XLIII CILAMCE
21-25, November, 2022
FOZ DO IGUAÇU - PR - BRAZIL

CHARACTERIZATION OF DISTURBANCE IN ELECTRIC POWER SIGNALS: A MACHINE LEARNING APPROACH

Felipe Lima de Abreu[1], Marcus Varanis [2], Pedro Augusto Beck [1] , Clivaldo de Oliveira [1], José Manoel Balthazar [3]

[1]*Faculdade de Engenharia, Universidade Federal da Grande Dourados*
felipeabreu507@gmail.com
[2]*Instituto de Física, Universidade Federal de Mato Grosso do Sul*
Marcus.varanis@ufms.br
[3]*Universidade Tecnológica Federal do Paraná*

Abstract.

Voltage variation in electrical networks is one of the problems that arise when it comes to electronic equipment that is sensitive to voltage variations. As a way of classifying voltage variation phenomena, such as Voltage Sag or Swell, this paper aims to use artificially created voltage signals in the time domain, which represent each electrical fault, which will be analyzed in the frequency domain with techniques of time frequency analysis (TFA) and entropy analysis, among other Features. Subsequently, with classic methods from the Machine Learning literature, classify the general electrical signals and identify the respective faults. As a working tool, the Python language is used, as it is easy to implement and learn, in addition to being widely documented. Additionally, the scikit-learn librarie is used, which are widely tested and documented in the literature.

Keywords: Power Quality Signals, Wavelet Packet Transform, Machine Learning.

1 Introduction

Electrical machines are widely used nowadays, whether within industry, on an assembly line, or within homes, in computers or household appliances. Because they are so important nowadays, more and more the need for a good supply and supplement of energy is necessary, as well as the stability of energy for the good operation and longevity of electrical equipment, since much of this electrical equipment is sensitive to variations in the electrical line.

Within this context, one of the relevant problems involving the power supply are electrical faults, commonly seen in households. This problem must be very well identified and avoided, because besides causing damage to the electrical equipment, damaging an entire assembly line or computers, it can also cause accidents, such as fires or explosions.

This paper aims to present the main electrical faults that can be found, characterized in the frequency domain, to classify them through 4 classifiers. The signals used in this paper are artificial signals added with noise. The tools used to obtain features are the Wavelet Packet Transform, kurtosis, and entropies.

2 Methodology

This paper aims to classify different types of electrical faults common in the literature [1]. In section (3), 9 different classes of signals are presented, which will be simulated in the time domain in a range from 0 to 0.5 [s],creating a set of 2250 different artificially generated signals, where will have 250 belonging to each class.

To extract features from these signals, it was chosen to use the Wavelet Packet Transform (WPT) with 8 levels of decomposition [2, 3], and from this decomposition, calculate the relative energy of each packet and use as a feature. The relative entropy of all packets and the signal in time is also used as a feature, and the kurtosis of the signal in the time domain as well. Through an initial analysis, it is identified that the initial packets are those that have more energy, then use only the energy of the first 16 packets, or 16 features. As a feature is also used the entropy of the packages of the signal as a whole (256 Packets), the entropy of the signal in the time domain and the kurtosis of the signal in the time domain, which totals 19 features.

Finally, for the classification, the machine learning library "scikit-learn" will be used as FrameWork, and all the classification will be done in the Python programming language. Initially, these 19 features will be normalized, so that there are no errors due to data scale, and then passed to a dimensionality reduction algorithm, the Principal Component Analyses (PCA), using the first three principal components. From this the data will be separated for testing and training, where 70 % of each class will be passed to training and 30 % to testing. Different classifiers will be evaluated, such as K-nearest neighbors (KNN), Support Vector Machine (SVM) using the one vs rest method, since it is a binary classifier, Random Forest (RF) and Multi-layer Perceptron (MLP) [4–6], and finally, the analysis and discussion of the results obtained will be done and define which is the best classifier for the present failures.

3 Data Set

The electrical faults discussed in this article, which will later be analyzed in the frequency domain, will be the following: Normal, Voltage sag, Voltage swell, Harmonics, Flicker, Oscillatory transient, Sag with harmonics, Swell with harmonics and Interruption. They are characterized as follows:

Voltage sag: is characterized with a rapid reduction in voltage in the power distribution. It can be caused by high voltage demand or a short circuit. This voltage reduction is characterized in the range from 10 to 90 percent of the nominal voltage.

Voltage swell: Unlike Voltage Sag, it is the significant increase of the nominal voltage in the electrical power distribution, mainly found in system failures. The increase in nominal voltage is characterized in the range of 110 to 180 percent for a short period of time.

Harmonics: are mainly characterized by overlapping base frequencies in the supply of power distribution networks. A good example is the sum of several sines and cosines with multiple, integer signal frequencies. Such behavior mischaracterizes the signal.

Flicker: electrical phenomenon characterized as fluctuations in the nominal voltages, an effect commonly explained as a light bulb that 'flickers' over time. This electrical fault is inconvenient and can cause damage to equipment and to humans, causing attention disturbances.

Oscillatory transient: effect visualized by the presence of brief transient periods in the nominal voltage, behavior that causes a total change in the system response.

Sag with harmonics: characterized by the same behavior as Voltage sag, but with the presence of harmonics in its nominal voltage decay.

Swell with harmonics: is characterized by following the behavior of the Voltage swell, but with the presence of harmonics in its nominal voltage increase.

Interruption: effect classified as being the momentary interruption of power supply to the system, which can be from 0 to 10 percent of the total value of the nominal voltage.

Next, in equation system 1, the equations that represent the Normal, Voltage sag, Voltage swell, Harmonics, Flicker, Oscillatory transient, Sag with harmonics, Swell with harmonics and Interruption signals, respectively, are presented in Fig.1. For the simulation a base frequency $\omega = 60$ $[Hz]$, was chosen, with a fixed time for all equations of 0.5 $[s]$. With this, a DataSet was created with these 9 classes having 250 signals from each class, thus creating a set of 2250 signals. These signals went through a feature extraction using the WPT transform with 8 levels of decomposition resulting in 256 packages, which will be calculated the energy contained in each, the entropy of all packages, and the calculation of entropy and kurtosis of the signal in the time domain, thus presenting a total of 259 possible features, of which will be used only the energy of the 16 initial packages, the entropies and kurtosis, totaling 19 features.

$$
\begin{cases}
x^{(1)}(t) = A * sin(\omega t * 2\pi) + noise(t) \\
x^{(2)}(t) = A * (1 - a((t > t_1) - (t > t_2)))sin(\omega t * 2\pi) + noise(t) \\
x^{(3)}(t) = A * (1 + a((t > t_1) - (t > t_2))) * sin(\omega t * 2pi) + noise(t) \\
x^{(4)}(t) = A * (sin(\omega t * 2\pi) + a_1 sin(2\omega t * 2\pi) + a_2 sin(3\omega t * 2\pi) + a_5 sin(5\omega t * 2\pi) + a_7 sin(7\omega t * 2\pi) + noise(t) \\
x^{(5)}(t) = A * (1 + af * sin(\omega t * 2\pi) * sin(50 * \omega t * 2/pi) + noise(t) \\
x^{(6)}(t) = A * (sin(\omega_1 t * 2\pi) + b - (t > t_l) * sin(\omega t * 2\pi * (t > t_l)) + noise(t) \\
x^{(7)}(t) = A * (1 - a * ((t > t_1) - (t < t_2))) * sin(\omega t * 2\pi) * (a_1 * sin(\omega t * 2\pi) + a_2 * sin(2\omega t * 2\pi) \\
+a_3 * sin(3\omega t * 2\pi) + a_5 * sin(5\omega t * 2\pi) + a_7 * sin(7\omega t * 2\pi) + noise(t) \\
x^{(8)}(t) = A * (1 + a * ((t > t_1) - (t > t_2))) * sin(\omega t * 2\pi) * (a_1 * sin(\omega t * 2\pi) + a_2 * sin(2\omega t * 2\pi) \\
+a_3 * sin(3\omega t * 2\pi) + a_5 * sin(5\omega t * 2\pi) + a_7 * sin(7\omega t * 2\pi) + noise(t) \\
x^{(9)}(t) = A * (1 - a((t > t_1) - (t > t_2))) * sin(\omega t * 2\pi) + noise(t)
\end{cases}
\tag{1}
$$

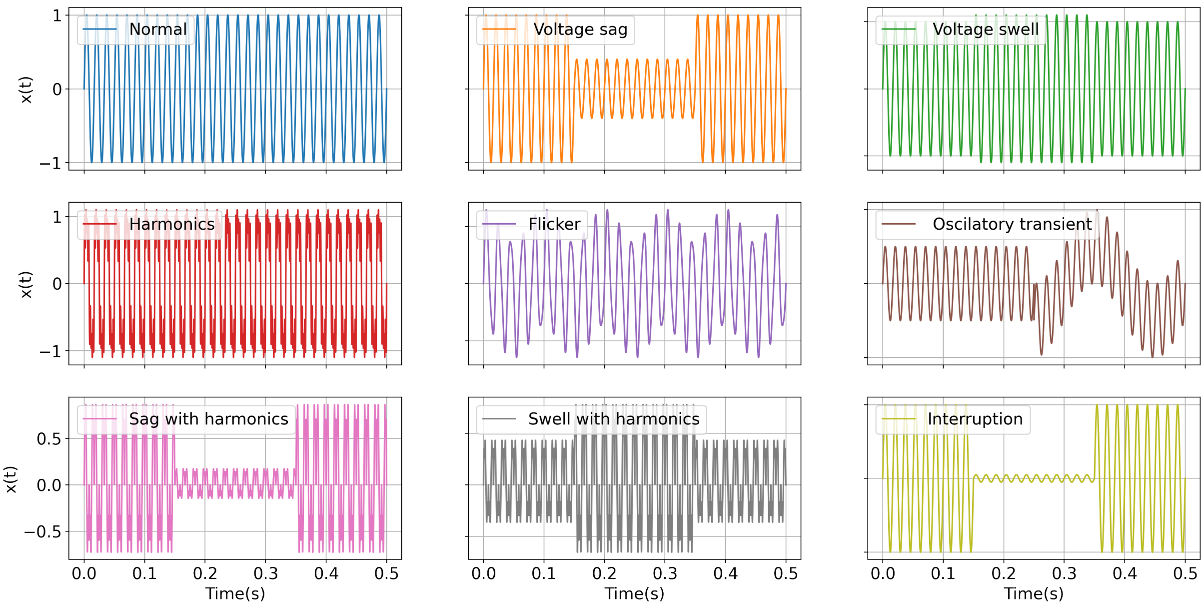

Figure 1. Artificial signals with their respective faults

4 Classification

For the classification, as already mentioned in the methodology, the PCA dimensionality reduction algorithm was used. Below, in Fig.2, have how much of the data variation was captured by the first ten components, which in this case, the first three components, seen in Fig.3, are shown to have captured 87.28 % of the data variation..

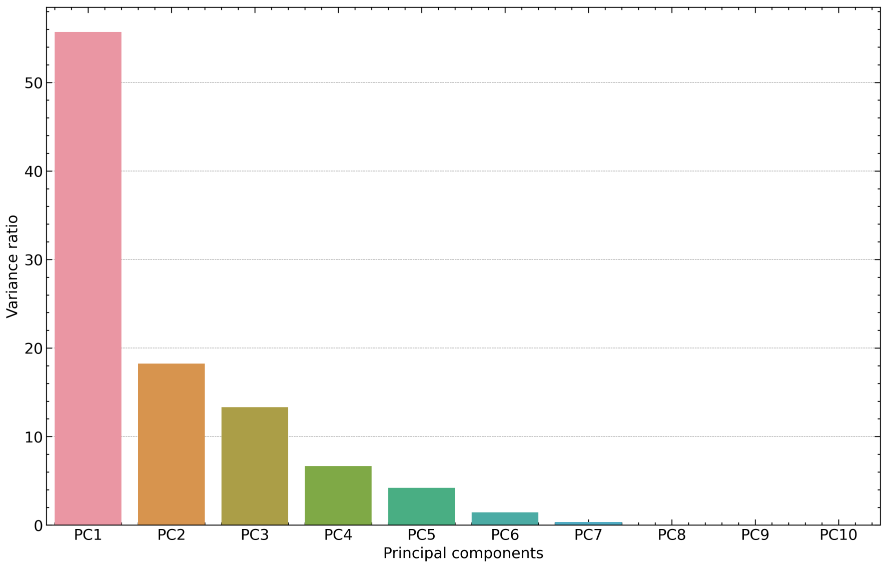

Figure 2. First 10 principal components

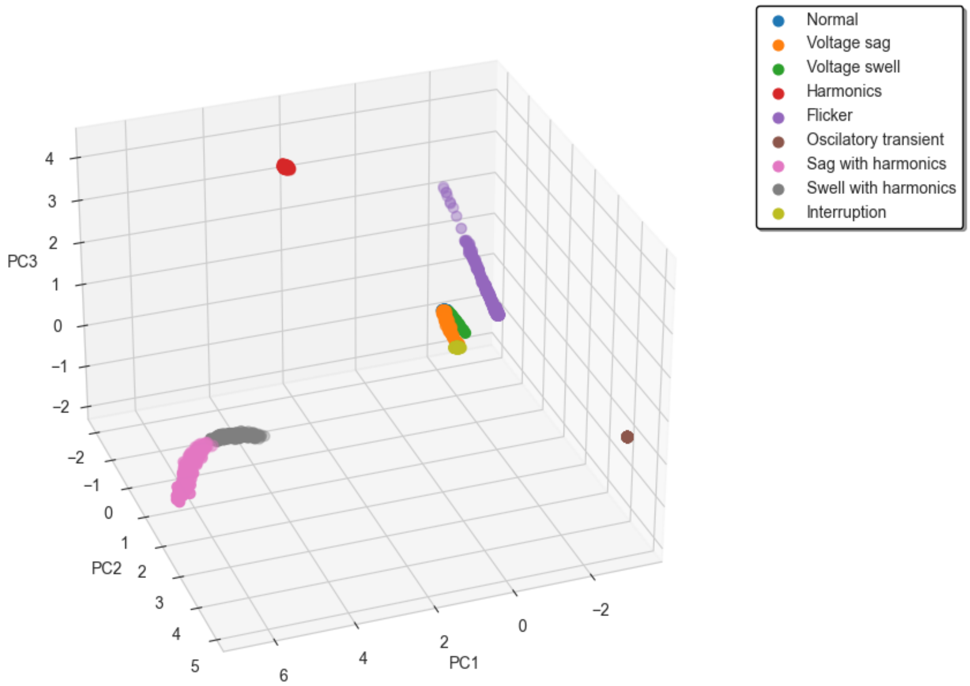

Figure 3. Visualization of the first 3 principal components

From the data visualization these features were able to capture well the representation of each class. These three main components will be used to train different algorithms, all already implemented in the scikit-learn library, which were iterated several parameters seeking to improve the accuracy of the classifier without experiencing overfitting problems. The amount of signals used for training and testing was given by a 70% and 30% ratio respectively, for each of the classes.

Below in Fig.4, have the confusion matrix of some trained classifiers, these are the KNN algorithm, SVM with RBF kernel, RF and finally a MLP with two layers with ten neurons each. The confusion matrices present on the x-axis the class predicted by the algorithm and on the y-axis the real class, where the colors present the hit magnitude, in this case they are all the same color due to the high hit magnitude of all classifiers. The classes 1 to 9 represent respectively the Normal, Voltage sag, Voltage swell, Harmonics, Flicker, Oscillatory transient, Sag with harmonics, Swell with harmonics and Interruption signals.

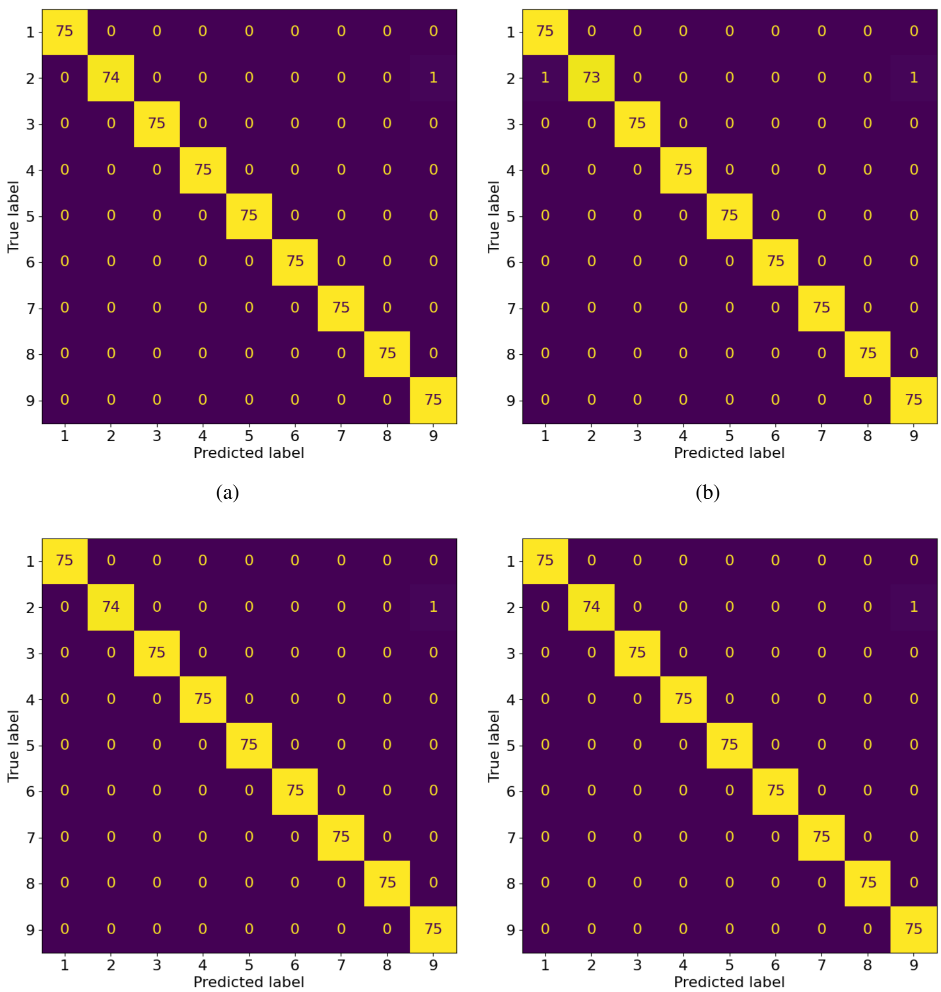

Figure 4. Confusion matrix of the classifiers: a) KNN, b) SVM, c)RF e d) MLP.

5 Discussion

In this section will present a comparison between the analysis of the classifiers of the first 16 packages plus entropies and kurtosis, and using all 256 packages plus entropies and kurtosis, both cases using the three principal components of the PCA.

The accuracy roughly is the amount of hits within the result returned by the classifier, for example, in an image identifier where an image has 11 circles and 5 squares, the classifier returns 1 result with 9 answers, where 6 are circles and 3 are squares, therefore the accuracy is 6/9 and the revocation is the amount of hits by the maximum number of hits, i.e. the maximum number of hits it can have is 11 circles, therefore the revocation would be 6/11. Following the previous example to explain 3 terms: the first being True positive (TP), where it refers precisely to the answers returned that are really true, in this case would be the 6 circles of the 9 answers returned, the False Positive (FP) that unlike TP refers to the answers that do not represent what was asked, In this case the 3 squares that were returned within the circles classification, and finally the False negative (FN) that are the answers that should be within the classification but are not, for example the 5 circles that were not included in the classification even though they were circles, in other words, were classified as not circles. The F1 score, shown by eq. (2), presents the harmonic mean of precision and revocation.

$$F_1 = \frac{2}{\frac{1}{precision} + \frac{1}{revocation}} = \frac{TP}{TP + \frac{FN+FP}{2}} \tag{2}$$

The analysis using only the first packages presents an F1 score higher than 99% for all classifiers, this may be due to a higher importance involving the initial packages since they present a higher energy. Applying the PCA in the first case, with the 19 features, it is obtained that 87.28% of the data variation in the first three components, which when training the algorithms with these components was enough to obtain good classifications, as presented in Tab.1.

Classifiers	Average F1 Score
KNN	**99.78%**
SVM	**99.67%**
RF	**99.78%**
MLP	**99.78%**

Table 1. 19 Features

Now in Tab.2, when using all the provided packages, the F1 score of all classifiers dropped considerably, which may be due to the rest of the packages left out in the first analysis having lower energy. The principal components analysis failed to keep most of the variation of all the data in the first three principal components, as it contains only 62.88 % of the variation of all the data, which did not prove sufficient for correct classification.

Classifiers	Average F1 Score
KNN	**95.55%**
SVM	**96.67%**
RF	**96.00%**
MLP	**96.89%**

Table 2. 259 Features

From an analysis of the confusion matrix of both cases, realized that the main failures in the classification are present in the failures Voltage sag and Interruption, which have similar responses, with the only difference being the magnitude of the failure, in addition another misclassification that occurs when using all the packages is the confusion of classifying signals that have Voltage sag and Voltage swell with low amplitudes of failure as normal signals, but still presenting a relative quality in the hits as can be seen in the table above, the similarity in the signals can be seen in Fig.1.

With this from the features extracted with the WPT, entropy, and kurtosis, plus a little data engineering, it was possible to use classical methods of machine learning to obtain a classification with good accuracy in classical data sets, where methods were used with low computational cost compared to other methods of frequency time using spectrograms and image readings with neural networks, which require much more computational power.

6 Conclusion

The analyses involving the classifiers present in this paper were as a whole efficient to characterize the electrical faults, even though the signals are similar and present similar behaviors, a good hit rate was found. The dataset used for training and testing is a first incursion, since it is data that has been extensively studied in the field of power quality signals, and in this dataset a small amount of signals (2250) was used compared to the amount needed for training other types of machine learning algorithms.

The methods presented in this paper for the extraction of features in the frequency domain, such as the energy of the WPT packages, the entropies and the kurtosis, were sufficient for a good data classification. However, for the classification to have a higher accuracy, it was necessary to find among the possible features those that presented a better representation of the data, so that we could perform the dimensionality reduction and a good overall classification of all algorithms. The extraction of information in the frequency domain from the WPT packages was confirmed as a good feature for classification, along with the remaining features.

In this paper, the K-nearest neighbors, Support Vector Machine using the one vs rest method, Random Florest and Multi-layer Perceptron algorithms were used, all already implemented in the ”scikit-learn” library, with all algorithms in the first case (19 features) and second case (259 features) showing a similar response. The hit rate for the first case in all algorithms is higher than 99%, and for the second case around 96%, as seen in section 5. The results also show that one of the reasons for the hit rate not being 100% is the similarity between the signals, seen in more depth in section 5.

Acknowledgements. The authors acknowledge support by CNPq (Brazilian research funding agency).

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] Minh Khoa, Ngo, and Le Van Dai. "Detection and classification of power quality disturbances in power system using modified-combination between the Stockwell transform and decision tree methods." Energies 13.14 (2020): 3623.

[2] Choe, Sangho, and Jeonghwa Yoo. "Wavelet packet transform modulus-based feature detection of stochastic power quality disturbance signals." Applied Sciences 11.6 (2021): 2825.

[3] Zhang, Yong, et al. "Comparison of classification methods on EEG signals based on wavelet packet decomposition." Neural Computing and Applications 26.5 (2015): 1217-1225.

[4] Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2019.

[5] XIAO, Fei et al. Maximal overlap discrete wavelet transform and deep learning for robust denoising and detection of power quality disturbance. IET Generation, Transmission & Distribution, v. 14, n. 1, p. 140-147, 2020.

[6] Khetarpal, Poras, and Madan Mohan Tripathi. "A critical and comprehensive review on power quality disturbance detection and classification." Sustainable Computing: Informatics and Systems 28 (2020): 100417.