Predictive Maintenance of Naval Assets Using Machine Learning Techniques

Lamas López, Francisco
CESADAR, Armada Española.
Arsenal de Cartagena
SPAIN

flamlop@mde.es

Novoa Paradela, David
CITIC
Universidad de A Coruña.
SPAIN

david.novoa@udc.es

Eiras Franco, Carlos
CITIC
Universidad de A Coruña.
SPAIN

carlos.eiras.franco@udc.es

Fontenla Romero, Óscar
CITIC
Universidad de A Coruña.
SPAIN

oscar.fontenla@udc.es

Sanz Muñoz, David
DigitalLabs / IA
INDRA
SPAIN

dsmunoz@indra.es

ABSTRACT

The appearance of anomalies during the operation of industrial assets can point to the presence of degradations and failures, which over time lead to undesired behavior, loss of operation conditions and the final breakdown of the system. Predictive maintenance techniques are in charge of monitoring the status of the systems in order to carry out the detection of these anomalies in incipient phases, allowing to schedule maintenance tasks in an optimal way. This paper presents a predictive maintenance solution for naval assets based on artificial intelligence techniques as Machine Learning. For this, the information from the sensors (temperatures, pressures, etc.) collected in real time by the vessels and transmitted through the control center is used. The system developed (SOPRENE) is able to predict the occurrence of different failure modes or abnormal operating conditions from a historical data of an engine on board of our warships. In addition, the use of this system is scalable to large fleets, the solution has been implemented using the Spark distributed environment in order to facilitate distributed computation of predictions.

Keywords: machine learning, predictive maintenance, Spark, SOPRENE.

1. INTRODUCTION

The cost of maintenance represents an important part of the operating costs in industry. In some cases, as in the metallurgical industry, these costs can amount to 15% -60% of total production costs. Furthermore, of these, a third of the investment is wasted as a result of unnecessary or incorrect activities [1]. However, maintenance is crucial as the failure of a system can lead to huge financial costs.

In the past, the impossibility of handling large and continuous data flows has led to the use, in many cases, of statistical techniques. Today's predictive maintenance, however, follows a more advanced philosophy:

Instead of relying on these industry statistics (e.g. mean time between failures) to schedule maintenance activities, real-time monitoring of the system is carried out to determine its status and real condition. The current computing capacity allows both processing larger amounts of data, as well as the use of more sophisticated techniques to carry out predictions, detection of abnormal conditions and a possible diagnosis of the system. Therefore, predictive maintenance can be understood as preventive maintenance [2] based on the current state or condition of the system and future predictions made from an operation history.

This research work presents the development of a predictive maintenance system framed within the SOPRENE project in its application to navy ship engines. The proposed system has analyzed and used machine learning techniques in distributed environments. In this sense, the considered methodologies can be divided according to what was stated by Ran et al. [3]:

1.1 Categories based on purpose

Based on the optimization criteria that could be followed, we can distinguish between different methods:

- Cost minimization: the metric used is usually the system's Remaining Useful Life (RUL), although it is also possible to define an ad hoc cost model [4].

- Reliability maximization and asset availability: this metrics are calculated in order to estimate the probability of a system to be in a normal operating state given a time interval [5] and the probability that the system is operational [6].

- Multi-objective optimization: it seeks to optimize multiple metrics simultaneously to achieve a better balance between objectives. In addition to the aforementioned, they use metrics such as risk, security or viability. Generally, it is impossible to obtain optimal values for all objectives at the same time, so a wide variety of multiobjective models have been developed [7] - [9].

1.2 Categories depending on the approximation

Based on the type of approximation used, we can distinguish between:

- Knowledge-based approaches. Expert knowledge and deductive reasoning processes are used. There are approximations based on ontologies [10], on rules [11] or on analytical models that try to link the physical processes of a system with mathematical models, such as Gaussian models [12], models of linear systems [13] or Markov models [14].
- Approaches based on classic machine learning techniques (machine learning, ML). Artificial neuron networks [15], decision trees [16] (including the Random Forest algorithm [17]) have been used. as well as vector support machines (SVM), both supervised [18] and unsupervised [19]. Finally, the nearest neighbor technique (k-NN) is one of the most common used methods for classifying failures [20], for predicting lifetime (RUL) [21] and early detection [22].
- Approaches based on deep learning. One of the most used are the Autoencoder Neural Networks, whose output layer seeks to reproduce the data presented in its input layer after having gone through a dimensional compression phase, allowing the creation of robust models against noise [23]. Recurrent neural networks (RNN) have also been used in litterature, based on Long Short Term Memory Networks (LSTM) cells [24], which can learn longer-term dependencies. These types of networks are very powerful for sequence analysis [25].

2. PROBLEM CONTEXTUALIZATION

The application of predictive maintenance techniques to the naval assets or engines of the Spanish Armada's ships is defined by the installed monitoring system: The vessels considered in this study (BAM) record the values of nearly 5000 variables by unit (about 300 associated with an example engine), registered every 10 s (sampling frequency depends also on the variable), representing a wide variety of aspects physics of the same (e.g., gas temperature, pressure in the filters, etc), and with more than 10 years of registered operating data history for this kind of vessels. However, the recorded data show heterogeneous quality due to sensor failures (erroneous or missing values) or communication problems (non-uniform sampling frequencies). Secondly. There are three engine operating modes depending on the revolutions per minute (RPM) at which it works: engine off (RPM close to zero), engine idling (RPM close to a threshold μ) and engine normal operation (RPM greater than a threshold μ), having to characterize and filter the states of interest.

Regarding the malfunctions in the data history, due to the extension of the files and their identification, manual labeling by expert is not feasible. Thus, there is no set of structured failure modes and, therefore, the detection of anomalies must be carried out in an unsupervised manner. To fill this gap, an FMECA (Failure Mode, Effects, and Criticality Analysis) analysis is used, which theoretically describes the failure modes that can occur over an engine (the variables or elements that intervene in these modes and their values. nominal, maximum and minimum).

Finally, since the proposed solution should not analyze a single vessel, but a fleet, the management and control is carried out centrally (on land, at CESADAR, the Spanish Armada's data center). Each ship sends the data recorded by its sensors to a central node where it is stored and processed. Since the datasets to be dealt with can be large, and for the system to be scalable to a large number of vessels, the system has to be distributed: the HDFS distributed file system has been used to store the data and the Apache Spark environment to train and run the models in a distributed way.

3. PROPOSED SOLUTION

The designed solution combines the solutions to several sub-problems. Thus, the general architecture is made up of four clearly defined blocks or tasks (see Figure 3-1):

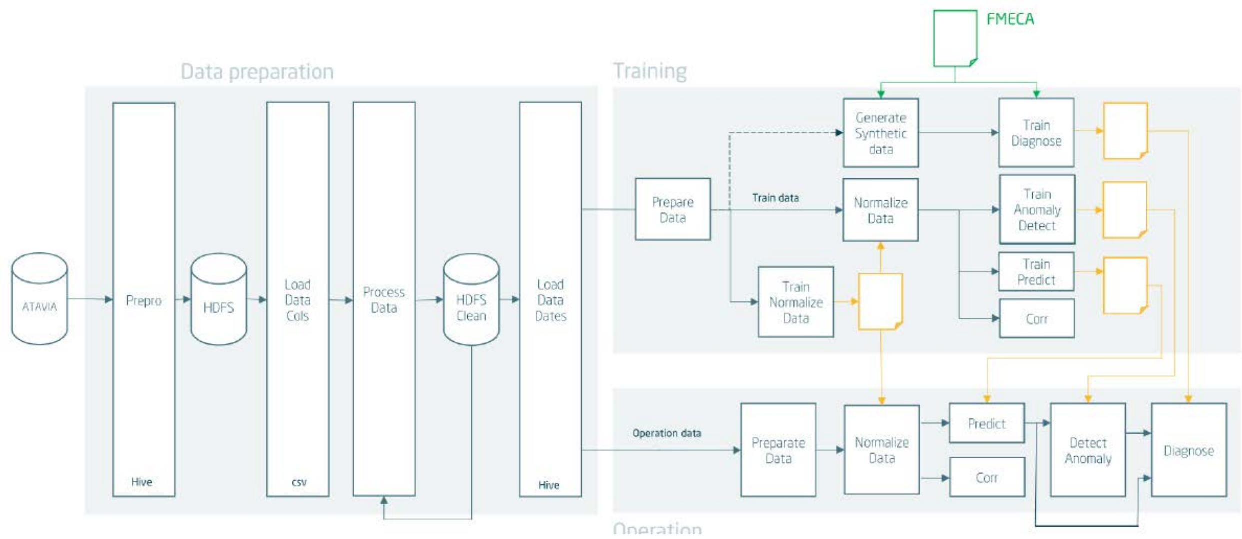

Figure 3-1. SOPRENE solution architecture: from Data Preprocessing to training and operation

3.1. Pre-processing module

In order to train ML models it is necessary to have a robust and representative dataset of the different operating states of the engine or naval asset. The values measured by the system sensors are stored in CSV files into the HDFS system, however these data need to be pre-processed. Thus, in this block, the existence of missing values is checked, the sampling frequency is unified and a variable selection process is carried out:

- Data loading: The flow begins by reading the dataset to be processed and stored in HDFS.
- Data processing: In order to solve the problem of the low quality of the original data (section 3), two actions are carried out:
    1. Variables selection: of all the variables, those that show sufficient variability in their values will be used. This selection is carried out both automatically (discarding variables that have constant values) and manually, eliminating those selected by the user.
    2. Sampling frequency standardization: in order to avoid the second problem, it is necessary to homogenize the sampling frequency. Thus, a set is created from the initial data in which there is a simultaneous measurement for all the variables every 60 s. For this, the absence of values of the variables that present a lower frequency is filled with the last available and validated value.

The processed data is stored in a consolidated and structured Hive database.

- Data normalization: Since the values collected by the sensors oscillate in very different ranges for each variable, a normalization process is executed individually. The trained normalizer is stored in HDFS to be used in production with the arrival of new data.
- Data preparation: To provide flexibility to the system, the user can establish the unit of the horizon with which to carry out the prediction (hours, days, weeks or months), so that the normalized data is temporarily grouped (eg if predictions are going to be made in days, the normalized data is grouped into a single data per day). In addition, before grouping the data, these are filtered based on the value of the RPM variable to eliminate those that correspond to moments of the engine off.

3.2. Prediction module

The main goal is to know the state of the engine at an instant in future time. In this way, to carry out a prediction from an instant in time t_i, to a future instant that is distant from the horizon units of engine use time (t_i + horizon), the system must receive the information collected by the sensors during the last window units of time before you (t_i-window), defined next to the horizon by the user. Figure 3-2 shows a graphical representation of these concepts.

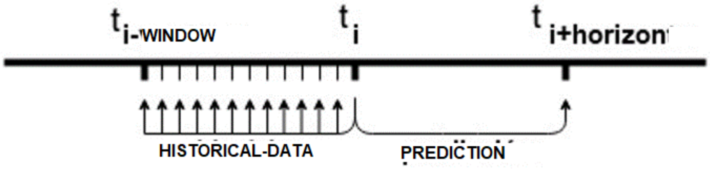

Figure 3-2. Using a previous data window to carry out a prediction.

The prediction methods available are: linear regression, and of LSTM and Elastic Net networks, as shown in Table 3-1. After training the models, they are stored in HDFS to be used in production in the same way as the normalizers.

Table 3-1. Mean squared errors (MSE) obtained by the algorithms using different data grouping modes: 1 data/day (D), 1 data/week (W) and 1 data/month (M). The prediction horizons used were 5 and 10 time units.

	Horizon = 5			Horizon = 10		
Algorithm	D	W	M	H	D	W
Linear Regression	0.65	0.36	0.22	0.64	0.35	0.27
L1	0.62	0.36	0.22	0.61	0.32	0.25
L2	0.60	0.33	0.20	0.64	0.31	0.24
Elastic Net	0.56	0.32	**0.15**	0.57	0.33	**0.06**
LSTM	**0.41**	**0.28**	0.16	**0.40**	**0.27**	0.16

3.3. Anomaly detection module

Once the prediction of the naval asset condition has been made (Figure 3-3), it is necessary to determine whether the condition corresponds to a normal value or not. With no tagged anomalies available, the detection process is performed unsupervised using an Autoencoder Neural Network. By means of the reconstruction error we can discern between normal data (low reconstruction error) and anomalous data (high errors). This value can be presented as the mean square error of all the input variables (a single value) or its decomposition, the error of each of the input variables or nodes of the network (Shadow area in Figure 3-3). The objective, given a set of records, is to determine which are anomalous and which variables cause these anomalies, which is done in three sequential sub-phases:

1. Detect anomalies: to determine which records are anomalous, a first filter is carried out using the mean square error. Based on a pre-calculated threshold error, the data that exceed it are classified as anomalous and the rest as normal. The user chooses whether the calculation of this threshold error is carried out by means of the interquartile range or by establishing a percentage of anomalous data in the set.

2. Separate contributions: to determine which variables have been the cause of the appearance of the anomaly in the data classified as anomalous, the decomposition of the reconstruction error is used. Thus, it goes from a single global error to as many as variables make up the record, being able to sort the variables by their reconstruction error and automatically determine the contribution of each variable using the Elbow method [26].

3. Construct an anomaly mask: from the selection of the previous sub-phase, a matrix or output mask of dimensions $m_x n$ is constructed (m being the number of rows or records and n the number of columns or variables) in which the anomalous variables are marked with a one and normal variables with a zero.

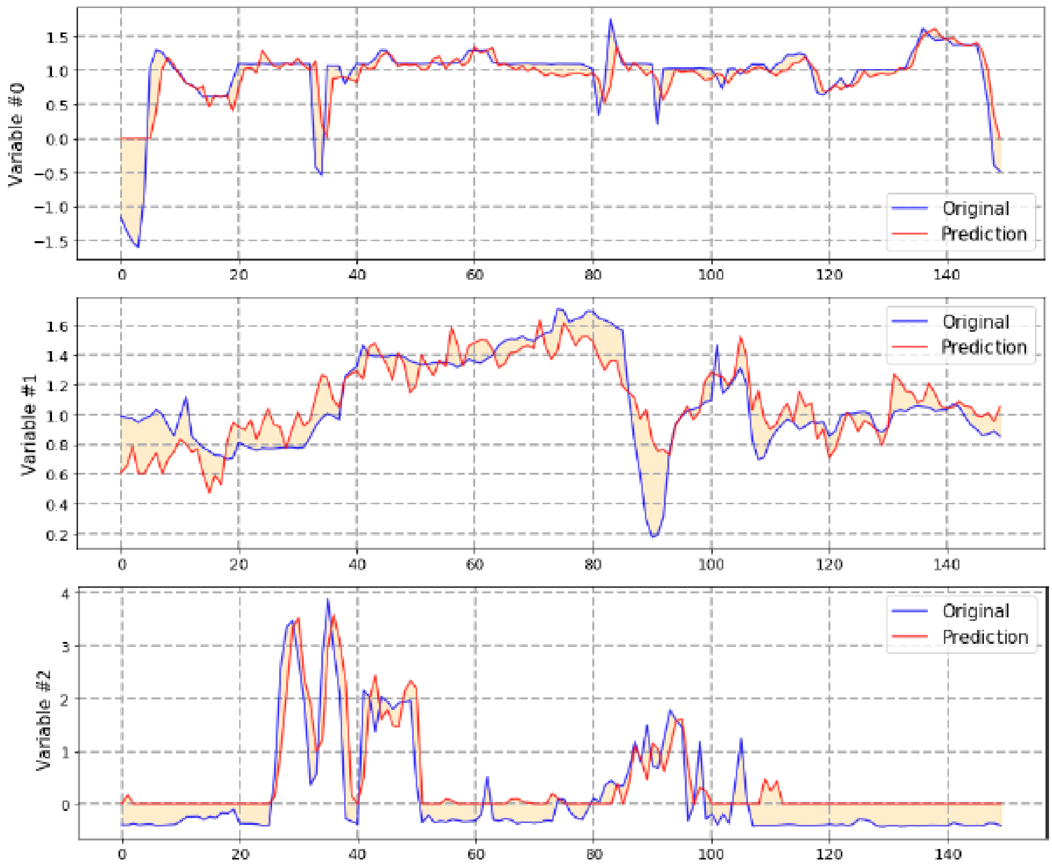

Figure 3-3. Long-Short Term Memory Networks (LSTM) prediction output (red) versus the real value to be predicted (blue) for three attributes of the diesel engine for propulsion. X-axis represents normalized data coming from the BAM ship; Y-axis contains an ordered index. Shadowed area represents prediction error.

3.4. Diagnostic module

The diagnostic module is responsible for, based on the prediction and the anomaly detection mask, determining which failure modes may occur and their probability. To determine the probability of each failure mode it has been decided to use a supervised classification model based on neural networks (Figure 3-4). Since there are no labeled datasets of all possible failure modes, an artificial data generator has been implemented to produce them from the theoretical characteristics of the failure modes included in the FMECA document (variables involved, nominal values, range stop values, etc.). Thus, the model is trained considering each of the failure modes as an output class, having as many neurons as motor variables in its input layer, and as many *softmax* as failure modes in its output layer.

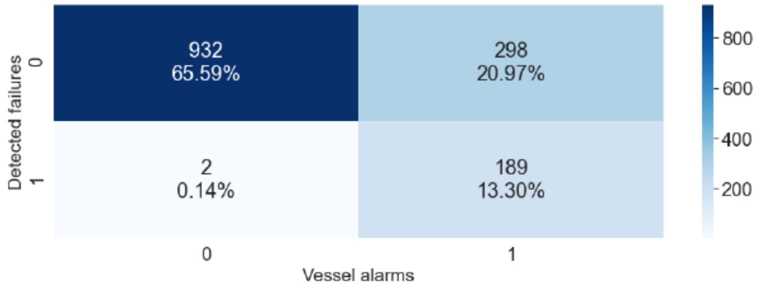

Figure 3-4. Confusion matrix of the failures detected by the BAM's engine alarm system versus the anomalies detected by the model.

The classifier of failure modes is used in combination with the output mask of the anomaly detection module to carry out the diagnosis of the engine. This was used to limit the number of possible failure modes, omitting the failure modes in which all its variables involved have been considered normal. The output of this module (and end of the system) contains the identifier of the failure mode in the FMECA, the expected date on which the failure mode will occur and the probability / certainty associated with it (Figure 3-5).

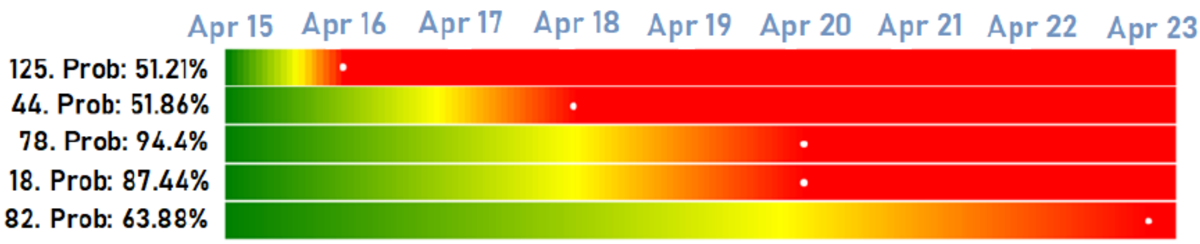

Figure 3-5. Example of system output for five failure modes (named with ID 125, 44, 78, 18 and 82) sorted temporally. Each row depicts the evolution of the probability of failure along the time. The points where the failure probability reaches its peak are marked with white dots.

4. RESULTS

Since there is no set of targets or reference values of abnormal operating conditions (conveniently structured for this purpose), it has not been possible to perform a quantitative evaluation of the system. The development of the different partial solutions that make up the final solution has been subject to a mostly qualitative evaluation by the experts of the organization involved:

- Prediction: Filtering the data by RPM and grouping it considerably reduces the number of available examples. This disables models such as LSTM networks that require large amounts of data. In these scenarios, it has been seen that the simplest models report better results, being possible to make predictions up to 10 days / weeks / months with reasonable quality.
- Anomaly detection: The results of the anomaly detector have been compared with the alarms recorded on the vessels over the four years. As can be seen in Figure 4-1, the model was able to detect most of these anomalies and even anticipate the occurrence of some of them.
- Diagnosis: the generation of artificial datasets based on FMECA has made it possible to build classification models that determine which failure modes may occur (or abnormal operating conditions). However, this theoretical behavior of the engine does not always correspond to reality, since its operation may vary with the use, replacement or repair of parts.

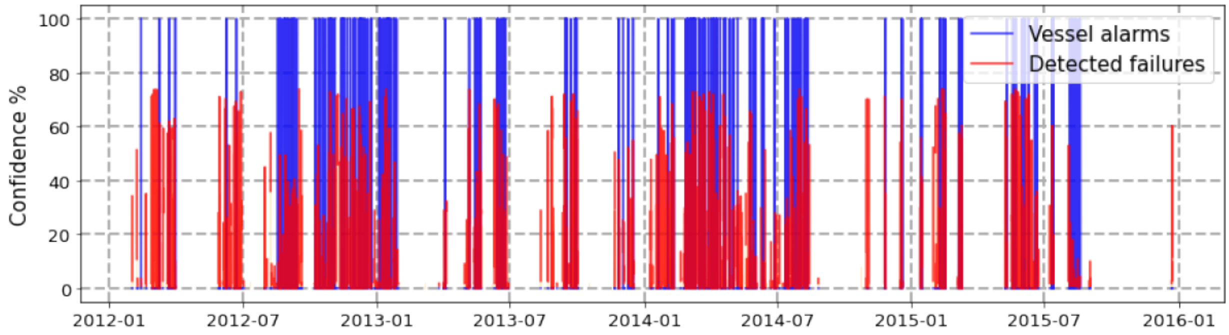

Figure 4-1. Normalized comparison of the vessel alarms (in blue) with the autoencoder's joint diagnostics for all failure modes. Data comes from a BAM ship. Y-axis represents the confidence given by the autoencoder. X-axis represents normalized data coming from the vessel.

5. CONCLUSIONS

The solution developed makes it possible to predict the occurrence of the different failure modes or abnormal operating conditions described in the FMECA of a warship's propulsion engine. The tasks of prediction and detection of anomalies are totally independent, so the latter can be carried out both for future moments (data from the prediction), present (real time) or past (a posteriori analysis). The greatest responsibility lies with the prediction module since subsequent operations start from the output of this module. To provide it with flexibility, a wide range of methods are provided in SOPRENE, easily scalable if needed, while allowing the user to configure the prediction parameters to obtain the most appropriate results in each situation. The system is highly configurable and its use can be extrapolated to other warship's assets with similar characteristics. Its computation is distributed, so prediction, detection and diagnosis times are low, while the biggest bottleneck is the pre-processing task, which is not performed in a distributed manner.

ACKNOWLEDGEMENTS

This research was funded by DGAM-PLATIN (General Directorate of Armament and Material—Spanish Ministry of Defence) with project funding number '2018/1003218001400'. This paper reports part of the results obtained within the 'SOPRENE' research and development project funded by DGAM-PLATIN (General Directorate of Armament and Material—Spanish Ministry of Defence). This project developed a technological demonstrator of AI techniques and applications over naval platform assets for predictive maintenance purposes. The authors are also grateful to all the Spanish Armada's units that have participated in some manner in the obtaining of these results and for providing their technical advice in FMECA development, without which this assessment of AI techniques would not be accomplished.

REFERENCES

[1] R. K. Mobley, An Introduction to Predictive Maintenance, 2nd ed., Elsevier, 1990.

[2] O. Motaghare, A. S. Pillai, and K. I. Ramachandran, "Predictive maintenance architecture," IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2018, pp. 1–4.

[3] Y. Ran, X. Zhou, P. Lin, Y. Wen, and R. Deng, "A survey of predictive maintenance: Systems,

purposes and approaches," arXiv:1912.07383, 2019.

[4] Y. He, X. Han, C. Gu, and Z. Chen, "Cost-oriented predictive maintenance based on mission reliability state for cyber manufacturing systems," Advances in Mechanical Engineering, 2018.

[5] S. Song, D. W. Coit, and Q. Feng, "Reliability analysis of multiple-component series systems subject to hard and soft failures with dependent shock effects," IIE Transactions, vol. 48, no. 8, pp.720–735, 2016.

[6] M. A. Gravette and K. Barker, "Achieved availability importance measure for enhancing reliability-centered maintenance decisions," Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 229, no. 1, pp. 62–72, 2015.

[7] L. Lin, B. Luo, and S. Zhong, "Multi-objective decision-making model based on CBM for an aircraft fleet with reliability constraint," International Journal of Production Research, vol. 56, no. 14, pp. 4831–4848, 2018.

[8] J. Zhao and L. Yang, "A bi-objective model for vessel emergency maintenance under a condition-based maintenance strategy," Simulation, vol. 94, no. 7, pp. 609–624, 2018.

[9] Y. Xiang, D. Coit, and Z. Zhu, "A multi-objective joint burn-in and imperfect condition-based maintenance model for degradation-based heterogeneous populations," Quality and Reliability Engineering International, vol. 32, no. 8, pp. 2739–2750, 2016.

[10] A. Konys, "An ontology-based knowledge modelling for a sustainability assessment domain," Sustainability, vol. 10, 300, 2018.

[11] M. D. Ying Peng and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," The International Journal of Advanced Manufacturing Technology, vol. 50, pp. 297–313, 2010.

[12] N. Chen, Z.-S. Ye, Y. Xiang, and L. Zhang, "Condition-based maintenance using the inverse Gaussian degradation model assessment domain," European Journal of Operational Research, vol. 243, no. 1, p. 190–199, 2015.

[13] A. Lucifredi, C. Mazzieri, and M. Rossi, "Application of multiregressive linear models, dynamic kriging models and neural network models to predictive maintenance of hydroelectric powers ystems," Mechanical Systems and Signal Processing, vol. 14, no. 3, pp. 471–494, 2000.

[14] M. Calder and M. Sevegnani, "Stochastic model checking for predicting component failures and service availability," IEEE Transactions on Dependable and Secure Computing, vol. 16, no. 1, pp. 174–187, 2019.

[15] B. Samanta and K. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features," Mechanical Systems and Signal Processing, vol. 17, pp. 317–328, 2003.

[16] J. R. Quinlan, "Improved use of continuous attributes in C4.5," CoRR, vol. cs.AI/9603103, 1996. [Online]. Available: https://arxiv.org/abs/cs/9603103

[17] J. Shi, N. Niu, and X. Zhu, "A fault diagnosis method for multi-condition system based on random

forest," Prognostics and System Health Management Conference (PHM-Paris), pp. 350–355, 2019.

[18] Y. Chen, H. Yigang, Z. Li, L. Chen, and C. Zhang, "Remaining useful life prediction and state of health diagnosis of lithium-ion battery based on second-order central difference particle filter," IEEE Access, vol. 8, pp. 37305-37313, 2020.

[19] G. Ratsch, S. Mika, B. Scholkopf, and K. Muller, "Constructing boosting algorithms from SVMs: an application to one-class classification," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1184–1199, 2002.

[20] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifier approach," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812–820, 2015.

[21] Z. Liu, W. Mei, X. Zeng, C. Yang, and X. Zhou, "Remaining useful life estimation of insulated gate bipolar transistors (IGBTs) based on a novel volterra k-nearest neighbor optimally pruned extreme learning machine (VKOPP) model using degradation data," Sensors, vol. 17, no. 11: 2524, 2017.

[22] X. long Chen, P. hong Wang, Y. sheng Hao, and M. Zhao, "Evidential KNN-based condition monitoring and early warning method with applications in power plant," Neurocomputing, vol. 315, pp. 18 – 32, 2018.

[23] M. Ma, C. Sun, and X. Chen, "Deep coupling autoencoder for fault diagnosis with multimodal sensory data," IEEE Transactions on Industrial Informatics, vol. 14, pp. 1137–1145, 2018.

[24] J. Yuan and Y. Tian, "An intelligent fault diagnosis method using GRU neural network towards sequential data in dynamic processes," Processes, vol. 7, no. 3: 152, 2019.

[25] R. Yang, M. Huang, Q. Lu, and M. Zhong, "Rotating machinery fault diagnosis using long-short-term memory recurrent neural network," IFAC-Papers OnLine, vol. 51, no. 24, pp. 228–232, 2018, 10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCES).

[26] P. Bholowalia and A. Kumar, "Article: Ebk-means: A clustering technique based on elbow method and k-means in WSN," International Journal of Computer Applications, vol. 105, no. 9, pp. 17–24, 2014.