

# Intelligent Incipient Fault Detection in Wind Turbines based on Industrial IoT Environment

Pedro H. Feijó de Sousa<sup>1</sup>, Navar de Medeiros M. e Nascimento<sup>1</sup>, Jefferson S. Almeida<sup>1</sup>, Pedro P. Rebouças Filho<sup>1</sup>, Victor Hugo C. de Albuquerque<sup>2</sup>

<sup>1</sup>Laboratório de Processamento de Imagens, Sinais e Computação Aplicada, Instituto Federal do Ceará, Fortaleza, CE, Brasil  
Email: pedrofeijo@ppger.ifce.edu.br; navar@lapisco.ifce.edu.br; jeffersonsilva@lapisco.ifce.edu.br; pedrosarf@ifce.edu.br

<sup>2</sup>Universidade de Fortaleza, Fortaleza, CE, Brasil

Email: victor.albuquerque@unifor.br

\*Corresponding Author: Victor Hugo C. de Albuquerque, Email: victor.albuquerque@unifor.br

How to cite this paper: Sousa, P. H. F.; Nascimento, N. M. M.; Almeida, J. S.; Rebouças Filho, P. P. and Albuquerque, V. H. C. (2019) Intelligent Incipient Fault Detection in Wind Turbines based on Industrial IoT Environment. Journal of Artificial Intelligence and Systems, 1, 1–19. <https://doi.org/10.33969/AIS.2019.11001>

**Received:** May 31, 2019

**Accepted:** June 13, 2019

**Published:** June 14, 2019

Copyright © 2019 by author(s) and Institute of Electronics and Computer. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



## Abstract

The eagerness and necessity to develop so-called smart applications has taken the Internet of Things (IoT) to a whole new level. Industry has been implementing services that use IoT to increase productivity as well as management systems over the past couple of years. Such services are now encroaching on wind energy, which nowadays is the most acceptable source among renewable energies for electricity generation. This work proposes an intelligent system to identify incipient faults in the electric generators of wind turbines to improve maintenance routines. Four feature extraction methods were applied to vibration signals, and different classifiers were used to predict the running status of the wind turbine. We correctly identified 94.44% of normal conditions, reducing the false positive and negative rates to 0.4% and 1.84%, respectively; a better result than other approaches already reported in the literature.

## Keywords

Intelligent IoT, Wind turbines, Feature Extraction, Vibration, Artificial Intelligence

## 1. Introduction

Recent trends in technology show that the Internet of Things (IoT) is a growing field in academic research and for real applications. Its potential usages are vast, ranging from low to hi-tech industries [31]. The number of connected devices will have tripled by 2020, and it will be surrounded by a trillion dollar global market [66]. Lee [33] classified IoT as an industrial Internet, since it is going to change today's computing paradigms and trigger new business models in order to satisfy industries and market needs.

The IoT is based on providing Internet for any kind of device, so they can easily exchange information between servers [27]. The purpose of IoT is to create automatic services, making these devices fully accessible to the operator; this is the so called smart device. A literature review made by Xu [67], showed that the number of papers published, surrounding IoT, increased 5 times in a four-year period. This growth in the field is due to the technical development in electronics, with miniaturization and power enhancement,

and computing with distributed processing improvements [57]. However, this global trend demands standards for production and usage, and this was perceived by Xu [67]. These authors indicated that the success of IoT relies on these standards, because they are going to make systems more interoperable and reliable.

IoT standardizations are being developed by international and highly qualified technical organizations, such as IEEE, International Organization for Standardization, and various others [4, 67], alongside the many countries that have invested in IoT development. In Europe, the United Kingdom invested more than US\$ 5 million in projects, while an international forum was created under the IoT European Research Cluster (IERC) to discuss the direction of development and to ensure proper funding [67]. Meanwhile, China wants to take IoT to a whole new level; they made US\$ 1.6 billion in grants and loans in 2015 and projections indicate their IoT manufacturing market is going to mobilize US\$ 127.5 billion in 2020 [8]. In the United States, companies like IBM indicate that the Internet of Things is a field of interest to improve technology infrastructure, and Japan also has shown strategies to create smart city standards [67]. Reviews in the literature indicate that (i) interests, concerns and investments in IoT are being applied all over the world, and (ii) China, a strong tech-developer, vouches this technology as one of the most promising, as can be seen from its large investments over the last few years and its future projections. Therefore, IoT is not only a trend, but is in fact being designed now for a future reality.

Diagnosing faults previously in wind turbines is very important to ensure a reliable and efficient operation, especially in offshore insulation. However, prior knowledge about the state of health and possible flaws that the wind turbine may have to succeed in the diagnosis is necessary [34]. Research uses data specific to wind turbine, for example, short circuit current [53], axial flow [13], and Vibration [35, 60, 41], used in this work. Others analyze the wind turbine dataset to detect failures [11, 70].

Roposing a new architecture of a Cyber-physical system (CPS) in the manufacture of modules of high Intensity Discharge (HID) cables, it used the technology Iot [32]. Due to multilocation and multi-product production, some factors limit process quality control and device application Iot aims to address this limitation by detecting system status and transmitting signals to the cloud. The author analyzes and classifies data using deep learning techniques. In the cloud, a process simulation is performed with data from the Iot devices and the results aid in real-time process decision making.

In other research, Lee [33] categorized IoT as (i) control and monitoring in terms of data collection from automated control systems to tracking variables and performance in real time, independent of location. This feature is important for technologies that require advanced control and monitoring such as power grids. (ii) Big data and business analysis is another category, and it is based on systems that allow to generate data in large scale, from sensors and / or usage information. These data provide relationships between different systems, enabling you to identify problems and develop business strategies, as well as providing customer-facing services. However, the system we will describe here, has both characteristics.

In order to increase reliability, predict failures, reduce operating / maintenance costs and prevent serious accidents that damage the ecosystem, this paper proposes a solution to a health monitoring system for wind generators using Iot technology. We present a framework for detecting incipient faults in wind turbines. Our system was developed based on an IoT infrastructure. A system, containing the framework, will analyze the vibration signal of the wind turbines and send information on the status of the wind turbines to devices connected to the network. More explicit information will be provided and in smaller time intervals, allowing technical managers to conduct preventive maintenance actions to improve the efficiency of wind farms, reduce operating costs and thus ensure cheaper electricity.

This paper is organized as follows: Section 2 presents a literature review of the main topics related to this work. In Section 3 all the details concerning the methodology used here are discussed. The results and discussions are presented in Section 4, which leads to

the conclusions in Section 5.

## 2. State of the art

In this section we present a literature review on the following topics: a wind energy overview in Section 2.1 and an IoT paradigm for wind farms in Section 2.2. The main faults that occur in wind turbines are discussed in Section 2.3. The feature extraction methods used to process vibration signals are exhibited in Section 2.4; and the machine learning methods used for classification are shown in 2.5.

### 2.1. Wind energy overview

The increase in the gross domestic product (GDP) is directly linked to industrial production and consequently economic growth in a country. The input generated in the electric energy industry influences the national economy due to its importance for the sectors of goods and services [69]. China is an example of a country where dependence on electricity is implicit in high output. The research conducted by Yuan [69], shows a positive correlation between energy production and the food industry, which is considered crucial, especially in densely populated countries. Renewable energy sources have emerged in the midst of energy needs, with the aim of reducing problems related to the burning of fossil fuels[48] and reducing the costs of generating energy by diversifying the energy matrix [58].

Countries with renewable energy sources strongly represented in their energy matrices demonstrate economic advantages, as reported by Valodka [64]. However, some countries have little land space for the installation of wind farms and the availability of natural resources is not favorable. So developing intelligent systems to manage and optimize the maximum power generation is important.

Also, according to the Global Wind Energy Council GWEC [21] wind energy, today, the most appropriate and stable source in the context of renewable energy sources. The global installed capacity reached 486.8 GW by the end of 2016 [21], illustrating the significant growth of the last years and projecting a promising scenario in 2030, where the wind energy will be able to supply up to 19% of the world's electricity needs[20].

Despite the present favorable growth of wind energy for investment, the implementation of a wind farm is a complex decision, since it is based on a trade-off between productive needs, economic interests, government policies and socio-environmental impacts. Sovacool [61] studied the energy market of the United States, and described the main impediments to include renewable energies in the energy matrix. The most important argument was related to the discontinuous nature of alternative sources, such as wind power. In contrast, Sovacool [61] showed evidence that the conventional energy sources also suffer variability of production and that wind energy when used on a large scale is extremely beneficial to the energy matrix. And an Internet of Things based system is important to manage multiple wind turbines by optimizing power generation and maintenance actions.

The very nature of wind generates power variability to the electric grid [26]. Katzenstein [26] analyzed the impacts on the final cost of electric power, according to the criterion of the variability in the power of wind farms. Compiling production data from 20 wind farms and energy price data, the authors noted that the price variability of up to US\$14/MWh, could be reduced to US\$3.16/MWh by increasing the capacity factor of a wind farm. Polinder [43] reports that 30% of the final cost of energy is due to operational costs of an offshore wind farm.

### 2.2. IoT system and operational costs of a wind farm

Kiviluoma [29] analyzed the behavior of the power variation in different regions and demonstrated that rapid changes in local wind conditions cause power fluctuations in the

order of 10% to 30%. Due to these characteristics, the economic viability concerns of a wind farm are complex and particular to each region, so studies are needed to promote methods to maximize the power extraction of a wind farm.

A reliable monitoring tool is one of the most efficient ways to mitigate operational costs of a wind farm, according to Asfani [2]. IoT systems play important roles in this field, because nowadays we have: efficient hardware for monitoring and data acquisition; high computer power for data processing, with cloud computing and infrastructure.

Several researchers analyze the benefits of the introduction of Internet devices in the energy industries. Faheem [16] optimized the smart grid integrations by applying algorithms and wireless sensor networks in the system. This area of research converges with another area, pattern recognition, built by several types of data-based methods with the objective of finding relevant information in large datasets. Qureshi [45] was able to predict wind energy by applying an algorithm based on artificial neural networks and Faheem [16] used it to forecast wind speed. Both pieces of research apply machine learning techniques to solve problems in wind energy.

Solid contributions from the IoT community have made this research possible, because there are various developments in microelectronic devices, standards and suggestions for data collection and analysis. Also, concepts of monitoring and management systems have established a solid basis to develop a solution that is integrated with local needs and aligned with global standards.

### 2.3. Faults on wind turbines

A report on data, made by Hahn, Durstewitz and Michael [22], has shown how the electrical generator is the most critical component in a wind turbine. A problem in this equipment decreases operational availability three times more than any other component. This also has an impact on the cost of energy, which is about 30% according to Polinder [43].

There is a conscious need to maintain the reliability of such equipment through predictive and preventive measures due to the high cost of generator maintenance when some failure occurs. Some authors claim that from 30% to 40% of induction motor failures are caused by short-circuits in the stator [30]. Early identification of these failures will reduce the cost of energy. In general failures are due to the bearings 69%; stator windings 21%; rotor bars 7% and shaft/coupling 3% [17].

One of the ways to detect short-circuit failures between stator turns is through current signature analysis, which characterizes the induction motor through current analysis and its frequency spectrum [63, 44, 18]. Another method is the vibration analysis [47, 24, 25], which is the focus of this work.

Vibration monitoring is one of the most widely used and most appropriate techniques for rotating machine analysis [5], so much so that it has been used in recent works as in Azizi [3], which used vibrational analysis for cavitation detection in centrifugal pumps. Stopa [62] presented a study of the frequency spectrum in vibration signals in rotary pumps that had failed. However, the alternative method to detect the cavitation phenomenon in centrifugal pumps proposed by the author failed to achieve results superior to the results presented by the vibration and pressure sensors methods typically used.

Ramalho [50] shows the vibration patterns of an electric motor submitted to operations with different types of loads and with misalignment that can be detected together with Wavelet decomposition and computational intelligence techniques. Some years later, Ramalho [51] also used a similar methodology, but for short-circuit detection between turns of stator coils in induction motors.

Royo [55] studied the application of the Fourier Transform (FT) in a squirrel cage induction motor with three different types of faults: broken rotor bars, short-circuits between turns and problems with the bearings. The authors used speed and current sensors for validation and concluded that it is possible to distinguish the three types of failure from the

normal situation using the FT. However, when the generator, some frequency components appear in the Fourier spectrum that may make it difficult to identify the normal conditions from the faults.

De Oliveira [12] used the FT, based on the frequency spectrum theory of Penman [42], and managed to identify 67% of short circuit failures, in 1.4% of turns, in an electric motor, using a current sensor.

## 2.4. Concepts of the feature extraction methods proposed

Applied in the digital processing of information, the Fourier transform is one of the leading mathematical tools, computationally implementable. The Fourier transform allows any function to be expressed in relation to the functions of the sine-base [15]. A peculiarity of the application of this transformation is that the function can be reconstructed with the inverse process and suffer the loss of information. This feature allows you to work on the Fourier spectrum and then return to the original spectrum of the function without losing signal information [19]. After the digital computers emerged and the creation of the Fast Fourier Transform (FFT) algorithm innovated the signal processing area [7]. There are several implementations of the Fourier transform, for example, analysis, filtering, reconstruction, and signal compression, as well as the extraction of characteristics for pattern recognition [19].

Higher-Order Statistics (HOS) describes time domain signals as opposed to the Fourier transform describing in the [37] frequency domain. Dwyer [14] was the first to propose the use of Kurtosis as a statistical tool to indicate non-Gaussian components in signals. This theory is the High-Order Statistics basis and it is described at the fourth statistical moment. However, Antoni [1] emphasizes in his research the efficiency of this method to characterize non-stationary signals. Mendel [37] pointed out that the advantage of HOS for signals depends on the robustness (i.e. filtering) to Gaussian noise when using moments higher than the second order. The features extracted from HOS are going to be skewness and kurtosis, alongside variance and RMS values.

The Structural Co-occurrence Matrix (SCM) performs a structural analysis of discrete signals, based on co-occurrence statistics, admitting existing connections between low-level structures of two discrete signals in n-dimensions [49]. The main characteristic of this extractor is to insert a prior knowledge in the analyzed signals, highlighting the detection of details. At the output, a two-dimensional histogram is generated, where the SCM presents the co-occurrences between the structures of the output signals [49]. In their article, the authors presented the calculations of the method and proposed 6 characteristics to be extracted from any signal in the entrance. In this work, SCM is the most contemporary extraction method and is able to recognize detailed patterns from [52] signals.

## 2.5. Concepts on machine learning approaches for classification

Multilayer Perceptron (MLP) stands out among several artificial neural networks because it presents powerful results in the modeling of input-output mappings typically encountered in function approximation (regression) and pattern classification problems. Formed by a simple system of interconnected neurons, the multilayer perceptron represents a non-linear mapping between an input vector and an output vector. The sum of the weight-weighted inputs and modified by an activation function results in the output signals, which are connected to the neurons through the weights [23].

Multilayer Perceptron (MLP) stands out among several artificial neural networks because it presents powerful results in the modeling of input-output mappings typically encountered function approximation (regression) and pattern classification problems.

The Bayesian classifier uses existing classes in data to indicate the class of the new samples, using probabilistic calculations [38]. Fisher [46] maximizes the posterior

probability and classifies an unknown sample to the highest posterior probability class. We can consider that in the Bayesian classifier the data is organized as a Gaussian probability density function. Assuming that all attributes are statically independent, we use a diagonal covariance matrix. The Naive-Bayes classifier uses the Russell [56] approach.

The Optimal Path Forest (OPF) classifier models the pattern recognition problem the partitions of a graph in a given feature space. In this one, nodes represent feature vectors and arcs connect all pairs of nodes, thus defining a complete graph [40]. This graph is formed by nos, which are characteristic vectors and by arcs that connect all pairs of nodes. Among the samples, there are arches with their respective weights. The weights vary according to the distance established between their resource vectors.

The Support Vector Machines (SVM) classifier uses an optimization process, based on the principle of structural risk minimization [65], establishing a separation function that respects a threshold between generation and overfitting during the classification. The nonlinearity cases in the decision surface can be solved by substituting the inner product of weights and vectors for a kernel function [9], applied to extend the concept of hyperplane-based classifiers for non-linear systems.

### 3. Methodology

In this study, we propose an IoT system architecture for wind farms as exhibited in Figure 1. A microcontroller with the framework and peripheral sensors, among them the accelerometer MEMs, are connected to the electric generator of the wind turbine and record the data of vibration; this gives intelligence to our status monitoring system, so we call it an intelligent system. The embedded system contains the **Feature Extractor** and the **Classifiers**, together they represent an intelligent decentralized application because they are able to identify the running status of the wind turbine generator.

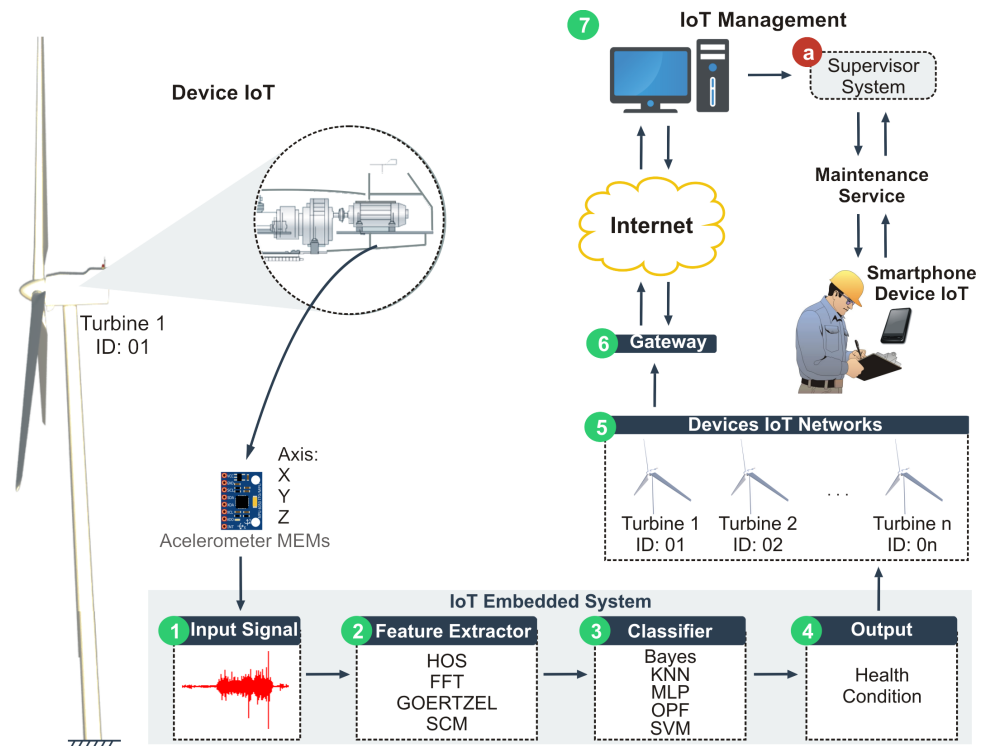
This information about the state of health of the system is visualized online by the responsible professionals, providing greater security and accuracy in the planning of predictive and corrective maintenance and at the ideal moment for execution, with these actions it is possible to reduce operating costs, costly failures and minimize displacements unnecessary maintenance teams, especially in offshore installations.

This paper proposes an intelligent decentralized application, from the database acquired by [53], we evaluate the performance of multiple classifiers on different feature extraction techniques, with the aim of identifying the best combination of extractor / classifier. With this information, the framework of fault detection is embedded in the system that emulates a wind turbine, which will be described later. The other parts of the system are part of our research as well. However, they are out of the scope of this paper and will be taken care of in future publications. The next steps outline the important steps so we can develop predict failures in wind turbines system.

#### 3.1. The wind turbine system

We used the same wind turbine test-bench as our previous work, as reported in Rebouças Filho [53]. The configurations consists of a wind turbine, using a Squirrel Cage Induction Generator, prepared to insert short circuits from the most incipient, 1.41% of the stator winding, to the more severe short applying with 9.26% of the amount of stator winding . And other Squirrel Cage Induction Generator to emulate the force of the wind. This generator has some advantages over other types, such as low maintenance costs and they represent 48% of all offshore installations [68, 71]. As we proposed before, the faults represent incipient conditions, in general long before total insulation degradation, i.e., thousands of seconds away. More details are provided in Rebouças Filho [53]. The faults considered in our work represent the increasing of the magnitude of a short-circuit as follows: HI-1, HI-2,





**Figure 1.** Diagram of proposed approach.

HI-3, LI-1, LI-2 and LI-3. The types of faults are characterize in high (HI) and low (LI) impedance.. The numbers 1, 2 and 3 are the percentage number of turns under short-circuits.

We imitate different conditions of wind-speed and energy generation, on the wind-turbine test-bench, following the procedures described in Rebouças Filho [53]. After the emulation of the acquisition system, there were 248 normal acquisitions and 1108 under failure conditions, totaling 1356 acquisitions. The faults were divided into 6 different types. In all experiments, the values of the frequency applied into the generator stator ( $f_g$ ) are registered, as well as the voltage in the dc-bus of the frequency converter on the generator side ( $V_{dc}$ ) and power generated ( $P$ ). The  $f_g$  ranges from 43.65 to 59.27 Hz,  $V_{dc}$  from 210 to 380 V.

### 3.2. The data acquisition system

The signals are acquired with a data acquisition system, developed using a National Instruments hardware, the NI-USB6009 module, programmed to read 10 seconds of a signal, sampled at 5 kHz, with 14-bits resolution. A LabVIEW interface running in a microcomputer was developed especially for this work.

The accelerometer transducer used was the MMA7360L, developed by Semiconductor [59]. It has a non linearity of 1%, sensitivity of 800mV/g, with a supply voltage of 3,3 V. This equipment is a Micro Electro Mechanical system (MEMs), which is a really small device (5 mm long), and is considered to be suitable to develop an IoT application for smart detection of incipient faults on wind turbines. The MEMs device is also able to measure three axis vibrations.

Other approaches in the literature have reported the use of this sensor to monitor rotating machinery, such as [50, 51]. The authors were able to characterize vibration patterns of load and misalignment in induction motors.

### 3.3. Parametrization of feature extraction methods to build a dataset

To compose our datasets we used three different feature extraction methods: a one time-domain method (i) the Higher-Order Statistics (HOS) and two frequency-domain methods (ii) the Fourier Transform and (iii) the Structural Co-occurrence Matrix (SCM).

In the Fourier transform we used the amplitude of frequencies  $0.5f_n$ ,  $1.5f_n$ ,  $2.5f_n$ ,  $3.0f_n$ ,  $5.0f_n$  and  $7f_n$  as features, wherein  $f_n$  is the fundamental frequency of the vibration signal. In the HOS, skewness and kurtosis, along with variance and RMS values are the features. In the SCM, we used the six original features proposed by Ramalho [49]. These feature extractors are applied to each of the axes. Furthermore we proposed using the dc-bus voltage,  $V_{dc}$ , and the generator frequency,  $f_g$ , measurements as features for each method. .

So, each feature extraction method resulted in one single dataset, and these were fed into different classifiers. We also, created combinations, *i.e.*, fusions, of these feature extraction methods, increasing the dimensionality of the problem. All datasets, and their dimensionalities are shown in Table 1. Further references in the text to any feature extraction method is followed by the tag shown in table 1. For example, if we wish to refer to the combination of Fourier and HOS as the feature extraction the initials FH shall be used.

All datasets<sup>1</sup> have seven classes, that is one Normal and six Faults. The number of samples are also the same, 1356, and they split into the following classes: a Normal condition with 248 samples, Fault HI-1 with 203, HI-2 with 179, HI-3 with 183, LI-1 with 177, LI-2 with 208 and LI-3 with 158 samples.

**Table 1.** All datasets created with the feature extraction methods, as well as their dimensionalities and tags used in the text.

Feature sets		Feature size	Tags
Single	Fourier	23	F
	Goertzel	23	G
	HOS	14	H
	SCM	17	S
Fusion	Fourier + Goertzel	46	FG
	Fourier + HOS	37	FH
	Fourier + SCM	40	FS
	Goertzel + HOS	37	GH
	Goertzel + SCM	40	GS
	HOS + SCM	31	HS
	Fourier + Goertzel + HOS	60	FGH
	Fourier + Goertzel + SCM	63	FGS
	Goertzel + HOS + SCM	54	GHS
	Fourier + HOS + SCM	54	FHS
Fourier + Goertzel + HOS + SCM	77	FGHS	

### 3.4. Classifier specifications and evaluation metrics

Similar to feature extractors, we adopted different types of classifiers with the objective of finding the best configuration for the task of detecting incipient short-circuits in wind-turbines, using vibration signals. Among the categories of existing techniques, we tested the Naive-Bayes classifier, which uses probability and statistics concepts and classifies the samples based on the probability density function. Its application in this case starts from the assumption that the variables are random and allow themselves to be modeled by a Gaussian

<sup>1</sup>Datasets are available at [https://github.com/navarmn/Wind\\_turbine\\_failure\\_prediction](https://github.com/navarmn/Wind_turbine_failure_prediction).



probability density function [54]. The classifier based on the Statistical Learning Theory, SVM, determines the class of samples with limits that increase the distance between them. Initially this technique was created by Cortes and Vapnik to solve binary problems. But its kernel machines are versatile and allow you to classify in a simpler way [39]. The MLP artificial neural network, inspired by the human neural network, it can be represented by a non-linear vector in the input vector and another vector in the output vector. Its learning takes place in the training phase, based on the backpropagation of the error [36]. The classifier using graph theory, OPF, is a non-parametric, multi-class, low cost computational classifier that uses a simple mathematical approach and is based on graph theory [40]. And a distance-based classifier, the KNN, uses one of the simplest and oldest methods for the classification pattern. This method classifies each unlabeled sample according to the majority of its nearest  $k$  in the training set [10].

To choose the best hyperparameters 10-fold cross-validation, with the random search, as proposed by Bergstra and Bengio [6] were used. Three types of kernels for the SVM classifier, the linear, polynomial and Radial Basis Function (RBF) were used. The range used for  $\gamma$  was  $[2^{-15}, 2^{33}]$  and  $[2^{-5}, 2^{15}]$  for  $C$ . For other types of kernels the same range as  $C$  was used. The MLP was trained with one hidden layer using ReLU activation function in all hidden neurons and softmax in the output layer, defined by the cross-entropy cost function, and optimized by the Adam algorithm [28]. In MLP, the number of neurons in the hidden layer ranged from 2 to 500. The Gaussian probability density function was applied in the Naive-Bayes classifier. The number of neighbors of KNN were tested within the range of 3 to 50. The OPF was tested with eight distance measures. The best configurations found for all classifiers and their tags are shown in Table 2.

After the feature extraction step, 15 datasets were created from the four extraction techniques. We preprocessed all databases equally. The training and the test datasets were normalized using the same measures (mean zero and unit variance). And to access the performance of all classifiers a Monte-Carlo simulation with fifty hold out cross-validation set to 80/20 was applied.

The metrics chosen to evaluate our model were accuracy and the confusion matrix. Accuracy (Acc) is measured in percentage and indicates which is the most promising result, ie, the most successful / extractor combination is the one with the nearest Accuracy of 100%. The confusion matrix, also measured in percentage, more specifically illustrates in which situation the extractor / classifier combination sets or misses the operation status of the generator. The highlighted line on the diagonal represents the hit percentage, that is, this line highlights the intersection of the correct operating statuses, the closer to 100% the more accurate the extractor / classifier combination was. And the sum of each line is set to 100 %, as it totalizes the total sample number of each generator status. The values highlighted in light gray are false negatives and dark gray are false positives. False negatives are when the classifier indicates that the status is faulted but in fact it is in normal operation. And the false positive is when classifier informs that it is working normal but in fact it is faulted.

The experiments were computed on a PC Intel i7 running at 3.1 GHz and 8Gb of RAM using a Linux Ubuntu operating system installed on a solid-state drive. We coded the extractors and the classifiers in the Python language.

## 4. Results and Discussions

Since the purpose of this work is to provide a IoT-based framework for smart systems on wind farms, we had to access the performance of different classifiers with different feature extraction methods. We would like to emphasize that all the results reported are on the test set, but the results on the training set were taken into consideration, as well. We paid attention to these results to evaluate overfitting of our models, but due to limitations of space we cannot exhibit them here.

Table 3 show the accuracy performance of all classifiers with the four feature extraction

**Table 2.** The best configurations found for all classifiers after a random search.

Classifier	Setups	Tags
Bayes	Gaussian	B
KNN	k=1	K1
KNN	k=3	K3
KNN	k=5	K5
MLP	IL/HL/OL	M1
MLP	IL/HL/OL	M2
OPF	Bray-Curtis	OB
OPF	Canberra	OC
OPF	Chi-Square	OCS
OPF	Euclidean	OE
OPF	Gaussian	OG
OPF	Manhattan	OMN
OPF	Squared Chi Squared	OSCS
OPF	Squared Chord	OSC
SVM	Linear	SL
SVM	Poly	SP
SVM	RBF	SR

methods used. Among the 17 classifiers, the four best combinations were HOS-SR, HOS-SP, HOS-Bayes and HOS-MLP2, with 94.44%, 93.49%, 91.76% and 91.20% accuracy, respectively. Among all the types of SVM, the RBF was better in all cases. The results using the SCM as the feature extractor obtained, in general, 10% less than HOS for this application. While using Fourier and Goertzel, the best combination was 15% less than the best with HOS. The HOS demonstrated to be the best among all these feature extraction methods. Also we achieved better results with the SR classifier than [53]. For the same problem they achieved 89.14% of overall accuracy.

The performance of SVM using RBF and the Polynomial kernel is quite low, so in further analyses we only used the RBF kernel. As we wanted to evaluate different machine learning paradigms, we choose SVM, Bayes and MLP for evaluations. Table 4 presents the confusion matrix of these classifiers. The sum in each line is 100%, so the values of the true positives represent the accuracy per class.

A result of 98.93% of all Normal conditions were correctly identified by HOS-SR, and this was the same with Bayes and MLP. A look into all the confusion matrices shows the accuracy on fault classes are, in general, lower than 80%. However, most of the misclassifications are the following: the HI-1 class was mostly mistaken for LI-1; HI-2 for LI-2 and HI-3 for LI-3. The pattern is the same in all classifiers. These results ratify the affirmative of Rebouças Filho [53], that knowing if a fault is of higher or lower impedance is more relevant to the classifier than the amount of coils turns under a short-circuit.

However, we must highlight that all faults were incipient, they are conditions very close to normal. Despite that, very few of them were classified as Normal. Since the problem we are evaluating in our work is to identify the status of the wind turbine, false positives and false negatives rates are more important than accuracy. False positives, *i.e.* type I error, represent normal conditions which are classified as fault, this represents an unnecessary cost to the wind farm, since the maintenance team has to stop the machine to inspect it. This error also gives discredit to the IoT expert system. On the other hand, false negatives, *i.e.* type II error, are fault conditions classified as Normal, which are particularly bad, because it means the system is mostly like to allow the wind turbine to operate under a short-circuit condition.

Taking these points under consideration, MLP is discard for further analyses as it has

more false negatives than Bayes and SVM. HOS-SR marked about 1.84% of false negatives, while HOS-Bayes marked, 16.48%. On the other hand HOS-Bayes produced less false positives than HOS- SR, with 0.4% and 1.07% respectively. So, HOS-Bayes is better to mitigate unnecessary maintenances and HOS-SR is better to prevent silent degradations.

The methodology proposed for mitigating the false positives and negatives used by [53] is based on ensemble learning of multiple MLPs, trained with different groupings of faults. The authors were able to identify 1.41% of turn short-circuits with 99.93% accuracy with less than 2% false negatives and 0% false positives. In our approach, the average accuracy of classes HI-1 and LI-1 was 81.69% for HOS-SR. But, for false negatives the HOS-Bayes was better, since it had only 0.4%, compared with the best results of [53] of 2%.

**Table 3.** Results of all classifiers for the feature extraction methods used. The metrics evaluated are accuracy (Acc) and the values in bold are the best configurations.

Classifiers	Extractors - Acc(%)			
	Fourier	Goertzel	HOS	SCM
B	77.77±0.14	76.06±0.19	<b>91.76±0.06</b>	84.89±0.22
K1	79.73±0.10	75.32±0.21	89.80±0.03	85.51±0.55
K3	78.88±0.25	76.11±0.57	89.84±0.54	85.68±0.54
K5	79.23±0.37	75.64±0.38	89.33±0.08	85.81±0.48
M1	78.55±0.66	76.74±0.28	88.54±2.83	84.87±2.73
M2	79.43±0.62	76.51±0.47	<b>91.20±1.42</b>	84.90±3.34
OB	76.25±0.48	75.87±0.56	75.53±0.51	78.72±0.58
OC	77.60±0.74	75.90±0.30	86.81±0.47	82.08±0.47
OCS	75.97±0.22	75.60±0.24	75.73±0.30	75.95±0.64
OE	79.23±0.53	76.02±0.53	89.28±0.71	85.00±0.48
OG	78.91±0.36	76.22±0.43	89.74±0.41	84.92±0.51
OMN	79.78±0.49	76.37±0.39	89.85±0.68	84.36±0.32
OCSC	78.19±0.33	76.02±0.40	88.24±0.65	84.56±0.25
OSC	75.49±0.38	75.26±0.54	74.75±1.12	76.12±0.31
SL	78.63±0.51	76.76±0.27	91.13±0.59	85.88±0.84
SP	80.29±0.58	76.46±0.57	<b>93.49±0.86</b>	85.81±0.73
SR	80.76±0.44	76.59±0.57	<b>94.44±0.36</b>	86.43±0.62

Nevertheless, in order to improve the system reliability we proposed a different approach to [53], who wanted to use the same feature extraction methods, but concluded that Fourier was more suitable for the application. However, they did not investigate multiple combinations of features. We raised the hypothesis that, combining multiple feature extractions might improve separability between classes, due to the high dimensionality, as discussed below.

#### 4.1. Results using multiple feature extraction

Since we intended to improve our system we will use the results until now as a baseline, in which case accuracies no better than 90% would be considered unacceptable for this application. We kept the same settings for the experiment as before, as describe in Section 3.4.

The attributes of the extractors were merged into pairs, in order to improve the accuracy of all classes. In the Table 5, the results of all the classifiers on these datasets are presented. The results of the fusion of the Fourier extractors with Goertzel were not included because the individual accuracy results of these extractors were not satisfactory when compared to the others.

The combination of the HOS and SCM extractors, which obtained the two best accuracy results in the individual results, with the SR classifier reached an accuracy of 94.56%,

**Table 4.** Average confusion matrix for the feature extractors HOS and SVM, Naive- Bayes and MLP classifier. The values are expressed in percentage, and the sum of each line is set to 100%. The values highlighted in light-gray are the false negatives and in dark-gray are the false-positives.

Predict/True	HOS-SR						
	Normal	HI-1	HI-2	HI-3	LI-1	LI-2	LI-3
Normal	<b>98.93%</b>	0%	0%	0%	0.27%	0.67%	0.13%
HI-1	0%	<b>82.46%</b>	2.30%	0%	9.84%	5.41%	0%
HI-2	1.11%	7.59%	<b>60.56%</b>	2.41%	5.74%	22.59%	0%
HI-3	0%	0%	1.09%	<b>73.27%</b>	0%	6.55%	19.09%
LI-1	0.74%	11.3%	7.04%	0%	<b>80.93%</b>	0%	0%
LI-2	0%	4.13%	13.17%	3.17%	0%	<b>78.89%</b>	0.63%
LI-3	0%	0%	0%	17.61%	0%	0.43%	<b>81.96%</b>
Predict/True	HOS-Bayes						
	Normal	HI-1	HI-2	HI-3	LI-1	LI-2	LI-3
Normal	<b>99.6%</b>	0%	0%	0%	0.40%	0%	0%
HI-1	4.43%	<b>62.62%</b>	11.48%	0.98%	15.57%	4.92%	0%
HI-2	6.30%	4.26%	<b>60.56%</b>	4.81%	0%	24.07%	0%
HI-3	0%	1.82%	1.09%	<b>52.73%</b>	0%	14.73%	29.64%
LI-1	4.81%	19.81%	14.81%	0%	<b>60.56%</b>	0.0	0%
LI-2	0.48%	9.21%	5.56%	4.76%	0%	<b>78.89%</b>	1.11%
LI-3	0%	0%	0%	25.22%	0%	2.71%	<b>72.61%</b>
Predict/True	HOS-MLP2						
	Normal	HI-1	HI-2	HI-3	LI-1	LI-2	LI-3
Normal	<b>96.00%</b>	0.53%	2.40%	0.13%	0.40%	0%	0.53%
HI-1	18.85%	<b>60.00%</b>	2.13%	0.66%	13.28%	4.92%	0.16%
HI-2	12.04%	4.63%	<b>58.89%</b>	3.70%	2.22%	17.96%	0.56%
HI-3	10.73%	0%	2.73%	<b>58.00%</b>	0.36%	7.82%	20.36%
LI-1	15.37%	11.11%	2.41%	0.19%	<b>70.93%</b>	0%	0%
LI-2	13.97%	3.81%	10.32%	6.35%	0.16%	<b>64.60%</b>	0.79%
LI-3	11.96%	0.22%	0.22%	18.48%	0.22%	1.52%	<b>67.39%</b>

slightly better than the result shown before. The SVM and MLP classifiers achieved a hit rate above 90% in all fusions where the HOS extractor was present. And, when the HOS extractor was not present in the fusions, the SVM and MLP classifier accuracy rate had an average of 87%, which represents a good value if compared to other classifiers, for example the OB, which presents an average of 75% accuracy.

The results of fusions of three and four features are shown in Table 6. We highlight the HOS results, because all datasets built with HOS maintained an accuracy rate above 90%, again in the SVM and MLP classifier. And, in the fusions where HOS was not present, the average accuracy of SVM and MLP classifiers reached 86.5%. The SR classifier achieved the highest hit rate among all results, with 94.09% when combined with the GHS fusion. And again, the HOS and SCM extractors are present in the most successful combination.

Considering the three tables presented, the best combinations of feature extraction method and classifiers are HOS-SR, with 94.44%, HS-SVM-RGB with 94.56% and GHS-SR with 94.09%. The SR classifier is present in the best results in all the fusions of extractors, and the HOS extractor was also present in the best results. However, the best result was when the HOS and SCM extractors were merged with the SR classifier.

Despite the results discussed above, the combination of multiple feature extraction methods was not conclusive for us. So, we cannot affirm that it improved the performance of the classifier.

**Table 5.** Results of all classifiers on datasets created by the fusion of feature extractors in pairs. The metrics evaluated are accuracy (Acc) and the values in bold are the best pairs.

Classifiers	Fusion of 2 Extractors				
	FH	FS	GH	GS	HS
B	89.52±0.48	84.07±0.27	76.90±0.45	76.87±0.30	76.89±0.52
K1	85.39±0.53	82.37±0.75	87.64±0.43	82.50±0.27	88.62±0.72
K3	85.36±0.38	83.28±0.03	86.97±0.42	86.71±0.29	88.84±0.06
K5	85.23±0.22	82.69±0.02	87.50±0.24	83.53±0.18	88.84±0.10
M1	91.37±0.87	86.31±0.57	<b>90.96±0.55</b>	90.96±0.55	<b>92.40±0.95</b>
M2	92.08±0.56	86.42±0.70	<b>92.02±0.80</b>	85.63±0.78	<b>93.43±0.42</b>
OB	77.46±0.76	77.73±0.57	76.81±0.22	77.77±0.39	80.81±0.22
OC	82.45±0.60	80.82±0.50	82.53±0.55	79.65±0.37	86.31±0.56
OCS	75.24±0.73	75.02±0.42	75.23±0.48	75.15±0.39	76.52±0.53
OE	85.91±0.73	82.53±0.53	86.54±0.54	82.45±0.55	88.72±0.77
OG	85.89±0.65	82.45±0.59	86.49±0.57	82.43±0.59	89.18±0.26
OMN	87.81±0.61	83.24±0.67	87.37±0.29	82.53±0.19	88.88±0.33
OCSC	83.82±0.43	81.35±0.29	84.68±0.42	81.23±0.42	88.38±0.64
OSC	73.31±0.32	75.20±0.62	73.36±0.41	75.48±0.57	74.90±0.61
SL	91.33±0.67	85.90±0.54	90.74±0.53	86.09±0.61	91.22±0.46
SP	<b>92.75±0.48</b>	87.61±0.66	<b>93.05±0.65</b>	87.57±0.55	<b>93.92±0.26</b>
SR	92.29±0.59	87.65±0.66	<b>93.67±0.50</b>	88.48±0.42	<b>94.56±0.40</b>

**Table 6.** Results of all classifiers on datasets created by the Fusion of three and four extractors. The metrics evaluated are accuracy (Acc) and the values in bold are the best combinations.

Classifiers	Fusion of 3 and 4 Extractors				
	FGH	FGS	GHS	FHS	FGHS
B	75.95±0.48	76.60±0.65	85.28±2.21	77.05±0.59	84.33±2.20
K1	84.14±0.37	82.54±0.62	86.97±0.26	85.73±0.46	84.74±0.66
K3	84.59±0.52	82.91±0.61	87.40±0.15	85.67±0.48	84.63±0.33
K5	84.80±0.84	82.36±0.44	87.46±0.26	85.43±0.23	85.22±0.37
M1	90.69±0.74	86.06±0.36	92.49±0.60	92.10±0.44	91.57±0.62
M2	91.22±0.38	86.20±1.01	92.39±0.48	92.42±0.69	92.16±0.54
OB	77.63±0.38	77.92±0.46	79.64±0.40	78.99±0.76	78.60±0.28
OC	81.38±0.29	79.97±0.54	83.57±0.57	83.11±0.35	82.34±0.48
OCS	75.58±0.59	75.86±0.45	75.79±0.49	75.68±0.42	75.36±0.61
OE	84.22±0.65	81.34±0.47	87.03±0.66	86.34±0.44	85.37±0.38
OG	84.26±0.59	81.56±0.31	86.49±0.58	80.33±0.35	85.12±0.60
OMN	86.11±0.47	82.75±0.89	87.77±0.47	87.70±0.48	86.93±0.53
OCSC	82.96±0.37	80.54±0.66	85.39±0.49	84.53±0.65	83.63±0.65
OSC	73.62±0.41	75.51±0.35	74.45±0.28	74.72±0.26	74.83±0.39
SL	90.96±0.71	85.53±0.49	91.12±0.54	90.96±0.38	90.71±0.49
SP	<b>93.14±0.73</b>	87.58±0.65	93.64±0.42	<b>93.04±0.57</b>	<b>92.94±0.62</b>
SR	92.47±0.51	87.36±0.58	<b>94.09±0.33</b>	92.87±0.69	92.74±0.23

## 5. Conclusions

We presented a IoT-based framework for wind farms to improve their maintenance routines. Our system is based on an intelligent decentralized application. In this paper we showed a design of different strategies to identify incipient faults of short-circuits in wind turbines.

Vibration signals were demonstrated to be suitable for this purpose, compared to other approaches in the literature that use electrical current. We were able to identify Normal conditions of the wind turbine with 98.93% of accuracy. We also, reduced the false positives rate to 0.4% and false negatives to 1.84%.

Among all combinations HOS is highlighted as the best approaches for this task. SVM with the RBF kernel is the best approaches, however the Naive-Bayes and MLP classifiers are also relevant. In future works we aim to combine an ensemble of multiple learning,

combining multiple information from different sensors. Another goal for future work is to apply this structure, based on Iot, to real proportions and considering the challenges of this technology, for example, type of connectivity, which is the power source of reliable actuators and sensors and the standardization of networks and sensors for secure integration. Install a microcontroller and its peripherals in a wind turbine and set the framework to identify incipient short-circuit failures and distribute the status information of the wind turbine to devices connected to the network in the shortest time. A review of the devices will be necessary to identify the limitations and advantages of the types of connections, for example, RFID, Wi-fi, Bluetooth and ZigBee, and to allow connection with other technologies such as GSM, 3G and LTE.

## 6. Acknowledgments

VHCA received support from the Brazilian National Council for Research and Development (CNPq, grants no. 304315/2017-6 and 430274/2018-1).

## Conflicts of Interest

There is no conflict of interest.

## References

- [1] Jérôme Antoni. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2):282–307, 2006. ISSN 08883270.
- [2] DA Asfani, MH Purnomo, and DR Sawitri. Naïve bayes classifier for temporary short circuit fault detection in stator winding. In *Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED), 2013 9th IEEE International Symposium on*, pages 288–294. IEEE, 2013.
- [3] Raziye Azizi, Behrooz Attaran, Ali Hajnayeb, Afshin Ghanbarzadeh, and Maziar Changizian. Improving accuracy of cavitation severity detection in centrifugal pumps using a hybrid feature selection technique. *Measurement*, 108:9–17, 2017.
- [4] Debasis Bandyopadhyay and Jaydip Sen. Internet of things: Applications and challenges in technology and standardization. *Wireless Personal Communications*, 58(1):49–69, 2011.
- [5] Raymond S Beebe. *Predictive maintenance of pumps using condition monitoring*. Number April. 2004.
- [6] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, 13:281–305, February 2012. ISSN 1532-4435.
- [7] E Oran Brigham and E Oran Brigham. *The fast Fourier transform and its applications*, volume 448. prentice Hall Englewood Cliffs, NJ, 1988.
- [8] China Telecom Americas. 5 things you need to know about IoT in China. Technical report, 2017.
- [9] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [10] Thomas M Cover, Peter E Hart, et al. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.



- [11] Phong B Dao, Wieslaw J Staszewski, Tomasz Barszcz, and Tadeusz Uhl. Condition monitoring and fault detection in wind turbines based on cointegration analysis of scada data. *Renewable Energy*, 116:107–122, 2018.
- [12] Atila Girao De Oliveira, Ricardo Silva Thé Pontes, and Cláudio Marques de Sá Medeiros. Neural network used to stator winding interturn short-circuit fault detection in an induction motor driven by frequency converter. In *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*, pages 459–464. IEEE, 2013.
- [13] Pedro Henrique Feijó de Sousa, Navar Medeiros M e Nascimento, Pedro Pedrosa Rebouças Filho, and Cláudio Marques Sá de Medeiros. Detection and classification of faults in induction generator applied into wind turbines through a machine learning approach. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2018.
- [14] R. Dwyer. Detection of non-Gaussian signals by frequency domain Kurtosis estimation. *ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 8:607–610, 1983.
- [15] Okan K Ersoy. A comparative review of real and complex fourier-related transforms. *Proceedings of the IEEE*, 82(3):429–447, 1994.
- [16] M Faheem and V C Gungor. Energy efficient and QoS-aware Routing Protocol for Wireless Sensor Network-based Smart Grid Applications in the Context of Industry 4.0. *Applied Soft Computing Journal*, 2017.
- [17] Fiorenzo Filippetti, Alberto Bellini, and Gerard-Andre Capolino. Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives. In *Electrical Machines Design Control and Diagnosis (WEMDCD), 2013 IEEE Workshop on*, pages 196–209. IEEE, 2013.
- [18] Xiang Gong and Wei Qiao. Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals. *IEEE Transactions on Industrial Electronics*, 60(8): 3419–3428, 2013.
- [19] Rafael C Gonzalez, Richard E Woods, Steven L Eddins, et al. *Digital image processing using MATLAB.*, volume 624. Pearson-Prentice-Hall Upper Saddle River, New Jersey, 2004.
- [20] GWEC. Wind Energy Outlook: 2000 gigawatts by 2030, 2017. URL <http://www.gwec.net/wind-energy-outlook-2000-gigawatts-2030/>.
- [21] GWEC. Global Wind Report 2016. Technical report, Global Wind Energy Council, 2017. URL <http://files.gwec.net/files/GWR2016.pdf>.
- [22] Berthold Hahn, Michael Durstewitz, and Kurt Rohrig. Reliability of Wind Turbines. *Wind Energy*, pages 1–4, 2007.
- [23] Simon Haykin. Neural networks: principles and practice. *Bookman*, 2001.
- [24] David He, Ruoyu Li, and Junda Zhu. Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Transactions on Industrial Electronics*, 60(8): 3429–3440, 2013.
- [25] Fabio Immovilli, Alberto Bellini, Riccardo Rubini, and Carla Tassoni. Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Transactions on Industry Applications*, 46(4):1350–1359, 2010.

- [26] Warren Katzenstein and Jay Apt. The cost of wind power variability. *Energy Policy*, 51:233–243, 2012.
- [27] Iman Khajenasiri, Abouzar Estebarsari, Marian Verhelst, and Georges Gielen. A Review on Internet of Things Solutions for Intelligent Energy Control in Buildings for Smart City Applications. *Energy Procedia*, 111(September 2016):770–779, 2017.
- [28] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.
- [29] Juha Kiviluoma, Hannele Holttinen, David Weir, Scharff Richard, Söder Lennart, Menemenli Nickie, A. Cutululis Nicolaos, Danti Lopez Irene, Lannoye Eamonn, Estanqueiro Ana, Gomez-Lazaro Emilio, Zhang Qin, Bai Jianhua, Wan Yih-Huei, and Michael Milligan. A simple atmospheric boundary layer model applied to large eddy simulations of wind turbine wakes. *Wind Energy*, 17(April 2013):657–669, 2015.
- [30] GB Kliman, WJ Premerlani, RA Koegl, and D Hoeweler. A new approach to on-line turn fault detection in ac motors. In *Industry Applications Conference, 1996. Thirty-First IAS Annual Meeting, IAS'96., Conference Record of the 1996 IEEE*, volume 1, pages 687–693. IEEE, 1996.
- [31] Vlad Krotov. The Internet of Things and new business opportunities. *Business Horizons*, 2017.
- [32] Hyunsoo Lee. Framework and development of fault detection classification using iot device and cloud environment. *Journal of Manufacturing Systems*, 43:257–270, 2017.
- [33] In Lee and Kyoochun Lee. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4):431–440, 2015.
- [34] Gustavo de Novaes Pires Leite, Alex Maurício Araújo, and Pedro André Carvalho Rosas. Prognostic techniques applied to maintenance of wind turbines: a concise and specific review. *Renewable and Sustainable Energy Reviews*, 81:1917–1925, 2018.
- [35] WY Liu, WH Zhang, JG Han, and GF Wang. A new wind turbine fault diagnosis method based on the local mean decomposition. *Renewable Energy*, 48:411–415, 2012.
- [36] Aldísio G Medeiros, Solon A Peixoto, Antonio Carlos S Barros, Victor Hugo C de Albuquerque, and Pedro P Rebouças Filho. Uma nova abordagem para a segmentação de pulmões utilizando o método de contorno ativo não paramétrico optimum path snakes em imagens de tomografia computadorizada. In *17<sup>o</sup> Workshop de Informática Médica (WIM 2017)*, volume 17. SBC, 2017.
- [37] J M Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3):278–305, 1991.
- [38] Nasser M Nasrabadi. Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4):049901, 2007.
- [39] Edson Cavalcanti Neto, Samuel Luz Gomes, Pedro Pedrosa Rebouças Filho, and Victor Hugo C de Albuquerque. Brazilian vehicle identification using a new embedded plate recognition system. *Measurement*, 70:36–46, 2015.
- [40] Joao P Papa, Alexandre X Falcao, and Celso TN Suzuki. Supervised pattern classification based on optimum-path forest. *International Journal of Imaging Systems and Technology*, 19(2):120–131, 2009.

- [41] Cédric Peeters, Patrick Guillaume, and Jan Helsen. Vibration-based bearing fault detection for operations and maintenance cost reduction in wind energy. *Renewable Energy*, 116:74–87, 2018.
- [42] J Penman, HG Sedding, BA Lloyd, and WT Fink. Detection and location of interturn short circuits in the stator windings of operating motors. *IEEE transactions on Energy conversion*, 9(4):652–658, 1994.
- [43] Henk Polinder, Jan Abraham Ferreira, Bogi Bech Jensen, Asger B Abrahamsen, Kais Atallah, and Richard a. McMahon. Trends in Wind Turbine Generator Systems. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 1(3):174–185, 2013. ISSN 2168-6777.
- [44] Miguel Delgado Prieto, Giansalvo Cirrincione, Antonio Garcia Espinosa, Juan Antonio Ortega, and Humberto Henao. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 60(8):3398–3407, 2013.
- [45] Aqsa Saeed Qureshi, Asifullah Khan, Aneela Zameer, and Anila Usman. Wind power prediction using deep neural network based meta regression and transfer learning. *Applied Soft Computing Journal*, 58, 2017.
- [46] MA RA Fisher. On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. Lond. A*, 222(594-604):309–368, 1922.
- [47] Mona Khatami Rad, Mohammadehsan Torabizadeh, and Amin Noshadi. Artificial neural network-based fault diagnostics of an electric motor using vibration monitoring. In *Transportation, Mechanical, and Electrical Engineering (TMEE), 2011 International Conference on*, pages 1512–1516. IEEE, 2011.
- [48] Ehsan Rahimi, Abdorreza Rabiee, Jamshid Aghaei, Kashem M. Muttaqi, and Ali Esmaeel Nezhad. On the management of wind power intermittency. *Renewable and Sustainable Energy Reviews*, 28(x):643–653, 2013.
- [49] Geraldo L Bezerra Ramalho, Daniel S Ferreira, Pedro P Rebouças Filho, and Fátima N Sombra de Medeiros. Rotation-invariant feature extraction using a structural co-occurrence matrix. *Measurement*, 94:406–415, 2016.
- [50] Geraldo Luis Bezerra Ramalho, Pedro Pedrosa Rebouças Filho, Celso Rogerio Schmidlin Júnior, and Samuel Vieira Dias. Detecção de falhas através de características do sinal de vibração e rede SOFM. *XI Simpósio Brasileiro de Automação Inteligente, 2013, Fortaleza-CE.Simpósio Brasileiro de Automação Inteligente 2013 (SBAI 2013)*, 2013. ISSN 1098-6596.
- [51] Geraldo Luis Bezerra Ramalho, Adriano Holanda Pereira, Pedro Pedrosa Rebouças Filho, and Cláudio Marques de Sá Medeiros. Detecção De Falhas Em Motores Elétricos Através Da Classificação De Padrões De Vibração Utilizando Uma Rede Neural Elm. *Holos*, 4(0):185, 2014.
- [52] Pedro P Rebouças Filho, Elizângela de S Rebouças, Leandro B Marinho, Róger M Sarmiento, João Manuel RS Tavares, and Victor Hugo C de Albuquerque. Analysis of human tissue densities: A new approach to extract features from medical images. *Pattern Recognition Letters*, 94:211–218, 2017.
- [53] Pedro Pedrosa Rebouças Filho, Navar MM Nascimento, Igor R Sousa, Cláudio MS Medeiros, and Victor Hugo C de Albuquerque. A reliable approach for detection of incipient faults of short-circuits in induction generators using machine learning. *Computers & Electrical Engineering*, 71:440–451, 2018.

- [54] Murillo B Rodrigues, Raul Victor M Da Nóbrega, Shara Shami A Alves, Pedro Pedrosa Rebouças Filho, Joao Batista F Duarte, Arun K Sangaiah, and Victor Hugo C De Albuquerque. Health of things algorithms for malignancy level classification of lung nodules. *IEEE Access*, 6:18592–18601, 2018.
- [55] J Royo and FJ Arcega. Machine current signature analysis as a way for fault detection in squirrel cage wind generators. In *Diagnostics for Electric Machines, Power Electronics and Drives, 2007. SDEMPED 2007. IEEE International Symposium on*, pages 383–387. IEEE, 2007.
- [56] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.
- [57] Ted Saarikko, Ulrika H. Westergren, and Tomas Blomquist. The Internet of Things: Are you ready for what’s coming? *Business Horizons*, 60(5):667–676, 2017.
- [58] Gervasio F. Santos, Eduardo A. Haddad, and Geoffrey J D Hewings. Energy policy and regional inequalities in the Brazilian economy. *Energy Economics*, 36:241–255, 2013.
- [59] Freescale Semiconductor. Three Axis Low-g Micromachined Accelerometer, 2008.
- [60] Suliman Shanbr, Faris Elasha, Mohamed Elforjani, and Joao Teixeira. Detection of natural crack in wind turbine gearbox. *Renewable energy*, 118:172–179, 2018.
- [61] Benjamin K. Sovacool. The intermittency of wind, solar, and renewable electricity generators: Technical barrier or rhetorical excuse? *Utilities Policy*, 17(3-4):288–296, 2009.
- [62] Marcelo Martins Stopa, Braz J Cardoso Filho, and Carlos B Martinez. Incipient detection of cavitation phenomenon in centrifugal pumps. *IEEE Transactions on Industry Applications*, 50(1):120–126, 2014.
- [63] William T Thomson, Ronald J Gilmore, et al. Motor current signature analysis to detect faults in induction motor drives-fundamentals, data interpretation, and industrial case histories. In *Proceedings of the 32nd Turbomachinery Symposium*. Texas A&M University. Turbomachinery Laboratories, 2003.
- [64] Ignas Valodka and Gitana Valodkienė. The impact of renewable energy on the economy of lithuania. *Procedia-Social and Behavioral Sciences*, 213:123–128, 2015.
- [65] Vladimir Naumovich Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5):988–999, 1999.
- [66] Bruce D. Weinberg, George R. Milne, Yana G. Andonova, and Fatima M. Hajjat. Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons*, 58(6): 615–624, 2015.
- [67] Li Da Xu, Wu He, and Shancang Li. Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4):2233–2243, 2014.
- [68] Venkata Yaramasu, Bin Wu, Paresh C. Sen, Samir Kouro, and Mehdi Narimani. High-Power Wind Energy Conversion Systems: State-of-the-Art and Emerging Technologies. *Proceedings of the IEEE*, 103(5), 2015.
- [69] Shen Yuan and Shaobing Peng. Trends in the economic return on energy use and energy use efficiency in China’s crop production. *Renewable and Sustainable Energy Reviews*, 70(May):836–844, 2017.

- [70] Dahai Zhang, Liyang Qian, Baijin Mao, Can Huang, Bin Huang, and Yulin Si. A data-driven design for fault detection of wind turbines using random forests and xgboost. *IEEE Access*, 6:21020–21031, 2018.
- [71] Z. Zhang, A. Chen, A. Matveev, R. Nilssen, and A. Nysveen. High-power generators for offshore wind turbines. *Energy Procedia*, 35(1876):52–61, 2013.