Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-3 (2020) 125–130

A comparison of machine learning techniques for LNG pumps fault prediction in regasification plants

De la Fuente A.*, Crespo Márquez A.*, Candón E.*, Gómez J.F*
Serra J.**

** Dept. of Industrial Management School of Engineering, University of Seville, Spain.*
(e-mail: antfuecar@alum.us.es, adolfo@us.es, eduardocandon@gmail.com)
*** Enagás S.A, Madrid, Spain, (e-mail: jserra@enagas.es)*

Abstract: We present a comparative study on the most popular machine learning methods applied to the challenging problem of Liquefied Natural Gas pumps fault prediction in regasification plants. The proposed solution tries to address the problem of pump failure during operation, this failure makes the pump unavailable, with a high cost of corrective maintenance. It must be taken into account that the condition monitoring may be insufficient because they are cryogenic and inaccessible equipment once the tanks have been started up. The use of machine learning techniques allows us to anticipate the response time by detecting anomalies in the operation, and to be able to do the maintenance before the failure occurs. In our experiments, we predict the power consumption based on the parameters captured in real time during operation. For the composition of the dataset, data was collected between 2007 and 2017, resulting in a dataset of over 15,000 lines for training and validation. First, all models were applied and evaluated on a dataset collected from a real case study. In the second phase, the performance improvement offered by boosting was studied. In order to determine the most efficient parameter combinations we compare Root Mean Squared Error, Absolute Error, Relative Error, Squared Error, Correlation, Training Time and Scoring Time. Our results demonstrate clear superiority of the boosted versions of the models against the plain (non-boosted) versions. The fastest scoring and total time was the Decision Tree and the best overall was Gradient Boosted Trees.

Keywords: Asset and maintenance management, Condition Monitoring, Sensors, Signal Analysis, Failure Analysis, Machine Learning, Predictive Maintenance.

1. INTRODUCTION

The condition and operation monitoring assets play an essential role in the management of the adequate control of the equipment of a company. A special mention refers to assets with a high level of capitalization and a long useful life, since the state of their health can have a great impact on the value factors of the organization (Yam *et al.*, 2001).

In this case, the analysis is focused on cryogenic centrifuge pumps. These pumps are located in regasification plants. These plants are designed to store natural gas in liquid state (-160ºC at atmospheric pressure) and elevate the pressure to 72 bars to gasify Liquefied Natural gas (LNG) and transport it through the pipelines across the country in gas state. The reason of that process is that the volume of LNG is 600 times smaller than Natural Gas (NG), and it is more efficient to transport and storage LNG than NG. On the other hand, is easier to transport NG that LNG that requires cryogenic temperatures.

The pressure elevation is divided in two steps (Fig.1). The first one is done into the LNG storage tanks, where primary pumps elevate the pressure from atmospheric pressure to 9 bars approximately. These pumps are submerged into LNG and are not accessible during operation period. There is other type that elevate pressure from 9 bars to 72 bars (secondary pumps) that are locate out of the tank and are directly accessible to the staff of the facility.

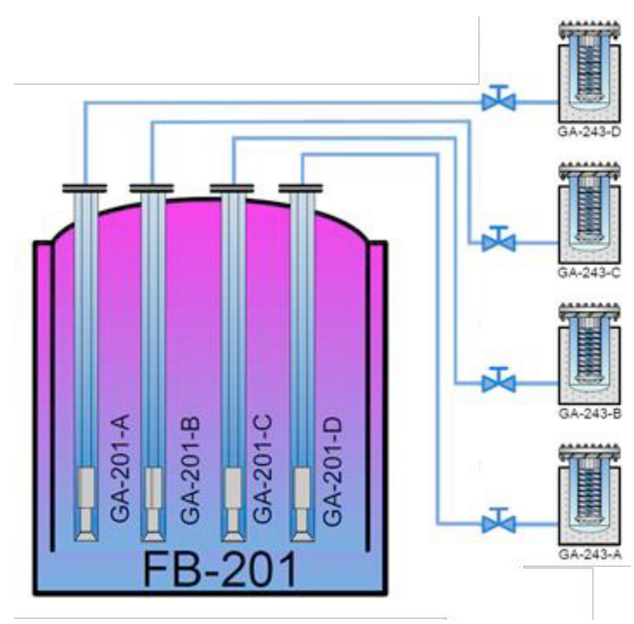

Fig. 1. LNG tank pumping configuration.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2020.11.020

Primary pumps (in-tank), are critical in terms of availability. The maintenance costs are high, not just because the design and materials are very specific, but any actuation requires the extraction of the pump to be manipulate into the workshop. Additionally, the location into the tanks, turns the standard monitoring process not very reliable. Even some parts of the monitoring system (accelerometer, cryogenic wire, amplifier…) are impossible to replace because of the location where are installed.

That is the reason why, even having good standard technology to apply predictive maintenance in centrifuge pumps, preventive plans in primary pumps have been basically based on time.

For the last years, other strategies have been considered, trying to optimize the lifecycle of the asset. The simplest and most effective has been to register operational signals as flow rate, temperature increase, pressure, performance… In terms of operation, these signals help us to ensure that the treatment of the item is under the designed parameters. In terms of maintenance, help us to detect anomalous behaviour. But the period between the observable anomaly for a technician and the failure caused is too close to use this monitoring as a reliably maintenance strategy.

For this reason, we use machine learning techniques, in order to detect anomalies in the operation, and early enough to avoid functional loss of the pump, caused by a failure. Allowing us to properly plan the maintenance of the asset, minimizing the impact on the business.

Several machine learning algorithms have been proposed in the research community to address the problem of predicting the energy consumption of process pumps for anomalies detection, related to the loss of performance.

Such methods include Artificial Neural Networks (Márquez, De la Fuente and Antomarioni, 2019; Wang *et al.*, 2019), Decision Trees learning (Hameed, Vaithiyanathan and Kesavan, 2019; Barrios Castellanos *et al.*, 2020), Deep Learning Model (Saufi *et al.*, 2019; Tang, Yuan and Zhu, 2020), Random Forest (Wang *et al.*, 2016), Gradient Boosted trees (Kusiak, Zeng and Zhang, 2013; Darmatasia and Arymurthy, 2017).

This work constitutes a comparison of five of the most used classification methods in the case of the prediction of the power consumption of rotating equipment. In particular, we compare the results of neural networks, decision trees, deep learning, random forest and the reinforced variation of decision trees to improve their performance. The motivation behind our study is to evaluate the suitability of advanced machine learning methods on the problem of predicting the power consumed by a cryogenic pump. In our experiments, we predict the power consumption based on the parameters captured in real time during operation. The variables for modelling are the flow, suction and outlet temperature, pressure, liquid level in the tank, the density of LNG and power consumption (Table 1). For the composition of the dataset, more than ten years of operation has been collected between 2007 and 2017, resulting in a data set of more than 15,000 clean database lines for training and validation. The sampling frequency is the minimum with which the data is stored in the historical data server, in our case one hour.

TABLE 1. Variables and values for data normalisation

Variable	Vmin	Vmax
Flow (m3/h)	250	700
Pressure (kgf/cm2)	6	14
Suction Temperature (Tin) (ºC)	-160,5	-150
Oulet Temperature (Tout) (ºC)	-160,5	-153,5
Liquid level in Tank (Level) (mm)	0	50.000
LNG Density (Dens GNL) (Kg/m3)	430	465
Power consumption (Kwh)	200	350

The rest of the paper is organized as follows. In Section 2, we give a brief presentation of the machine learning techniques that were selected. The evaluation criteria are presented in sections 3. Simulation settings and results are given in Section 4, and in Section 5 we take out Our conclusions.

2. MACHINE LEARNING TECHNIQUES

Below, the five techniques used in the prediction of the power consumed in pumps for fault detection are briefly presented, taking into account reliability, efficiency and popularity in the research community.

2.1 Artificial Neural Network

Artificial neural networks (ANN) are a very popular solution to solve complex problems, such as the problem of prediction of energetic consumption in pumps. Neural networks can be based on hardware (neurons are represented by physical components) or software (computer models) and can use a wide variety of learning algorithms. A popular supervised model is the multilayer Perceptron trained with variations of the Backward Propagation Algorithm (BPN). BPN is a feedback model with supervised learning (Zhang, Eddy Patuwo and Y. Hu, 1998; Basheer and Hajmeer, 2000). In this case, a feed-forward neural network trained by a back-propagation algorithm is used, with 6 input neurons, 5 neurons in the hidden layer and one in the output layer. There will be a normalization between -1 and 1 prior to training, in order to implement the sigmoid activation function.

2.2 Decision Trees Learning

Decision trees (DT) are tree-shaped structures that represent clusters of detections that can produce class rules for a set of singular data, or as Berry and Linoff pointed out “A structure that can be used to divide a large collection of records in successively smaller sets of records applying a sequence of simple decision rules" (Berry and Linoff, 2004). The most used names for the models are Classification Trees or Regression Trees. In these tree structures, the sheets represent class tags and the branches represent feature rules that lead to those class tags.

2.3 Deep Learning Model

Deep learning Model (DLM) is a subset of machine learning that has three learning techniques, supervised, semi-supervised and unsupervised learning. It consists of many layers of artificial neural networks. Each of the layers contains some neurons with activation functions that can be used to produce non-linear outputs. It is said that this methodology is inspired by the neuronal structure of the human brain. The scientific community is quite open and there are a series of deep learning tutorials and good quality books (Ian Goodfellow, Yoshua Bengio, 2017; Ahmad, Farman and Jan, 2019)

2.4 Random Forest

Random forests (RF) are effective and intuitive models used in regression and classification problems. They are intuitive because they provide a clear path to an outcome and are based on the underlying structures of the Decision Tree. A decision tree is a machine learning model created using a series of decisions based on variable values to take one route or another.

So, each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges as stochastic up to a limit as the number of trees in the forest increases. The generalization error of a tree-classifying forest depends on the strength of the individual trees in the forest and the correlation between them. Internal estimates monitor error, strength and correlation, and are used to show the response to the increase in the number of features used in the division. Internal estimates are also used to measure variable importance. These ideas are also applicable to regression (Breiman, 2001).

2.5 Gradient Boosted Trees

The Gradient Boosted Trees (GBT) (Hill, Lewicki and Lewicki, 2006) results from applying boosting methods to regression trees. Calculate a sequence of trees in which each successive tree is constructed from the prediction residues of the previous tree. At each step of the reinforcement algorithm, the data is divided into two samples in each divided node, determining the best partition and regression errors are calculated. Then, the next tree is adjusted to reduce the error. The driven tree generally improves the performance of a single model by adjusting many models and combining them for prediction. The tree driven algorithm has become one of the most powerful methods in the data mining domain. To avoid overfitting, the maximum number of additive trees is set to 200. The subsample ratio for building a tree is 0.5, and the maximum number of levels is set to 10.

3. EVALUATION CRITERIA

In order to compare the different prediction models, a series of measures are established, which are calculated from the error between the actual value and the prediction value. Of course, the model might have overfitted training data and does not work well with future forecasts unseen. For this reason, training waste is not a good way to measure how forecasting models will perform in the real world. That is why, the same data is used for the validation of the model as for obtaining these meters, which in our case represents 70/30 % for training and validation respectively. We use the measures of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Error (MRE), Squared Error (MSE) and Correlation.

3.1 Mean Absolute Error

The mean absolute error calculation is based on the prediction error, the prediction error e_i is determined by the difference between the value predicted by the model $\hat{y}_i$ and the actual value y_i. The prediction or residual error and MAE for a value it is given by eq. (1):

$$e_i = y_i - \hat{y}_i$$
$$MAE = mean(|e_i|) \tag{1}$$

The error of the individual data point may be positive or negative and may cancel each other out. To derive the overall prediction for the model, calculate the absolute error to aggregate all the residuals and average it. This measure is used to know the accuracy of it.

3.2 Root Mean Squared Error

In some cases, it is advantageous to penalize the individual point error with higher residues. Even though two models have the same MAE, one might have consistent error and the other might have low errors for some points and high error for other points. RMSE penalizes the latter. The RMSE is given by eq. (2):

$$RMSE = \sqrt{mean(e^2)} \tag{2}$$

3.3 Relative Error

The average relative error is the average of the absolute deviation of the prediction of the real value divided by the real value. This measure is used to know the quality of it, and the MRE is given by eq. (3):

$$MRE = mean\left(\frac{(|y_i - \hat{y}_i|)}{y_i}\right) \tag{3}$$

Where $\hat{y}_i$ is the value predicted by the model and y_i is the actual value.

3.4 Squared Error

It is an estimator that measures the average value of errors squared, that is, the difference between the prediction value and the actual value. The difference occurs due to randomness or because the prediction value does not consider the information that could produce a more accurate estimate. The MSE is a risk function, corresponding to the expected value of the quadratic loss. The MSE is given by eq. (4):

$$MSE = \frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2 \qquad (4)$$

Where n represents the number of points in the dataset.

3.5 Correlation

Correlation measures the statistical relationship between two attributes, particularly dependence of one attribute on another attribute. When two attributes are highly correlated with each other, they both vary at the same rate with each other either in the same or in opposite directions. However, correlation between two attributes does not imply causation, that is, one doesn't necessarily cause the other. Correlation between two attributes is commonly measured by the Pearson correlation coefficient (r), which measures the strength of linear dependence. Correlation coefficients take a value from $-1 \leq r \leq 1$. A value closer to 1 or -1 indicates the two attributes are highly correlated, with perfect correlation at 1 or -1. The Pearson correlation coefficient between two attributes $\hat{y}_i$ and y_i is calculated with the eq. (5):

$$r_{y_i\hat{y}_i} = \frac{\sum_{i=0}^{n}(y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}}_i)^2}} \qquad (5)$$

4. RESULTS AND MODEL COMPARISON

This section shows the results obtained from the different machine learning models and the detailed comparison between them by applying the established evaluation criteria (Table 2.).

For each of the models, the times in seconds of the training and scoring processes are taken into account, to take them into account as an additional parameter to put it into production in the future.

The dispersion diagrams of the different models are shown in (Fig.2). In the horizontal axis the true value "y" is represented, in the vertical axis the prediction "ŷ" of powers consumption. It is clear that the most correlated models are DLM and GBT.

In view of the results, three groups of models can be distinguished. A first group is positioned in a worse place and is formed by ANN, DT and RF. The three models show similar results, in terms of errors and correlations, DT and RF with better criteria measured against ANN. DT is the fastest of the three, closely followed by ANN.

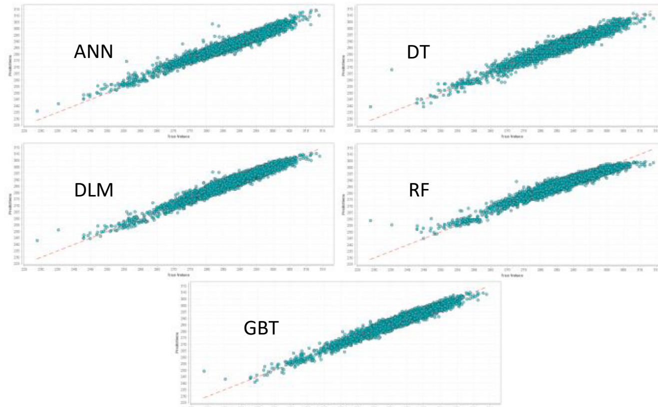

Fig.2. Scattergrams for different machine learning techniques in cryogenic LNG pumps.

The next one, positioned in the middle of all the models is DLM. Unlike the ANN, the results of the measured criteria are superior, but against this benefit, the training and production times increase considerably. As many authors comment in their research, this type of learning improves in every way compared to traditional learning methods at the expense of a longer time in carrying out the process.

It is apparent that the use of boosting improves significantly the classifiers' performance. DT algorithm with and without boosting are shown in the performance comparison. There is a 1% increase in the correlation of the results and a general decrease in the RMSE, MAE, MRE and MSE errors, however, the application of boosting leads to an increase in training and scoring time.

Our experiments suggest that the use of boosting can significantly improve the classification performance.

After selecting the best machine learning technique, it is time to implement its use to detect anomalies using operating variables. Figures 3 and 4 show the prediction results of the energy consumption and the actual energy consumption of the pump, and the cumulative square error is obtained as an indicator for two known situations of the pump condition. In figure 3, the state of the pump corresponds to "normal condition", where the behaviour of the variables is equivalent to a failureless and under normal conditions operation, for 15,000 hours of operation. In figure 4, the state of the same

TABLE 2. Machine learning evaluation criteria comparison

Model	Root Mean Squared Error		Absolute Error		Relative Error		Squared Error		Correlation		Training Time (s)	Scoring Time (s)
	RMSE	**Standard Deviation**	**MAE**	**Standard Deviation**	**MRE**	**Standard Deviation**	**MSE**	**Standard Deviation**	**Correlation**	**Standard Deviation**		
ANN	2,851	0,080	2,266	0,041	0,007	0,000	8,126	0,434	0,967	0,002	38	8
DLM	2,537	0,042	2,071	0,030	0,007	0,000	6,435	0,212	0,969	0,001	471	13
DT	2,744	0,077	2,135	0,077	0,007	0,000	7,534	0,422	0,962	0,002	11	3
RF	2,760	0,065	2,204	0,055	0,008	0,000	7,620	0,354	0,963	0,002	174	63
GBT	2,333	0,058	1,869	0,020	0,006	0,000	5,446	0,274	0,973	0,001	215	45

corresponds to "deteriorated condition", in which a period of approximately 5,000 hours of operation has been selected, where the pump was deteriorated, after a corrective maintenance intervention.

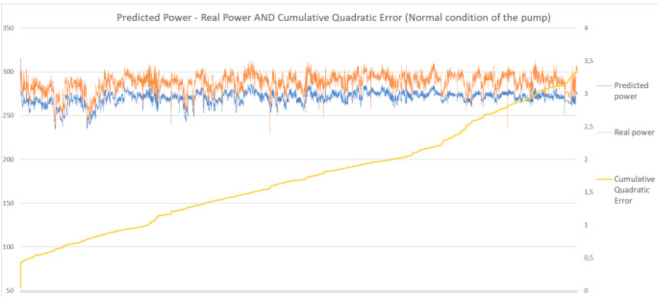

Fig.3. Scattergrams for different machine learning techniques in cryogenic LNG pumps.

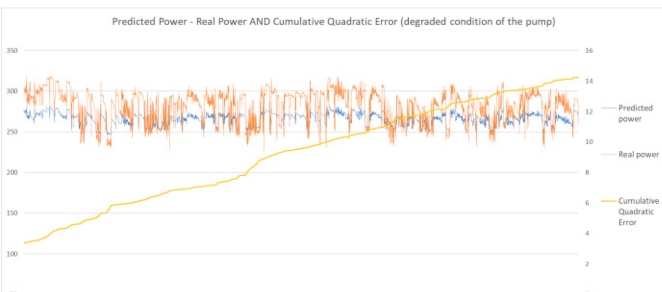

Fig.3. Scattergrams for different machine learning techniques in cryogenic LNG pumps.

It is clearly observed that the difference between prediction and actual value is greater when the pump operates in a degraded condition. The increase in the squared error accumulation becomes 7 times greater in degraded condition versus normal condition, for the same number of hours of operation.

5. CONCLUSIONS

In the problem statement, there is talk of the need to resort to machine learning techniques to detect anomalies related to aging or deterioration of the pump.

After a pre-processing of the variables, extracted from the information systems, different machine learning techniques are compared. Finally, the results of the model obtained are analysed by applying them in known periods with failures, demonstrating that it is possible to detect the anomaly early enough to be able to plan maintenance and reduce the overall impact on the business.

6. REFERENCES

Ahmad, J., Farman, H. and Jan, Z. (2019) ‘Deep Learning Methods and Applications’, in SpringerBriefs in Computer Science. Springer, pp. 31–42. doi: 10.1007/978-981-13-3459-7_3.

Barrios Castellanos, M. et al. (2020) ‘Fault identification using a chain of decision trees in an electrical submersible pump operating in a liquid-gas flow’, Journal of Petroleum Science and Engineering. Elsevier, 184, p. 106490. doi: 10.1016/J.PETROL.2019.106490.

Basheer, I. A. and Hajmeer, M. (2000) ‘Artificial neural networks: Fundamentals, computing, design, and application’, Journal of Microbiological Methods. doi: 10.1016/S0167-7012(00)00201-3.

Berry, M. J. A. and Linoff, G. (2004) Data mining techniques: for marketing, sales, and customer relationship management. Wiley.

Breiman, L. (2001) ‘Random forests’, Machine Learning, 45(1), pp. 5–32. doi: 10.1023/A:1010933404324.

Darmatasia and Arymurthy, A. M. (2017) ‘Predicting the status of water pumps using data mining approach’, in 2016 International Workshop on Big Data and Information Security, IWBIS 2016. Institute of Electrical and Electronics Engineers Inc., pp. 57–63. doi: 10.1109/IWBIS.2016.7872890.

Hameed, S. S., Vaithiyanathan, M. and Kesavan, M. (2019) ‘Fault Detection in Single Stage Helical Planetary Gearbox Using Artificial Neural Networks (ANN) and Decision Tree with Histogram Features’, in SAE Technical Paper. SAE International. doi: 10.4271/2019-28-0151.

Hill, T., Lewicki, P. and Lewicki, P. (2006) Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining. StatSoft. Available at: https://books.google.es/books?id=TI2TGjeilMAC.

Ian Goodfellow, Yoshua Bengio, and A. C. (2017) ‘Deep lLearning’, Genetic Programming and Evolvable Machines. Springer US, 19(1–2), pp. 305–307. doi: 10.1007/s10710-017-9314-z.

Kusiak, A., Zeng, Y. and Zhang, Z. (2013) ‘Modeling and analysis of pumps in a wastewater treatment plant: A data-mining approach’, Engineering Applications of Artificial Intelligence, 26(7), pp. 1643–1651. doi: 10.1016/j.engappai.2013.04.001.

Márquez, A. C., De la Fuente, A. and Antomarioni, S. (2019) ‘A Process to Implement an Artificial Neural Network and Association Rules Techniques to Improve Asset Performance and Energy Efficiency’, Energies 2019, Vol. 12, Page 3454. Multidisciplinary Digital Publishing Institute, 12(18), p. 3454. doi: 10.3390/EN12183454.

Saufi, S. R. et al. (2019) ‘Challenges and Opportunities of Deep Learning Models for Machinery Fault Detection and Diagnosis: A Review’, IEEE Access. Institute of Electrical and Electronics Engineers (IEEE), 7, pp. 122644–122662. doi: 10.1109/access.2019.2938227.

Tang, S., Yuan, S. and Zhu, Y. (2020) ‘Deep Learning-Based Intelligent Fault Diagnosis Methods Toward Rotating Machinery’, IEEE Access, 8, pp. 9335–9346. doi: 10.1109/ACCESS.2019.2963092.

Wang, W. et al. (2019) ‘Artificial neural network for the performance improvement of a centrifugal pump’, IOP Conference Series: Earth and Environmental Science, 240, p. 032024. doi: 10.1088/1755-1315/240/3/032024.

Wang, Yang et al. (2016) ‘Fault diagnosis for centrifugal pumps based on complementary ensemble empirical mode decomposition, sample entropy and random forest’, in Proceedings of the World Congress on

Intelligent Control and Automation (WCICA). Institute of Electrical and Electronics Engineers Inc., pp. 1317–1320. doi: 10.1109/WCICA.2016.7578401.
Yam, R. C. M. et al. (2001) 'Intelligent predictive decision support system for condition-based maintenance', International Journal of Advanced Manufacturing Technology. doi: 10.1007/s001700170173.
Zhang, G., Eddy Patuwo, B. and Y. Hu, M. (1998) 'Forecasting with artificial neural networks: The state of the art', International Journal of Forecasting. doi: 10.1016/S0169-2070(97)00044-7.