MISFIRE DETECTION ON INTERNAL COMBUSTION ENGINE THRORUGH DIFFERENT TYPES OF MACHINE LEARNING MODELS

Eduardo V.T. de M. Andrade[1], Antonio A. de S. Neto[2], Marcelo C. Rodrigues[3]

[1]Dept. of Mechanical Engineering, Federal University of Paraiba

edut0v0res2012@gmail.com

[2]Dept. of Mechanical Engineering, Federal University of Paraiba

antonioneto.mec@gmail.com

[3]Dept. of Mechanical Engineering, Federal University of Paraiba

marcelo.labii@gmail.com

Abstract. Misfire is a phenomenon that can jeopardize the good yield of an engines function, this might result in a lower efficiency than expected and an increase on pollution produced by elevated gas emission. However, there are certain systems that are capable of detecting this type of flaw. This article presents a model based on artificial intelligence that is capable to spot the misfire caused by a malfunction of the spark plug in a 2006 Zetec-Rocan ford motor that consists in a 4 stroke engine of internal combustion. As a result of its usage, identifying the main source of the problem becomes an easier task and therefore reducing the time spent on its maintenance. The model mentioned above uses vibration signals generated by an accelerometer. This signals went through a pre-processing proceedure to extract features in which were used a multiscreen analysis and FFT (Fast Fourier transform). After the extraction, those features were used as an entry on machine learning models that allow them to be classified according to its signal so we are able to identify if theres a defect and where the problem is located. The 5 machicne learning techniques used were Random Forest, Forest Tree, SVM(Support Vector Machine), KNN(K-Nearest Neighbors) and Neural Network. The results showed that they were all accurate both in train and in tests. External data validation also showed solid performance.

Keywords: Machine Learning, Combustion Engine, Misfire.

1 Introduction

According to SINDIPEÇAS, Brazil presents a fleet of over 46 million automobiles, with that being said, issues like ignition failure becomes more usual and affect not only the driver´s security but also the enviroment once this phenomenon reduces the engines efficency and increases the emission of polluting gas to the atmosphere. In this scenario, industries are investing in new technologies for a better detection of that failure to reduce costs and increase its precision.

The most commom method to diagnose ignition failures is vibration analysis [1] [2]. There are different ways to elaborate an vibration analysis in a system to detect an ignition failure using short term Fourier´s transformation Cavina et. A1 [3] accomplished an analysis on the time frequency of a signal vibrations processing all properties of velocity variantion due to ignition failure. Wong e Wong [4] accomplished a study that analyzed the ignition signal of the engine with Support Vector Machines (SVM) e Wavelet Packet Transform. A different way to analyze fibration while using Machine Learning was made by Sugumaran et al [5] using the algorithm J48 with Decision Tree methods and SVM. [15]

There are different methods of apprenticeship to process signals, in this article we will use Machine Learning method as they have the best precision and its data is transparent. We are able to measure a signal of vibrations aswell as other essencial features with an accelerometer.

From data acquired from an acquisition of vibration developed by the inspection and integrity lab of the Universidade Federal da Paraíba [6], apllied 6 differente types of Machine Learning: Support Vector Machine, Neural Network, Random Forest, Decision Tree e K-Nearest Neighbours, in order to compare the results accurancy and choose which one of them is the best method to determine of a better predictive maintenance using vibration analysis.

2 METHODS AND MATERIALS

2.1 FFT

Fourier´s transform is an important tool to process signals. It allows to transform a signal form a domain in time to a signal in a domain of frequency. For signals that underwent the sampling process the transform is given by:

$$X_k = \sum_{n=1}^{N-1} x_n e^{-i*\frac{2\pi}{N}*k_n} \tag{1}$$

Where N is the number os samples, k is the the current frequency, n is nth sample and x_n are complex numbers of the singal value at the moment t.

However, despite the widespread use of DWT, its more commom to use Fourier rapid transform, one of it´s variations have proven efficency and simplicity by multiple authors such as [7],[8] e [9]. This technique, separates DWT recursevily into smaller DWT ensuring a smaller computational resource during its application.

2.2 WAVELET´S DECOMPOSITION

Wavelet is an analysis of multiple resolution in frequency and time domains. It decomposes a signal in the domain of time in different resolutions, controlling the scale and translation, The decomposed signal possess good properties of location in both types of domain, as affirmed by [10], In top of that, with the use of Wavelet its possible to choose the type of transform function, known as Wavelet mother, from the signals entry chatacteristics which guarantees more flexibility during the analysis. Wavelet continuous transform (CWT) of a singal x(t) is given by:

$$CWT_x^{\Psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t)\Psi^*\left(\frac{t-\tau}{s}\right) dt \tag{2}$$

Where s is the parameter of scale , τ is the parameter of translation, * is the complext conjugate Ψ is Wavelet mother.

However, despite CWT applicability, its use requires a heavy computational resource. Therefore, Wavelet discrete transform is used, which is given by:

$$\Psi_{jk}(t) = \frac{1}{\sqrt{|s_0^j|}} \Psi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right) \tag{3}$$

Where j e k are whole numbers that control scale and translation, repespectively and τ_0 is the parameter of location.

In 1989, Mallat [11], developed an algorithm based in the signals decomposition using high and low pass filters. These filters allow the analysis of high and low components of frequency, repespectively. It is known as multiresolution analysis algorithm. This proceedure guarantees the approximate detalis of a signal in the domain of time. The approximations keep a general trend of the original signal e and the details describe its components in high frequency.

The importance of Machine Learning is rising day after day in the competitive market because it is capable to guarantee good solution to complex problems. Machine Learning is able to predict complex results that are boderline impossible for the human brain to resolve.

Machine Learning uses a range of algorithms for training data that can predict results. While the algorithms receive these trained data it is possible to fabricate various methods using this data with predicting results intent. ML is essential to analytical models for example, this happens because after the data is trained it provides an entry to a determined model and we receive and exit where its prediction is based on the data trained. [12]. In this article we will use 5 types of ML models: Support Vector Machine, Decision Tree, Random Forest, K-Nearest Neightboors e Neural Network.

2.3.1 SUPPORT VECTOR MACHINE (SVM)

SVM is a type of Machine Learning algorithm that was developed by a linear divisible model, its goal is to create an hyperplane in a space with N dimensions (where N= number of resources) to classify with high precision the data points. [13] In figure 2, we can take a glimpse in how it works:

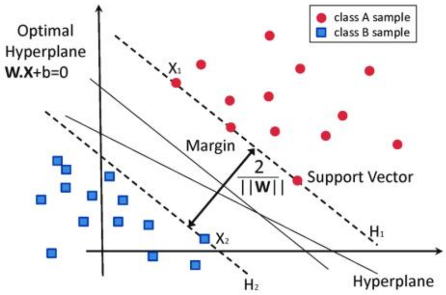

Figure 2 – Basic structure of a Support Vector Maching.

The prediction function of the distance of an hyperplane w to and support vector x is given by:

$$f(x) = < w, x > \; + \rho = w^T x + \rho \quad (4)$$

Where ρ is the trend.

2.3.2 DECISION TREE

Decision trees are a type of Machine Learning that develop diagrams to structure alternatives and results, its fundamental advantagem is that multiple types of classification are able to obtain as logic to various systems. [14]

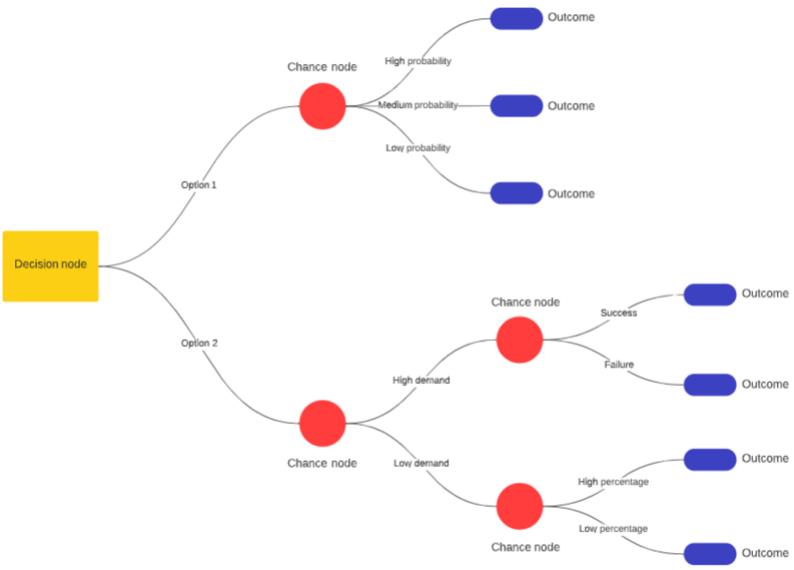

Figura 3 – Decision Tree strutcture of an algorithm.

Decision trees learning is given through a binary rcursive division of entry in space, i.e, in each stage of the flowchart the best predictor Xj and the best cutoff points are selected for a minor cost.

2.3.3 RANDOM FOREST

Random Forest is a type of machine learning based on the creation of various decision trees in order to combine all the decisions and come up with an more stable and a better accurancy prediction. [15]. This technique creats a system of decisions based in the entry of our machine learning.

2.3.4 K-NEAREST NEIGHBOR (KNN)

KNN is a machine learning algorithm that make prediction based on the apprenticeship of similar k instances, by its distance and synthesizing its exit based on neighbors samples that are determined by a set of training proposed by Fukunaga e Narenda em 1975. [16]

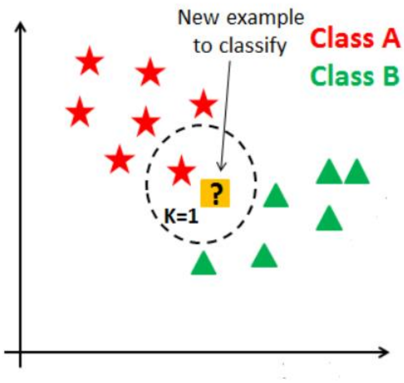

Figure 4 – KNN basic structure.

The distance calculaltion is very importanto to the KNN, the most used distance is Euclidean that is calculated by the equation below:

$$D_E(p,q) = \sqrt{(p_1 - q_1)^2 + \cdots + (p_n - q_n)^2} \tag{5}$$

Where p and q describe samples. In fig (4) p are the samples in blue while q are the samples in yellow.

In this article, the best value for k was found experimentally.

2.3.5 NEURAL NETWORK

Neural Network or Neural Artificial Network are paralel systems distributed as an complex structure associated by simple processing components, called neurons, that present a the capability to develop multiples types of operations at the same time for a better data processing, where the layers are linked by various connections. This connections, in most part, are assigned weights to come up with a better result, The physical structure is based in the human brain. [17]

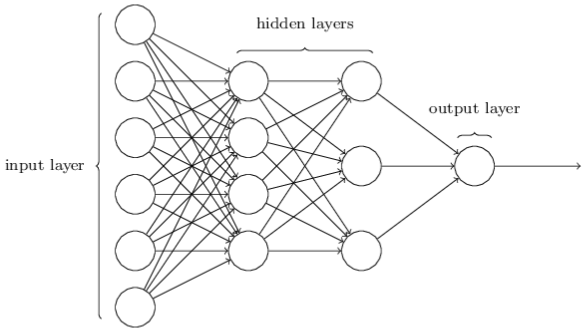

Figure 5 –Neural Artificial Network structure

2.3 EXPERIMENTAL SETUP AND DATA ACQUISITION

The tests were conducted in a Ford Zetec-Roman 2006 engine, that consists in a 4 stroke engine, with spark ignition and direct injection with maximum power of 77,6 kW (at 5500 rpm) and a maximum torque of 110,6 Nm (at 4250 rpm). The data was obtained at 900 rpm which is equivalent to a rotation frequency of 15 Hz and an explosion frequency of 30 Hz.

In this experiment the test was conducted in order to obtain vibration signals of the engine to identify possible ignition failures. Thus, it was used close to the engine block, where the combustion vibration can be detected more efficiently, an acquisition system of singals, like the one used by [6]. In this system, an accelerometer MPU6050 and an Arduino UNO plate put apart by, about, 30 centimeters above the block, The system has a sampling frequency of 512 Hz which is enough according to Nyquist theorem as the angular velocity of the engine was defined by 900 rpm measure read by the car´s panel. The system can be seen in the figure below:

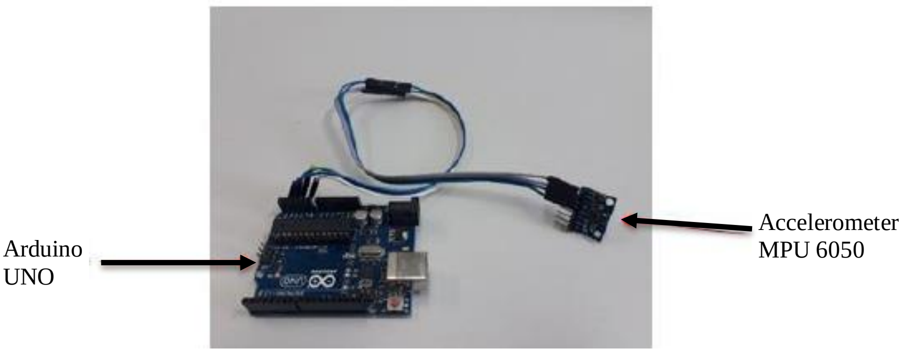

Figure 6 – Vibration Signal Acquistion System

.

In this experiment, each spark plug was disconnected individually from the cylinder in order to obtain a guaranteed failure. In top of that, other types of ignition failure were avoided, The dataset was collected in different days for each spark plug so that , in total, 30 signlas with 30 seconds duration for each type of failure induced. This dataset went through a preprocessing and went into machine learning algorithms for classification developed by this study.

3 RESULTS

After retrieving the data obtained by the acquisition system developed, a preproceedure is accomplished to extract the necessary features to the entry of the machine learning algorithms. All algoriths of machine learning used in this article were written in Python.

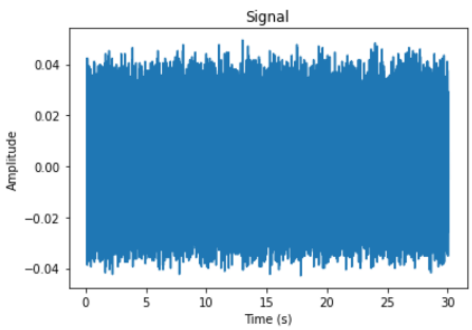

Figure 7 – Healthy Engine Vibration Signal

After the preprocessing stage, an FFT of the signal´s amplitude in time was obtained for decompositionin multiresolution of the signals. With this decompostion, the analysis of bigger amplitude levels that serve , with FFT, to extract features. Thus, 4 bigger types of peaks were used and the frequency corresponding to the FFT together with the 4 biggest peaks and frequencies corresponding to the biggest level. In the eighth level of detail, obtained through multiresolution analysis, summing for a total of 16 features for entry in algorithms. .

In the image below, its possible to notice the FFT and its frequencies of rotation and explosion of the engine with 13.53 e 27.07 Hz, respectively.

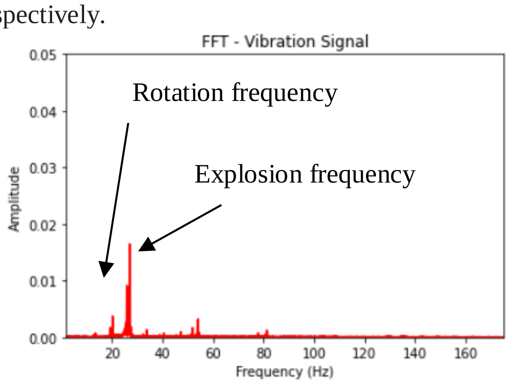

Figure 8 – FFT of the Vibration Signal of an Healthy Engine.

In the image below, its possibe to notice that each defect induced in a cylinder has a pattern. From this patterns that the features were extracted that were able to enalbe the ignition failure detection and analyze where it happened.

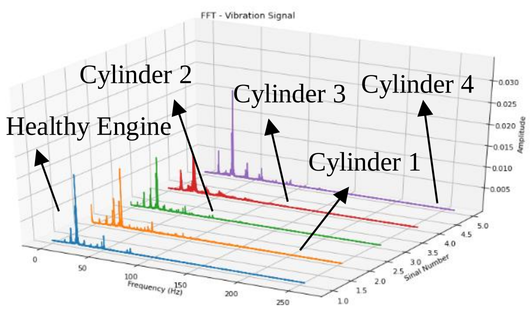

Figure 9 – Vibration signal FFT for a healthy engineand for an engine with ignition failure.

Thus, the resultdos obtained form Machine Learning algorithms are descbribed below..

Tabela 1. Resultdos obtained by Machine Learning

Algorithm	Parameters	Train accurancy	Test accurancy	Validation accurancy
SVM	C = 50	100,00%	96,67%	94,44%
Decision Tree	Max_depht = 3	100,00%	83,33%	88,89%
Random Forest	n_estimators = 100	100,00%	96,67%	94,44%
KNN	n_neighbors = 1	100,00%	93,33%	100,00%
RNA	1 Hidden layout with 70 neurons	94,81%	90,00%	94,44%

4 CONCLUSION

As proposed, this article brings a failure in ignition detection model based in artificial intelligence that allows to identify if there some type of flaw in the candles and in which on the flaw occurs. For that, the signals obtained through an accelometer, that after preprocessing were used as entry to 5 types of Machine Learning algorithms. As a result, all algorithms presented good accurancy both in training and tests and validation, presenting a reliable of detection. Furthermore, the best 2 results were given by Support Vector Machine and Random Forest with 96,67% and 94,44% accurancy in the validation, respectively.

References

[1] Lee J, Hwang S, Lim J, Jeon D and Cho Y. 1998. A new knock-detection method using cylinder pressure, block vibration and sound pressure signals from a SI engine SAE Paper 981436
[2] Chun K and Kim K 1994 Measurement and analysis of knock in a SI engine using the cylinder pressure and block vibration signals SAE Paper 940146
[3] Cavina, N., Corti, E., Minelli, G., Serra, G., 2002. Misfire detection based on engine speed time-frequency analysis. SAE Tech. Pap. https://doi.org/10.4271/2002-01-0480
[4] C.M. Wong, P.K. Wong, Engine ignition signal diagnosis with Wavelet Packet Transform and Multi-class Least Squares Support Vector Machines, Expert Syst. Appl. 38 (7) (2011) 8563–8570
[5] V. Sugumaran, K.I. Ramachandrana, S. Babu Devasenapatia, Misfire identification in a four-stroke four-cylinder petrol engine using decision tree, Expert Syst. Appl. 37 (3) (2010) 2150–2160
[6] Firmino, J.L., Neto, J.M., Oliveira, A.G., Silva, J.C., Mishina, K.V., Rodrigues, M.C., 2020. Misfire Detection of an Internal Combustion Engine Based on Vibration and Acoustic Analysis,
[7] Frederick, H., Winda, A., Solihin, M.I., 2019. Automatic petrol and diesel engine sound identification based on machine learning approaches. E3S Web Conf. 130.
[8] Glowacz, A., Glowacz, W., Kozik, J., Piech, K., Gutten, M., Caesarendra, W., Liu, H., Brumercik, F., Irfan, M., Faizal Khan, Z., 2019. Detection of Deterioration of Three-phase Induction Motor using Vibration Signals. Meas. Sci. Rev. 19, 241–249.
[9] Hagan, M.T., Demuth, H.B., Beale, M.H., 1995. Neural Network Design. Bost. Massachusetts PWS 2, 734.
[10] Nejad, F.H., Nourani, V., 2012. Elevation of wavelet denoising performance via an ANN based streamflow forecasting model. Int. J. Comput. Sci. Manag. Res. 1 (4), 764-770.
[11] Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern. Anal. Mach. Intell. 11 (7), 674-693.
[12] Hurwitz, Judith. Kirsh, Daniel. Machine learning for dummies. IBM Limited Edition. 2018.
[13] Liu, Wenyi & Zhenfeng, Wang & Jiguang, Han & Guangfeng, Wang. (2013). Wind turbine fault diagnosis method based on diagonal spectrum and clustering binary tree SVM. Renewable Energy. 50. 1–6.
[14] V. Sugumaran, K.I. Ramachandrana, S. Babu Devasenapatia, Misfire detection in a spark ignition engine using support vector machines, Int. J. Comput. Appl. 5 (6) (2010) 0975–8887.
[15] Breiman, L., Cutler, A., 2003. Setting up, Using, and Understanding Random Forests V4.0. Available at: http://www.stat.berkeley.edu/wbreiman/RandomForests/ (Acessado em 29/05/2021.).
[16] Fukunaga, K.; Narendra, P. M. A branch and bound algorithm for computing k-nearest neighbors. IEEE Transactions on Computers, v. 100, n. 7, p. 750–753, 1975.
[17] Braga, Antônio de Pádua. et al. Redes Neurais Artificiais: Teoria e Aplicações. 2 ed. LTC Editora. 2007.