Failure diagnosis through Machine Learning

C. Ruiz Páez[1] , P. García Florencio[1].

Aviónica Avanzada

Escuela Técnica Superior de Ingeniería

[1]University of Seville Student

Seville, 2018

Abstract: Machine failures sometimes cause sudden and catastrophic damage to machinery. The prevention of these supposes great costs in maintenance, as well as in the replacement of complete pieces. With the monitoring of the operation of the machine these sudden failures can be prevented, or even extend the life of many pieces, with a minimum cost. This paper proposes an implementation of the machine learning theory, based on the detection of excitations of frequencies outside of normal operation. It can also be used to determine the type of failure and its severity. This approach has been validated using synthetic and real data, providing satisfactory results.

I. INTRODUCTION

Recent research has been focused on failure detection before it occurs or produces damage. The studies related to the operating an anomalous frequency offers the possibility to distinguish the correct operation from the malfunctioning, and the type of malfunction.

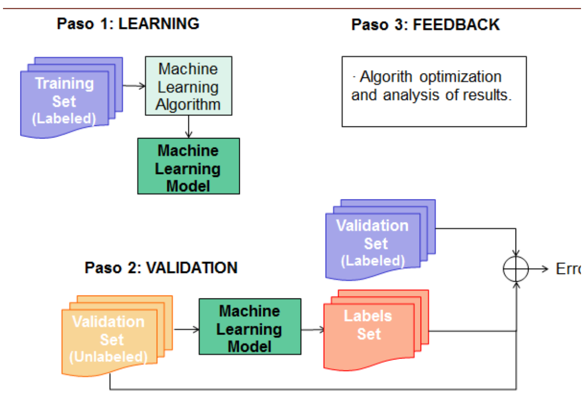

Fig 1. Machine learning approach.

In the last few years, the scientific study of algorithms and statistical models that computer systems use to progressively improve their performance on a specific task has increased [8][9].

Machine Learning algorithms build a mathematical model of sample data, training data set, in order to make predictions or decisions without being explicitly programmed to perform the task. Documentation from principles to algorithms can be found in references [1][2][3][5] and some applications [6][7].The purpose of the study is to test if it is possible to monitor the data frequency on real time using a ML algorithm to discern and discriminate the information obtained to verify the well-functioning or malfunctioning.

II. FAILURE MODES

Fig 2. Demagnetized rotor for data acquisition.

As a result of a research, the obtained data are divided into hydraulic, electric and mechanical failures. Each failure mode inside each of these type of failures has a different behaviour on the frequency spectrum. It is important to highlight that the failures shown are just some of them.

The typical mechanical failures are due to misalignment, imbalance, bent axle, friction and ball bearings damages. The most common electric failures are overheating, overpower, power loss, crown effect, interaction between rotor and stator. A special case of electrical failure is the demagnetization; because the real data provided come from this failure. Finally, the hydraulic failures are vortex, torch effect, cavitation and turbulences [4].

The frequency harmonics excited are different between them and, naturally, from the well-functioning frequency.

III. MACHINE LEARNING

Machine Learning is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on models and inference instead. It is seen as a subset of artificial intelligence. Nowadays there is a progressively use of the machine learning in all areas, not only on the engineering.

Big Data is one of the visible examples, the machine learning use of, identifies, through a complex algorithm model, the patterns of useful information of a huge amount of data, helping the behaviour analysis in economics, social media, IoT, etc.

IV. ML ALGORITHMS

Machine Learning algorithms build a mathematical model of sample data in order to make predictions or decisions without being explicitly programmed to perform the task.

In general terms, there are four categories of ML algorithms,

- Classification
- Regression
- Association
- Clustering

Classification algorithms are part of the supervised learning; such as Decision Trees, Support Vector Machines or K-Nearest Neighbours; estimates a classifier through on information from older observations, as a result of this information there is given a classifier, therefore the data is tagged, that is always being actualized from the early information. When the algorithms analyse the unseen data uses the classifiers to know to which class it belongs.

Regression algorithms are part of the supervised algorithms; like Support Vector Regression, Linear Regression or Logistic Regression; uses one or more continue variables and, based on the early observation, it predicts the behaviour providing a numerical number of the actual estimated value.

Association algorithms are part of the non-supervised algorithms; as A Priori, FP-Growth, Eclat Algorithm; is based on find relations in a big amount of data, this algorithm, as it is mention before, is used in Big Data.

Clustering algorithms are part of the non-supervised algorithms; for example, K-Means Clustering, Expectation Maximization or Hierarchical Clustering; use previous data, that there is no information about it, and try to find the common relationship; the common information is grouped into clusters, and the non-common data are not grouped into clusters.

The selected algorithm, after the evaluation, was the K-Nearest Algorithm (Classification), that is based on density tagged information. It provides a nominal and quantified tag.

VIBRATION SIGNAL	
MAXIMUM OPERATING VOLTAGE	2V
MAXIMUM TIME (s)	50
NOISE FUNCTION	Included
OUTPUT	Frequency Domain

Table 1. Signal information

In a n-dimensional space, the algorithm selects the

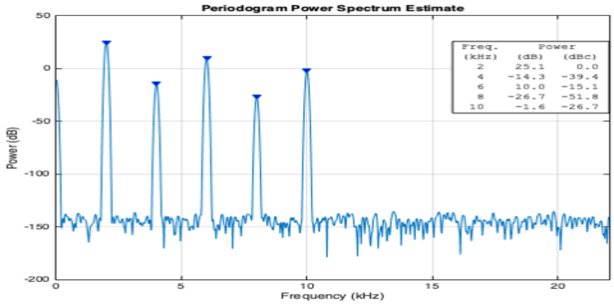

Fig 3. Period gram Power Spectrum Estimate.

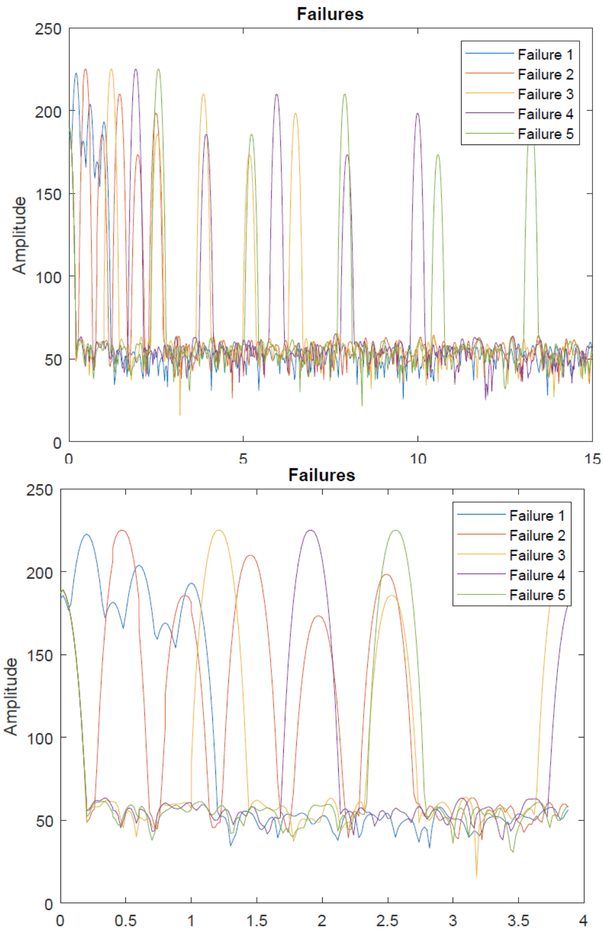

Fig 4. Synthetic data for simulation 1 and 2.

nearest k neighbours with similar behaviour and select the most repeated tag.

In this case, the n-dimensional space utilizes the frequencies and amplitudes data, and the classifier is the failure mode and the severity.

V. DATA ACQUISITION

The synthetic data was built through Matlab; every failure mode is related with a different phi value function in the frequency spectre shown in Equation (1) and the Table (1) parameters.

$$S = V \cdot \sin(2 \cdot \pi \cdot \varphi \cdot t) \qquad (1)$$

It was generated five types of failure modes with five different phi values. At the same time, one-hundred lectures were generated for the "training data" and perform the density function.

The real data was obtained from the Skylife laboratory of a real case of demagnetization. In order to perform a realistic analysis, there were different lectures with different speed velocities of a normal operation behaviour and from a demagnetization problem.

VI. IMPLEMENTATION AND OPTIMIZATION

The algorithm has been implemented in Matlab®, where the parameters of the algorithm are known through an optimization of the a priori results.

The simulations have been performed with synthetic data and real data. The data has been pre-processed to be useful for the algorithm. The frequency spectrum is discretized with a step of 0.02 Hz, having one variable for each frequency of the discrete spectrum. The algorithm calculates de distance or

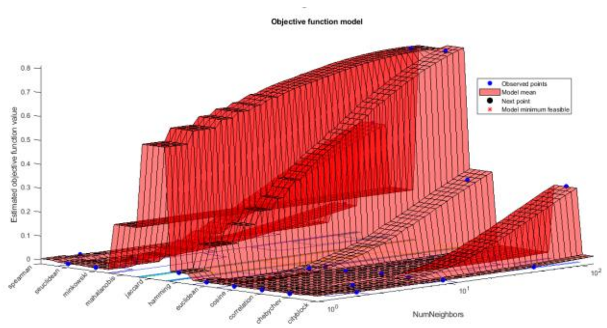

Fig 5. Algorithm optimization (1)

Iter	Eval result	Objective	Objective runtime	BestSoFar (observed)	BestSoFar (estim.)	NumNeighbors	Distance
1	Best	0	1.5498	0	0	2	spearman
2	Accept	0.8	0.72534	0	0.031809	92	jaccard
3	Accept	0	0.37718	0	0.016256	2	cityblock
4	Accept	0	0.38118	0	0.010925	8	euclidean
5	Accept	0.28	0.52508	0	0.0060839	95	euclidean
6	Accept	0	0.32757	0	-0.00025896	6	euclidean
7	Accept	0.004	0.32788	0	-0.00031859	44	cityblock
8	Accept	0	0.57747	0	-0.0062716	59	spearman
9	Accept	0	0.334	0	-0.005369	1	euclidean
10	Accept	0	0.52328	0	-0.0052675	11	seuclidean
11	Accept	0.092	0.41435	0	-0.0053886	124	seuclidean
12	Accept	0	0.36103	0	-0.0054051	1	seuclidean
13	Accept	0	0.47349	0	-0.0053947	8	chebychev
14	Accept	0.292	0.31073	0	-0.005281	125	chebychev
15	Accept	0	0.33449	0	-0.0052767	1	chebychev
16	Accept	0	0.36572	0	-0.0051685	2	minkowski
17	Accept	0.448	0.32884	0	-0.0047218	125	minkowski
18	Accept	0	0.36964	0	-0.0044724	1	cosine
19	Accept	0	0.34545	0	-0.0046978	54	cosine
20	Accept	0	0.3422	0	-0.00446	1	correlation

Fig 6. Algorithm optimization (2)

discrepancy between the actual signal and the training data set, and then it identifies the most likely state of functioning.

VII. RESULTS

The simulations done with the synthetic data have been done for two possible cases.

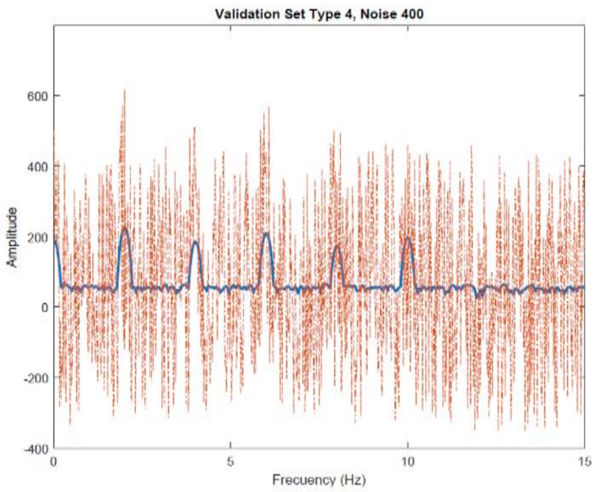

Fig 7. Validation set.

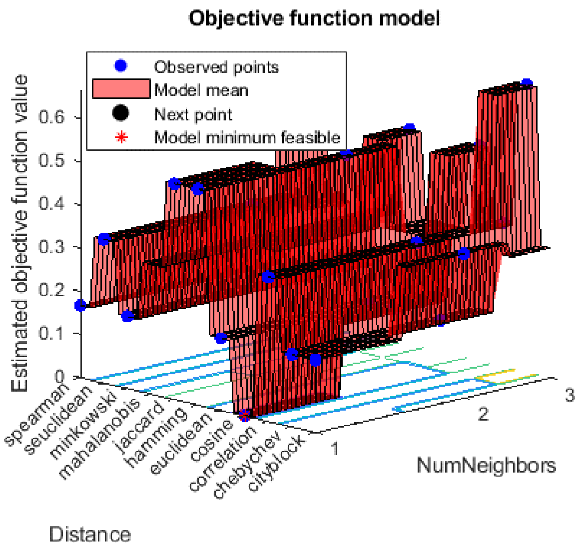

Fig 9. Real case optimization.

On the first one, the different failure modes have been well-differentiated. as it is shown on Figure (4)(a). However, in the second simulation they are not so easily differentiable, with similar frequency working and amplitudes, as we can see con Figure (4)(b).

The obtained data for each simulation is separated into training set (to build up the model) and validation set (to quantify its validity). Each one has a total of 250 signals.

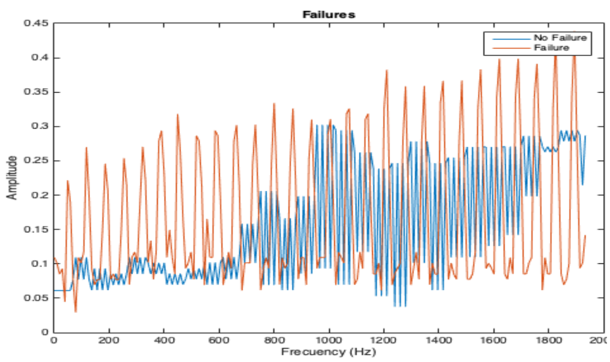

Fig 8. Real case data.

As we have already explained, two important parameters in the model of K-Nearest Neighbours are the number of neighbours (K) and the way to measure the distance or discrepancy between the current signal to be labelled and those of the model.

In the results, it can be observed that for 30 iterations, in simulation 1 with synthetic data, the expected optimum is an algorithm which uses the nearest 6 neighbours, calculating the Euclidean distance, in which the objective and runtime are minimum, as it can be found in Figure (5)(6).

In validation set have been added noise of different amplitude. In the Figure (7) is possible to observe the validation set failure signal without and with a random noise. The red one has noise with a maximum amplitude of 400, in which we observed that it is impossible to extract information at first sight.

The Table (2) shows the percentages of correct labelling for different noises in the signals even for optimized models of simulations 1 and 2. An acceptable performance is observed, being lower for simulation 2 as expected. It is also observed that the optimized model performs better for signals with a high noise level.

Finally, analysing the worst case, it can be observed the percentage of failure labelled signals, the correct percentage is on the diagonal. The most detectable failures are 3, 4 and 5, and that the most confused among them are 1 and 2.

Accuracy

	Simulation 1		Simulation 2	
Noise	K=1	K= optimized	K=1	K= optimized
-	100 %	100 %	100 %	100 %
100	100 %	100 %	100 %	88.4 %
200	100 %	100 %	82.8 %	82.8 %
300	100 %	98.4 %	70.4 %	75.2 %
400	92.4 %	91.2 %	54.8 %	68.8 %
500	82.0 %	88.4 %	50.0 %	58.0 %

Precision

		Predicted As				
		Failure 1	Failure 2	Failure 3	Failure 4	Failure 5
Actual	Failure 1	32	42	12	6	8
	Failure 2	36	32	20	8	4
	Failure 3	10	6	70	8	6
	Failure 4	4	0	16	66	14
	Failure 5	2	2	2	4	90

Table 2. Precision and accuracy.

Introducing the real data set in the ML Code, similar results were obtained. The Figure (8) shows the difference between normal working and demagnetization:

The model was optimizated during the ML process and the result of this process is shown at the figure (9). The best option for real data is K=1 and cosine mode. Finally, the accuracy obtained was 87.5%.

VIII. CONCLUSIONS

The machine learning algorithm for failure mode detection has been implemented successfully. The Real-time detection in the machine is possible due to the low execution times of the algorithm. It has been proven that the algorithm is robust when simulated with very noisy signals. It has also been proven that its accuracy increases as more differentiable the errors are. It has been shown that the more complex the algorithm is, the more precise it is with very noisy signals, being equivalent in the case of low noise signals. It has been proven that the scheme is valid for real data, obtaining acceptable results. However, the improvement of the real data extension will give us a better estimation.

REFERENCES

[1] Alex Smola, S. V. N. Vishwanathan, *"Introduction to Machine Learning"*, Departments of Statics and Computer Science Purdue University, College of Engineering and Computer Science Australian National University.

[2] Rui Pedro Paiva, *"Machine Learning: Applications Process and Techniques"*, University of Coimbra, Portugal, May 2013.

[3] Tom M. Mitchell, *"Machine Learning"*, McGraw-Hill Science / Engineer / Math, ISBN: 0070428077, March 1, 1997.

[4] *"Diagnóstico de daños"*, Dinámica de Sólidos. Máster en Diseño Avanzado en Ingeniería Mecáncia. Universidad de Sevilla.

[5] Shai Shalev-Shwartz, Shai Ben-David, *"Understanding Machine Learning From Theory to Algorithms"*, Cambridge University, 2014.

[6] F.T.K Au, Y.S. Cheng, L.G. Tham et. al. *"Structural damage detection based on a micro-genetic algorithm using incomplete and noisy modal test data".* Departament of civil engineering. University of Hong Kong.

[7] W.J. Staszewsky, S. Mahzan et. al. *"Health monitoring of aerospace composite structures",* Composite Science and Technology 69 (2009) 1678-1685.

[8] Eloi Figueiredo, Gyuhae Park, et. Al. *"Machine learning algorithms for damage detection under operational and environmental variability".* Structured health monitoring 10(6) 559-572.

[9] Young-Jin Cha, Wooram Choi, Deeo *"Learning-based crack damage detection using convolutional neural networks".* Computer-aided civil and Infraestructure Engineering 32 (2017) 361-378.