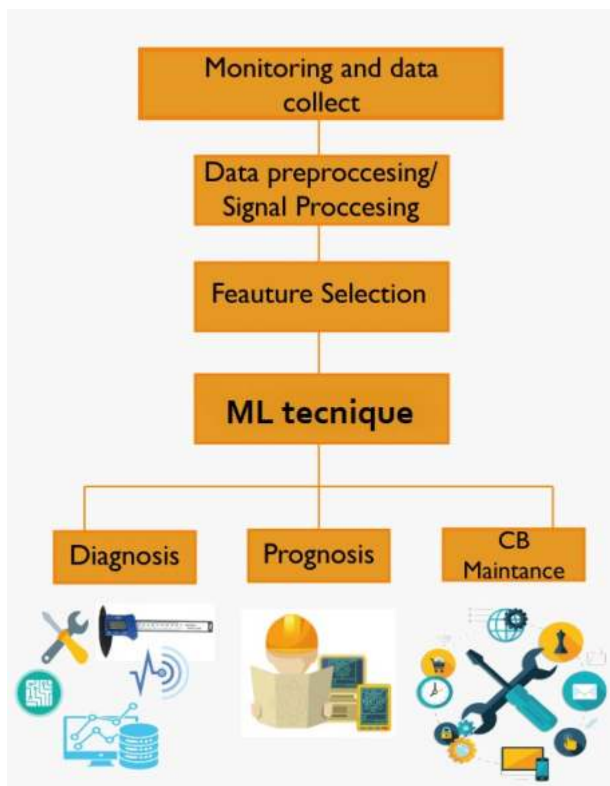

Fig. 2. Prognosis and health management in a process

the NASA. Using an unsupervised classification tool, they generated the classes for the state prediction. A *Hidden*

Table 2. Fault Prognosis in automotive and aerospace industries

Application	Algorithms	Results	Researches
Hybrid systems with intermittent faults and uncertain parameters	Adaptive Reinforcement Unscented Kalman Filter (ARUKF)	Consistent accuracy of 95% and always obtaining better results than regular Unscented Kalman Filters.	Developed. Yu et al. (2018)
Distribution Networks	Data Mining (Neural Networks)	The developed method predicted 27 faults from a real fault count of 32. The accuracy of the developed rule for prognosis is of 84.4%.	Developed. Wang et al. (2018)
Bearings	Generative Adversarial Network (GAN)	No percentage of accuracy is presented. Algorithm was tested with a data set from IMS, University of Cincinnati having acceptable results, suggesting that GANs are a suitable method to use for bearing fault prognosis with non-Gaussian noise systems.	Tested. Khan et al. (2018).
Turbofans	Recurrent Neural Network (LSTM)	The proposed approach performs especially great in the latter half of the lifetime which is more crucial than the first half in practical applications.	Developed. Wu et al. (2018).
Electric Vehicles Battery	Sparse Bayesian Predictive Model (SBPM)	In the first cycles, the RUL has relatively large standard deviation compared to measured data. However, as more state-of-health (SOH) estimates become available, the predicted RUL has a shrinking standard deviation, and its average has excellent agreement with the true one.	Tested. Hu et al. (2016)

Markov Models (*HMM*) based prognosis model was used to predict an "Imminence of the Next Degradation State" (INDS) and a *Adaptive Neuro-Fuzzy Inference Systems* (*ANFIS*) as an extrapolation tool to estimate the "Remaining Time before Next Degradation State" (RNTDS).

Khan et al. (2018) used *Generative Adversarial Networks* (*GAN*) on run-to-fail bearing data from IMS, University of Cincinnati to model the deterioration of bearings and be able to diagnose and predict according to wear when these will fail. The modelled indicator is called a Health Indicator (HI) and it is obtained by using RMS on the sampled vibration acceleration signals. The *GAN* uses a Generator Network, which tries to learn and imitate the behaviour of a given indicator and a Discriminator Network, which tries to differentiate between the real sample data and the data generated by the Generator Network. Both networks are trained until the Discriminator Network can no longer tell the difference between the real and the generated data. The data consisted in three data sets of bearings ran to failure. Only two of them were used for testing the proposed *GAN*. While using it on the first dataset showed great results, it showed regular ones on the second one (no accuracy data is given by the authors). Further inspection of the scale of the HI is advised.

Yuan Xie (2016) used an *Echo State Network* (*ESN*) and a *Recurrent Multilayer Perceptron* (*RMLP*) to predict failures in bearings using 2000 samples from the CWRU dataset containing normal and faulty conditions. The networks were trained first with 500 training samples and then 1000 samples. Results showed that compared against the autoregressive moving average (ARMA) and the Support Vector Machine (SVM), both of the proposed networks (ESN and RMLP) had better Root Mean Squared Errors than the ARMA and the SVM. When using 1000 samples the accuracy of the ESN increased while the one from RMLP only did so slightly and got the same result as the ARMA.

Kiakojoori and Khorasani (2014) compared an *Autoregressive neural network with exogenous input* (*NARX*) with an *Elman neural network* to predict degradation due to temperature in a jet engine. Data was obtained using a Simulink model with a given deterioration rate and validated against a gas turbine simulation program developed by the National Aerospace Laboratory (NLR). Both networks were compared using the Normalized Bayesian

Information Criterion (NBIC), showing that for the same training and testing conditions, the Elman Neural Network outperforms the NARX for health prognosis in an aircraft jet engine.
In Lu et al. (2013) Particle filters (PF), Extended Kalman Filters (EKF) and constrained Extended Kalman Particle Filters (cEKPF) have been used to monitor the status of the gas path for a turbofan engine based with non-Gaussian noise.
Juesas et al. (2017), and Giantomassi (2011) used decision processes of semi-Markov, Markov Switching Models and DBN and HMM to predict the *RUL* for critical engineering systems as those found in aircrafts.

Prognosis has also been used extensively in the area of medicine as a way to create correlations between specific genes or proteins and various types of cancer, although most of them are using some statistic method like Kaplan-Meier and Cox analysis instead of a proper Deep Learning technique. For example, in investigations related to ovarian cancer and its relationship with some proteins Shen et al. (2019), cervical cancer Feng et al. (2019), gastric cancer Zheng et al. (2019) and some genes related to it Moon et al. (2019), Liu et al. (2019) as the same in pancreatic cancer, Vundavilli et al. (2018), and several studies focused on the early prognosis of breast cancer, Sun et al. (2018), Fu et al. (2018), Peng et al. (2019). It has also been highly used in the relationship of different types of carcinoma certain genes Wu et al. (2019) or proteins Ma et al. (2018).

4. DISCUSSION

Deep Learning is one of the most commonly used techniques for data-driven analysis. There are many different architectures, such as those mentioned in this review. But, regardless of which architecture is used, the type of data and characteristics that are needed to select the network that best fits with these must be considered. While most of the authors use a supervised training technique, unsupervised training should be most considered as it is the most common way data is found for real applications.

Among the most studied applications in the automotive and aerospace industries are bearings, internal combustion engines, batteries of electric vehicles, and turbofans. Most used methods for prognosis are Neural Networks and some variations of them, like *GAN*, *LSTM-RNN*, *CNN*, *DBN* and *BNN*, among others. It is also worth mentioning that most of the used algorithms for fault prognosis have been developed by the authors, adjusting one of the mentioned main algorithms to fit their own particular available data.

Further research to test and to develop new algorithms dedicated to the automotive and aerospace industries must be made. This article showed a variety of algorithms and their applications for predicting faults in different systems, their advantages and limitations, and some areas of application of interest at present.

REFERENCES

Diez Oliván, A. (2017). *Machine Learning for Data-driven Prognostics: Methods and Applications*. Ph.D. thesis, Industriales.

Feng, L.L., Shen, F.R., Zhou, J.H., and Chen, Y.G. (2019). Expression of the lncRNA ZFAS1 in Cervical Cancer and its Correlation with Prognosis and Chemosensitivity. *Gene*, 696, 105 – 112.

Fu, B., Liu, P., Lin, J., Deng, L., Hu, K., and Zheng, H. (2018). Predicting Invasive Disease-Free Survival for Early-stage Breast Cancer Patients Using Follow-up Clinical Data. *IEEE Transs on Biomedical Eng*, 1–1.

Giantomassi, Andrea. Ferracuti, F.B.A.I.G.L.S.P.A. (2011). Hidden Markov Model for Health Estimation and Prognosis of Turbofan Engines. volume 3, 681 – 689.

Gómez-Ruiz, J.A., Jerez-Aragonés, J.M., Muñoz-Pérez, J., and Alba-Conejo, E. (2004). A neural network based model for prognosis of early breast cancer. *Applied Intelligence*, 20(3), 231–238.

Hong, Z. (2011). A Preliminary Study on Artificial Neural Network. In *6th IEEE Joint Int Information Tech y and Art Intelligence Conf*, volume 2, 336–338.

Hu, X., Jiang, J., Cao, D., and Egardt, B. (2016). Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling. *IEEE Trans on Ind Electronics*, 63(4), 2645–2656.

Juesas, P., Ramasso, E., Drujont, S., and Placet, V. (2017). On Partially Supervised Learning and Inference in Dynamic Bayesian Networks for Prognostics with Uncertain Factual Evidence : Illustration with Markov Switching Models.

Khan, S.A., Prosvirin, A.E., and Kim, J. (2018). Towards Bearing Health Prognosis Using Generative Adversarial Networks: Modeling Bearing Degradation. In *Int Conf on Advancements in Computational Sc*, 1–6.

Kiakojoori, S. and Khorasani, K. (2014). Dynamic Neural Networks for Jet Engine Degradation Prediction and Prognosis. In *Int Joint Conf on Neural Networks*, 2531–2538.

Kim, H.E., Tan, A.C., Mathew, J., and Choi, B.K. (2012a). Bearing Fault Prognosis Based on Health State Probability Estimation. *Expert Systems with Applications*, 39(5), 5200 – 5213.

Kim, H.E., Tan, A.C., Mathew, J., and Choi, B.K. (2012b). Bearing fault prognosis based on health state probability estimation. *Expert Systems with Applications*, 39(5), 5200–5213.

Komyakov, A.A., Nikiforov, M.M., Erbes, V.V., Cheremisin, V.T., and Ivanchenko, V.I. (2016). Construction of Electricity Consumption Mathematical Models on Railway Transport Using Artificial Neural Network and Fuzzy Neural Network. In *IEEE 16th Int Conf on Env and Elecl Eng*, 1–4.

Kordestani, M., Samadi, M.F., Saif, M., and Khorasani, K. (2018). A New Fault Prognosis of MFS System Using Integrated Extended Kalman Filter and Bayesian Method. *IEEE Trans on Ind Inf*, 1–1.

Lehman, L.w., Ghassemi, M., Snoek, J., and Nemati, S. (2015). Patient prognosis from vital sign time series: Combining convolutional neural networks with a dynamical systems approach. In *2015 Computing in Cardiology Conference (CinC)*, 1069–1072. IEEE.

Liu, S., Mao, Q., Xue, W., Zhang, X., Qi, Y., Wang, Y., Chen, P., and Zhou, Q. (2019). High Expression of ALPPL2 is Associated with Poor Prognosis in Gastric Cancer. *Human Pathology*, 86, 49 – 56.

Lu, F., Huang, J., and Lv, Y. (2013). Gas Path Health Monitoring for a Turbofan Engine Based on a Nonlinear Filtering Approach. *Energies*, 6(1), 492–513.

Luo, J., Tu, H., Pattipati, K., Qiao, L., and Chigusa, S. (2005). Graphical models for diagnosis knowledge representation and inference. In *IEEE Autotestcon*, 483–489. IEEE.

Ma, H., Wu, J., Li, X., and Kang, R. (2018). Condition-Based Maintenance Optimization for Multi-Component Systems under Imperfect Repair-Based on RFADT Model. *IEEE Trans on Fuzzy Systems*, 1–1.

Ma, Y.Y., Zhang, G.H., Li, J., Wang, S.B., Hu, Z.M., Zhang, C.W., and Li, E. (2018). The Correlation of NLRC3 Expression with the Progression and Prognosis of Hepatocellular carcinoma. *Human Pathology*, 82, 273 – 281.

Moon, S., Balch, C., Park, S., Lee, J., Sung, J., and Nam, S. (2019). Systematic Inspection of the Clinical Relevance of TP53 Missense Mutations in Gastric Cancer. *IEEE/ACM Trans on Computational Biology and Bioinformatics*, 1–1.

Peng, C., Zheng, Y., and Huang, D. (2019). Capsule Network based Modeling of Multi-omics Data for Discovery of Breast Cancer-related Genes. *IEEE/ACM Trans on Computational Biology and Bioinformatics*, 1–1.

Ramasso, E. and Gouriveau, R. (2010). Prognostics in switching systems: Evidential markovian classification of real-time neuro-fuzzy predictions. In *2010 Prognostics and System Health Management Conference*, 1–10. IEEE.

Sankavaram, C., Pattipati, B., Kodali, A., Pattipati, K., Azam, M., Kumar, S., and Pecht, M. (2009). Model-Based and Data-Driven Prognosis of Automotive and Electronic Systems. In *2009 IEEE Int Conf on Auto Science and Eng*, 96–101.

Shen, W., Niu, N., Lawson, B., Qi, L., Zhang, J., Li, T., Zhang, H., and Liu, J. (2019). GATA6: A New Predictor for Prognosis in Ovarian Cancer. *Human Pathology*, 86, 163 – 169.

Shridhar, K. (2019). Bayesian Neural Network Series Post 1: Need for Bayesian Neural Networks.

Shubhangi Vashisth, Alexander Linden, J.H.P.K. (2019). Hype cycle for data science and machine learning.

Soualhi, A., Razik, H., Clerc, G., and Doan, D.D. (2014). Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System. *IEEE Trans on Ind Elec*, 61(6), 2864–2874.

Sun, D., Wang, M., and Li, A. (2018). A Multimodal Deep Neural Network for Human Breast Cancer Prognosis Prediction by Integrating Multi-Dimensional Data. *IEEE/ACM Trans on Computational Biology and Bioinformatics*, 1–1.

Vachtsevanos G., Lewis Frank. Roemer M, H.A. and Wu, B. (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*.

Vundavilli, H., Datta, A., Sima, C., Hua, J., Lopes, R., and Bittner, M.L. (2018). In Silico Design and Experimental Validation of Combination Therapy for Pancreatic Cancer. *IEEE/ACM Trans on Computational Biology and Bioinformatics*, 1–1.

Wang, X., McArthur, S.D.J., Strachan, S.M., Kirkwood, J.D., and Paisley, B. (2018). A Data Analytic Approach to Automatic Fault Diagnosis and Prognosis for Distribution Automation. *IEEE Trans on Smart Grid*, 9(6), 6265–6273.

Wu, Q., Ding, K., and Huang, B. (2018). Approach for Fault Prognosis using Recurrent Neural Network. *Jrnl of Intelligent Manufacturing*, 1–13.

Wu, S., Wu, D., Pan, Y., Liu, H., Shao, Z., and Wang, M. (2019). Correlation Between EZH2 and CEP55 and Lung Adenocarcinoma Prognosis. *Pathology - Research and Practice*, 215(2), 292 – 301.

Yu, M., Lan, D., Huang, Y., Wang, H., Jiang, C., and Zhao, L. (2018). Event-Based Sequential Prognosis for Uncertain Hybrid Systems With Intermittent Faults. *IEEE Trans on Ind Inf*, 1–1.

Yuan Xie, T.Z. (2016). The Application of Echo State Network and Recurrent Multilayer Perceptron in Rotating Machinery Fault Prognosis. In *IEEE Chinese Guidance, Navigation and Control Conf*, 2286–2291.

Zheng, H., Yu, J., Li, W., Yang, D., Gao, C., Zhang, Q., and Xu, L. (2019). Is Co-Expression of USP22 and HSP90 More Effective in Predicting Prognosis of Gastric Cancer. *Pathology - Research and Practice*, 215(4), 653 – 659.

