A Machine Learning-Based Internal Fault Identification in Power Transformers

Lucas D. Simões, Adinara L. R. de Oliveira, Flavio Bezerra Costa
School of Science and Technology
Federal University of Rio Grande do Norte
Natal, RN, Brazil
Emails: lucassimoesee@gmail.com, adinara_larissa@hotmail.com, flaviocosta@ect.ufrn.br

Rodrigo Prado de Medeiros
Multidisciplinary Center of Caraúbas
Federal Rural University of the Semi-Arid
Caraúbas, RN, Brazil
Email: rodrigo.prado@ufersa.edu.br

***Abstract*—Currently, several techniques based in digital signal processing and artificial intelligence tools have been applied in power transformer protection in order to detect and discriminate internal faults from external faults. This paper presents the application of two machine learning algorithms, the support-vector machine and random forest, in order to distinguish between internal and external faults. A performance comparison between the both techniques regarding the success rate in discriminating the events is accomplished. The results reveal the feasibility regarding the application of these techniques in association to the power transformer protection schemes.**

***Keywords* – machine learning, power transformers, random forest, support-vector machines, wavelet transform**

I. Introduction

Power transformers are expensive components in electrical power systems and of high strategic importance [1], interconnecting the generation system to the transmission and distribution systems and ensuring the system operation in appropriated voltage levels. These components are subject to failures, of which 70-80% of faults result from turn-to-turn or internal faults, and 10% of faults in power systems take place on power transformers. [2] In addition, the maintaining of a faulted transformer demands an expensive financial cost due to its high commercial cost and the imposition of fines due to power outages [3]. Therefore, it is of utmost importance the protection and monitoring of these equipments, providing a fast and accurate disturbance detection and maintaining the security and reliability of the system.

The percentage differential protection has been the main protection for transformers with rated power above 10 MVA [4], providing a reliable discrimination between internal and external faults. However, this protection may fail in the presence of CT saturation and inrush currents during transformer energization [5], causing an incorrect relay operation.

In order to overcome this limitation, many modern percentage transformer differential relays have incorporated additional harmonic restraint and harmonic blocking methods [6]. However, these methods can fail for some inrush situations in which the second harmonic content of the differential currents may fall below 15% [7]. In addition, the conventional harmonic restraint and blocking methods present an inherently delay due to convergence of phasor estimation in pre- and post-fault regimes [8].

Recently, several techniques based on artificial intelligence and some digital signal processing tools have been developed to discriminate efficiently internal faults from other events in power transformer operation [8]–[17]. Among these techniques, some of them have presented promising results in power transformer fault detection and discrimination, especially those implementing wavelet transform and machine learning algorithms [8], [13]–[17]. For instance, a power transformer differential protection algorithm based on the boundary wavelet coefficient energy of the differential currents was proposed in [8]. In [13] a multilayer perceptron neural network was implemented in order to discriminate between internal faults, external faults and transformer energizations. In [14], a support-vector machine-based method for discriminating between internal faults and other disturbances was proposed with good results obtained even during CT saturation. Further, a random forest algorithm with the purpose of discriminating internal faults from other disturbances was implemented, and good results regarding the events classification were achieved [15]. All these techniques are able to provide a possible way for making power transformer protection schemes more reliable and faster.

This paper proposes the application of two machine learning (ML) algorithms, the support-vector machine (SVM) and random forest (RF), which were implemented to appropriately discriminate internal faults from external faults to the transformer protection zone. Further on, a comparison between the two techniques is carried out in order to evaluate which one performs better regarding the discrimination of these events. The input signals for the ML algorithms are the differential wavelet coefficient energies, which were computed by means of the real-time boundary stationary wavelet transform (RT-BSWT), according to [17]. For training and validating the models, several internal and external faults were simulated with variations of fault inception angle, fault resistance, and fault type.

Both ML-algorithms presented very significant success rates regarding the discrimination between internal and external faults. The presented technique can be implemented for real-time applications and is capable of operating in conjunction with traditional power transformer differential protection algorithms.

II. Machine Learning Algorithms

Machine learning is an important concept of the artificial intelligence area. According to [18], this concept is based on the idea that systems can learn from data, recognize patterns and make decisions with minimal human intervention. Tipically, the machine learning philosophy can be divided into two important categories: the supervised learning and unsupervised learning. With regard to a supervised learning model, the purpose is to classify unknown data from a database subdivided into groups according to their similarities. From the characteristics of each group, it is possible to infer a value judgment to the unknown data by comparing it with the existing groups and classifying it in one of these groups [18]. This paper used two supervised ML algorithms: the support-vector machine (SVM) and the random-forest (RF).

A. Support Vector Machine

The SVM is a supervised learning tool applied in both classification and data regression. The purpose of the classification is to partition two classes through an optimal separation hyperplane, making it possible to obtain the support vectors, which will produce a maximum margin capable of separating these classes. The support vectors delimit the positive and negative hyperplane sides, allowing the classification of data as belonging to one of these classes [19].

The classification proceeds from a set of training samples $\{(x_i,d_i)\}$, where x_i is the input pattern for the *i*-th term. This data set is linearly separable if it is possible to separate the samples into two classes, delimited by the positive and negative hyperplanes. The support vectors are the points closest to the optimal hyperplane and, through them, it is possible to generate a maximized separation margin, represented by ρ. Fig 1 depicts an optimal hyperplane.

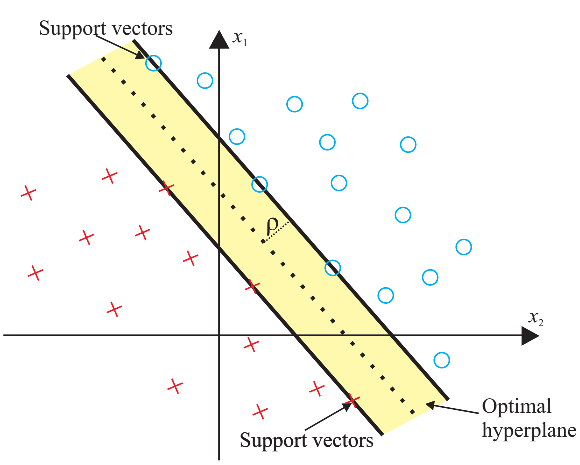

Figure 1. Illustration of an optimal hyperplane

The representation of the optimal hyperplane is defined as follows:

$$\mathbf{w}^T\mathbf{x} + b = 0, \tag{1}$$

where **w** is an adjustable weight vector; **x** is a vector containing the input data; and b is a bias. The solution of (1) is greater than or equal to zero for d_i = 1 and less than or equal to zero for d_i = -1.

The desired vector length is computed from the optimal values of the weight vector w_0 and bias b_0, and taking into account the existing multiplanes. Applying this analysis in (1) and computing the distance between the planes formed by the support vectors, the desired algebraic distance (r) is defined as follows:

$$r = \frac{2}{\|\mathbf{w}\|}. \tag{2}$$

B. Random Forest

Random Forest is an ensemble ML method which uses several typically weak predictive models, such as decision trees, in order to build more robust and stronger models. This is usually achieved by injecting randomness and diversity into the learning algorithm. Therefore, RF is a technique based on many random-generated decision trees.

In this method, randomness can be injected by means of one random vector θ_k, which is generated for the kth tree, independently from the previous θ_1,...,θ_{k-1} random vectors. According to [20], the parameters of the random vector define the behaviour of the decision trees, such as split nodes (j) thresholds and random feature selection from the input training data vector **x**. Furthermore, randomness is also carried into the algorithm by means of randomized subset sampling from the input **x**, thus growing each tree with a randomly different training subset. This strategy is commonly known as bootstrap-aggregating (bagging). These trees grown with different subsets of randomly selected samples and features differ from each other, allowing an improved generalization capability.

After growing, each tree is a classifier represented by $h(\mathbf{v}, \theta_k)$, which votes in a determined class for a given input vector **v**. The class with majority of votes is the winner, becoming the model output [21]. An example of one of these trees is depicted in Fig. 2, as well as the split nodes j, a given testing or validating input vector **v** and the selection for a class in each of the tree's nodes, represented by the coloured squares.

III. Proposed Method

Fig. 3 depicts the block diagram of the proposed ML-based transformer fault identifier, which is executed every sampling time k. The ML-based scheme uses the operating and restraining wavelet coefficient energy signals, as described in [17], as inputs for properly discrimination between internal and external faults. The functions of each block in this diagram are addressed in the remainder of this section.

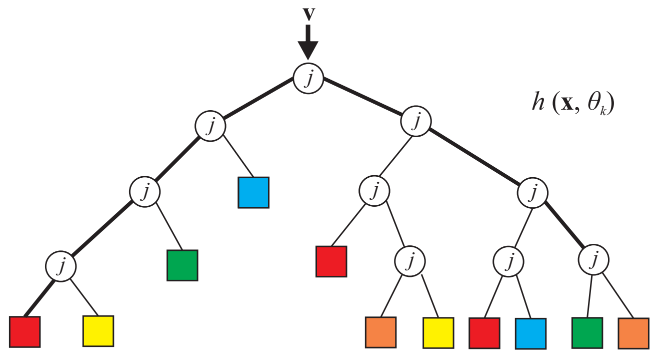

Figure 2. Example of a kth decision tree among the k trees in the random forest

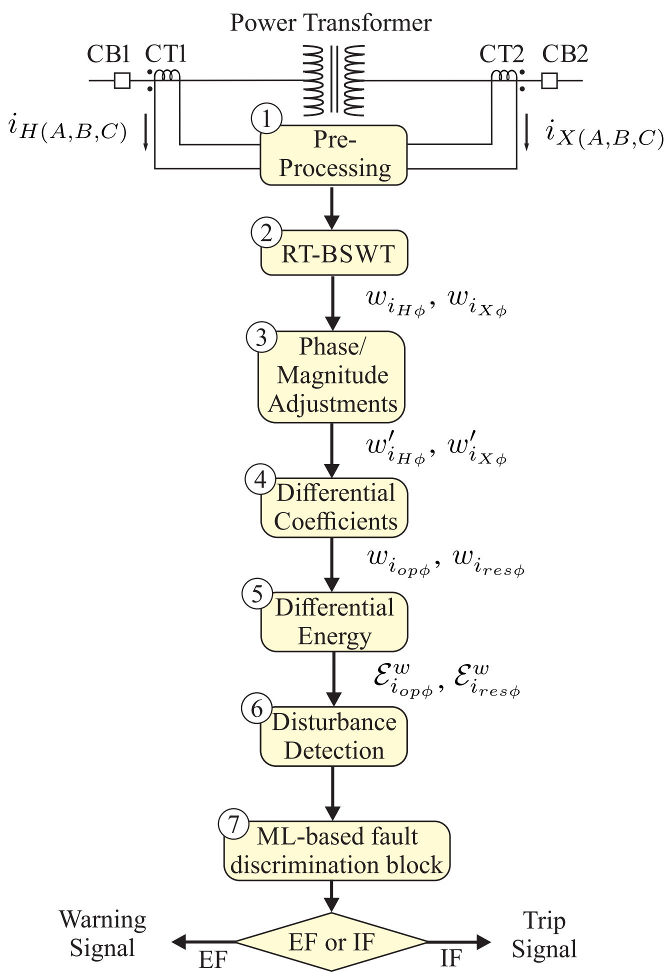

Figure 3. Proposed fault discrimination scheme flowchart

A. Pre-Processing (Blocks 1 - 6)

In this paper, the blocks 1-6 execute all required pre-processing computations, as follows:

- Analog Filtering (block 1): The relay performs the digital acquisition of the three-phase CT secondary currents, by means of an anti-aliasing filter and an A-D converter, in order to get the time-discrete secondary currents ($i_{H\phi}$ and $i_{X\phi}$). The variable ϕ corresponds to phases A, B, and C.
- RT-BWST (block 2): After the analog filtering of the CT currents, the real-time boundary wavelet coefficients w = $\{w_{i_{H\phi}}$ and $w_{i_{X\phi}}\}$ of the currents i = $\{i_{H\phi}$ and $i_{X\phi}\}$, respectively, are computed, as described in [22].
- Phase/magnitude adjustments (block 3): The amplitude, phase shift and zero-sequence correction is performed on the wavelet coefficients of the currents [17].
- Differential coefficients (block 4): The differential wavelet coefficients are given by [16]:

$$w_{i_{op\phi}}(0,k) = \frac{1}{2}(w'_{i_{H\phi}}(0,k) + w'_{i_{X\phi}}(0,k)), \quad (3)$$

$$w_{i_{op\phi}}(l \neq 0,k) = w'_{i_{H\phi}}(l,k) + w'_{i_{X\phi}}(l,k), \quad (4)$$

$$w_{i_{res\phi}}(l,k) = w'_{i_{H\phi}}(l,k) - w'_{i_{X\phi}}(l,k), \quad (5)$$

where $0 \leq l < L$ and w_{diff} = $\{w_{i_{op\phi}}, w_{i_{res\phi}}\}$.
- Differential energy (block 5): The boundary wavelet coefficient energies $\mathcal{E}^w_{diff}$ = $\{\mathcal{E}^w_{i_{op\phi}}, \mathcal{E}^w_{i_{res\phi}}\}$ are computed from the respective differential wavelet coefficients w_{diff} = $\{w_{i_{op\phi}}, w_{i_{res\phi}}\}$, as follows [16]:

$$\mathcal{E}^w_{diff}(k) = \mathcal{E}^{wa}_{diff}(k) + \mathcal{E}^{wb}_{diff}(k), \quad (6)$$

in which the terms $\mathcal{E}^{wa}_{diff}$ and $\mathcal{E}^{wb}_{diff}$ are computed as [22]:

$$\mathcal{E}^{wa}_{diff}(k) = \sum_{l=1}^{L-1} w^2_{diff}(l,k), \quad (7)$$

$$\mathcal{E}^{wb}_{diff}(k) = \sum_{n=k-\Delta k+L}^{k} w^2_{diff}(0,n). \quad (8)$$

- Event detector (block 6): Based on [8], any transient event, such as internal and external faults, can be detected when:

$$\begin{cases} \mathcal{E}^w_{diff}(k-1) \leq E_{diff}, \\ \mathcal{E}^w_{diff}(k) > E_{diff}, \end{cases} \quad (9)$$

where k_d is the first sample in which both inequalities in (9) are valid; E_{diff} = $\{E_{op\phi}, E_{res\phi}\}$ are the steady-state energy thresholds determined by [8]:

$$E_{diff} = \frac{3}{k_2 - k_1 + 1} \sum_{n=k_1}^{k_2} \mathcal{E}^w_{diff}(n), \quad (10)$$

where $[k_1/f_s \; k_2/f_s]$ is an arbitrary prior steady-sate time range.

B. ML-based fault discrimination block (Block 7)

As soon as an event is detected, the ML-based fault discriminator is enabled in order to distinguish between internal and external faults to the transformer protection zone.

In this paper, two different ML-based fault classification techniques were used: the RF and SVM. Both algorithms receive as inputs the boundary differential energies $\mathcal{E}^w_{diff}$, which are stored in a sliding window with the last four samples $(k-3, k-2, k-1, k)$, and $k > k_d + 2$. The sliding window with length of four samples was strategically chosen in order to limit the number of inputs of the ML-based algorithms, providing a better convergence and a faster training of the models [11]. Therefore, in the postfault regime, the windowing is performed sample by sample, discarding the first sample from the window and adding the next sample. Fig. 4 depicts the operating and restraining wavelet coefficient energies of a faulted signal, with their respective sliding windows as inputs for the ML-based fault discrimination block.

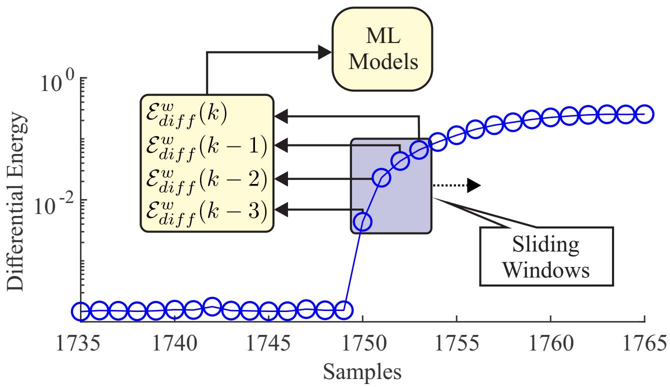

Figure 4. Differential energy sliding windows as inputs for ML models

According to Fig. 4, the ML models will have as inputs sliding windows from the differential energy signals with 4 samples each, for both restraining and operating energies, and for phases A, B and C. Thus, this results in 24 features as inputs for the ML models for each training patterns.

The ML-based fault discrimination logic is essential in order to define the type of event, whether it is a fault outside or within the transformer protection zone. Therefore, if an external fault is detected, the algorithm is able to send a warning signal. Otherwise, if an internal fault is detected, a trip signal is addressed.

C. Training and validation stages for ML-based algorithms

For training and validating the models for both ML blocks, the dataset was divided as depicted in Fig. 5.

According to Fig. 5, the total simulated records was divided randomly into training and validation databases. Firstly, for each record, the 64 first post-fault sliding windows were gathered as training and validating patterns. Afterward, the training patterns were shuffled and partitioned into training and test sets. These sets were used in order to train and adjust the ML models.

After proper parameter adjustment, 10-fold cross-validation was performed on the models in order to achieve the better generalization capability and avoid overfitting.

IV. Performance Assessment

Fig. 6 depicts a single-line diagram of the electrical power system used for performance assessment of the ML models regarding the fault discrimination and classification. The system was simulated by using the Alternative Transients Program (ATP), consisting in a power transformer with their primary and secondary windings connected to the 230 kV and 69 kV

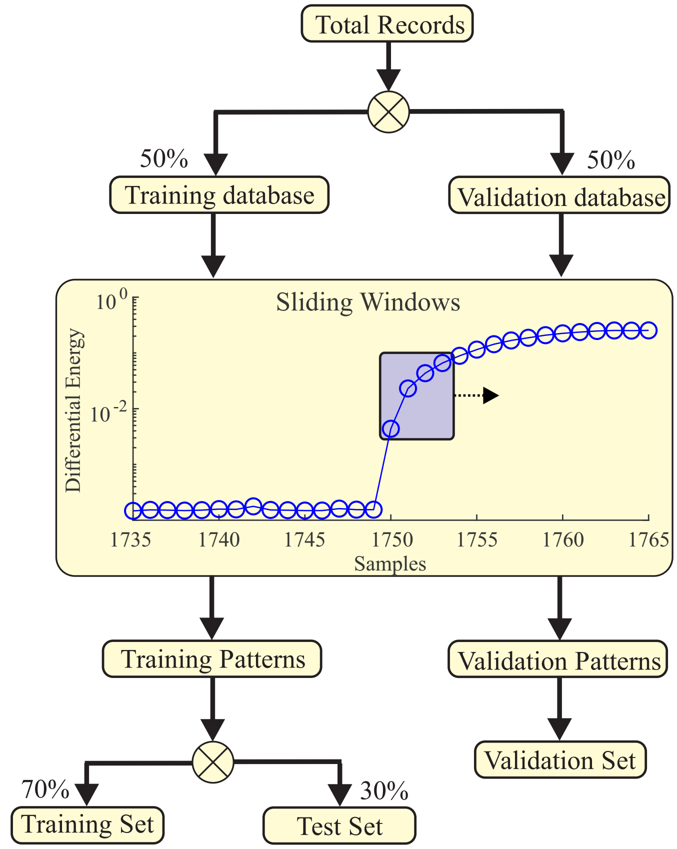

Figure 5. Training and Validation Flowchart

Thevenin equivalent systems, respectively. Details about the parameterization of the system components are described in [16].

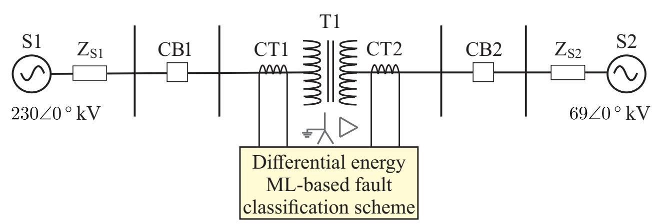

Figure 6. Electrical system single phase diagram

The ML models performance was assessed from the following databases of internal and external faults, generated in ATP simulations:

- Database 1 (external faults): AG, AB, ABG, AC, ACG and ABC faults, on both high and low voltage sides of the power transformer, with varying fault inception angle $\theta_f = \{0, 30, 60, 90, 120, 150, 180\}$ electrical degrees and fault resistance $R_f = \{1, 2, 3, ..., 9, 10\}$ Ω (840 records).
- Database 2 (internal faults): AG, BG, CG, AB, ABG, AC, ACG, BC, BCG, ABC faults, on both sides of high and low voltage windings of the power transformer, while varying fault inception angle θ_f

= $\{0, 30, 60, 90, 120, 150, 180\}$ electrical degrees and fault resistance $R_f = \{10, 20, 30, ..., 90, 100\}$ Ω (1400 records).

A. *Fault discriminator performance assessment*

Fig. 7 depicts the confusion matrix obtained for the performance assessment of the SVM regarding the correct recognition of external and internal faults. According to Fig. 7, the main diagonal of the matrix corresponds to the total of cases that were correctly classified, for each class. On the other hand, the elements outside of the main diagonal represent the misclassified cases. For instance, a total of four internal faults were classified as external faults. Therefore, the SVM presented a success rate of 99.64 % for discriminating between both events.

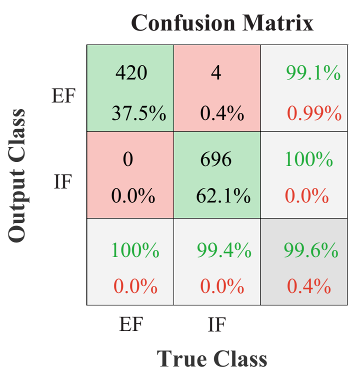

Figure 7. Confusion matrix obtained from the SVM algorithm for discriminating between internal and external faults

With regard to the RF algorithm, Fig. 8 illustrates its performance. According to Fig. 8 , none of the events was misclassified by the RF algorithm (success rate of 100 %). Therefore, the RF model showed greater efficiency in discriminating between external and internal faults in the transformer.

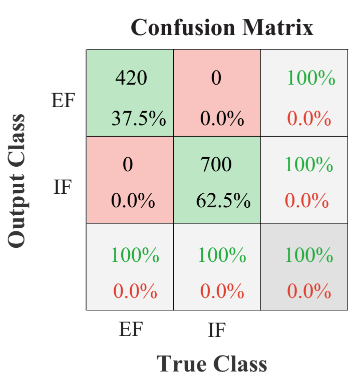

Figure 8. Confusion matrix obtained from the RF algorithm for discriminating between internal and external faults

V. Conclusion

This paper presented a power transformer fault discrimination scheme based on SVM and RF algorithms. The two techniques performed well, with both of them presenting good results in fault discrimination. However, the RF technique has shown a superiority in discriminating between internal and external faults, presenting a success rate of 100 % against 99.6 %. Therefore, the RF algorithm has proven itself a more robust, accurate and with greater generalization capability model in discriminating faults in power transformers.

Regarding the results, they also show that with such high success rates, both algorithms can handle more critical fault cases classification, such as turn-to-turn faults or faults during CT saturation, hence providing a possibility for future investigations. Furthermore, the proposed scheme could yet be integrated to a protection relay for further applications to power transformers protection.

Acknowledgment

The authors would like to thank CAPES and CNPq for the financial support.

References

[1] O. Ozgonenel and S. Karagol, “Power transformer protection based on decision tree approach,” *Electric Power Applications, IET*, vol. 8, no. 7, pp. 251–256, August 2014.

[2] O. Ozgonenel, E. Kilic, D. Thomas, and A. E. Ozdemir, “Identification of transformer internal faults by using an rbf network based on dynamical principle component analysis,” in *2007 International Conference on Power Engineering, Energy and Electrical Drives*, April 2007, pp. 719–724.

[3] M. O. Oliveira, “Proteção diferencial adaptativa de transformadores de potência baseada na análise de componentes wavelets,” 2013.

[4] IEEE Std C37.91-2008, “Ieee guide for protecting power transformers,” pp. c1–139, 2008.

[5] J. L. Blackburn and T. J. Domin, *Protective Relaying: Principles and Applications, Third Edition*. Taylor & Francis, 2006.

[6] R. Hamilton, “Analysis of transformer inrush current and comparison of harmonic restraint methods in transformer protection,” *IEEE Trans. Ind. Applications*, vol. 49, no. 4, pp. 1890–1899, July 2013.

[7] S. Hodder, B. Kasztenny, N. Fischer, and Y. Xia, “Low second-harmonic content in transformer inrush currents - analysis and practical solutions for protection security,” in *67th Annual Conference for Protective Relay Engineers*, March 2014, pp. 705–722.

[8] R. P. Medeiros, F. B. Costa, and K. M. Silva, “Power transformer differential protection using the boundary discrete wavelet transform,” *IEEE Trans. Power Del.*, vol. 31, no. 5, pp. 2083–2095, Oct 2016.

[9] O. Ozgonenel and S. Karagol, “Maximum overlap discrete wavelet based transformer differential protection,” in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, May 2017, pp. 1–4.

[10] Yong Sheng and S. M. Rovnyak, “Decision trees and wavelet analysis for power transformer protection,” *IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 429–433, April 2002.

[11] P. L. Mao and R. K. Aggarwal, “A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network,” *IEEE Transactions on Power Delivery*, vol. 16, no. 4, pp. 654–660, Oct 2001.

[12] M. O. Oliveira and A. S. Bretas, “Application of discrete wavelet transform for differential protection of power transformers,” in *2009 IEEE Bucharest PowerTech*, 2009, pp. 1–8.

[13] J. F. Fernandes, F. B. Costa, and R. P. de Medeiros, “Power transformer disturbance classification based on the wavelet transform and artificial neural networks,” in *2016 International Joint Conference on Neural Networks (IJCNN)*, July 2016, pp. 640–646.

[14] A. M. Shah and B. R. Bhalja, “Discrimination between internal faults and other disturbances in transformer using the support vector machine-based protection scheme,” *IEEE Trans. Power Del.*, vol. 28, no. 3, pp. 1508–1515, 2013.

[15] A. M. Shah and B. R. Bhalja, “Fault discrimination scheme for power transformer using random forest technique,” *IET Generation, Transmission Distribution*, vol. 10, no. 6, pp. 1431–1439, 2016.

[16] R. P. Medeiros and F. B. Costa, “A wavelet-based transformer differential protection with differential current transformer saturation and cross-country fault detection,” *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 789–799, Apr 2018.

[17] R. P. Medeiros and F. B. Costa, “A wavelet-based transformer differential protection: Internal fault detection during inrush conditions,” *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 2965–2977, Dec 2018.

[18] K. Faceli, *Inteligência Artificial. Uma Abordagem de Aprendizado de Máquina*. LTC, 2011.

[19] S. S. Haykin, *Neural networks and learning machines*, 3rd ed. Upper Saddle River, NJ: Pearson Education, 2009.

[20] A. Criminisi, E. Konukoglu, and J. Shotton, “Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning,” Tech. Rep. MSR-TR-2011-114, October 2011. [Online]. Available: https://www.microsoft.com/en-us/research/publication/decision-forests-for-classification-regression-density-estimation-manifold-learning-and-semi-supervised-learning/

[21] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324

[22] F. B. Costa, “Fault-induced transient detection based on real-time analysis of the wavelet coefficient energy,” *IEEE Trans. Power Del.*, vol. 29, no. 1, pp. 140–153, Feb 2014.