

DIAGNÓSTICO DE DISTÚRBIOS DE TENSÃO EM SISTEMAS DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA EMPREGANDO TÉCNICAS DE MINERAÇÃO DE DADOS

DIAGNOSIS OF VOLTAGE DISTURBANCES IN ELECTRIC POWER DISTRIBUTION SYSTEMS USING A DATA MINING TECHNIQUES

Entrega dos originais à redação em: 26/12/2022

Editoração em: 15/01/2023

Nayara Fernanda Santos Delfino de Sousa

Bacharelado em Engenharia Elétrica

José Guilherme Magalini Santos Decanini

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Presidente Epitácio

Leonardo Ataíde Carniato

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Presidente Epitácio

Tiago Veronese Ortunho

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Presidente Epitácio

Neste artigo propõem-se uma metodologia para detectar e classificar cinco tipos de distúrbios de tensão em sistemas de distribuição de energia elétrica empregando técnicas de processamento de sinais e de mineração de dados. Esta ferramenta de apoio à tomada de decisão provê maior segurança, confiabilidade e rentabilidade às concessionárias de energia elétrica. O procedimento adotado para diagnóstico de distúrbios contempla a extração de características dos sinais de tensão por meio da transformada wavelet, análise multirresolução e do conceito de energia, as quais são analisadas utilizando técnicas de mineração de dados, mais especificamente, os algoritmos J48, IBK e a rede neural artificial Perceptron Multicamadas. Dentro desta perspectiva, empregou-se o software Matlab para composição do banco de dados baseado nas equações representativas da operação normal e dos distúrbios sag, swell, outage, harmonics e flicker, e para extração das características dos sinais. Neste sentido, o sistema de diagnóstico compreende a análise e avaliação de um conjunto de 1800 eventos relativos ao estado operativo do sistema de energia elétrica. O diagnóstico foi realizado via software WEKA, no qual implementou-se também a seleção de atributos por meio do método InfoGainAttributeEval Ranker, a fim de otimizar o processo e obter melhor performance computacional. Os módulos classificadores concebidos por meio dos algoritmos J48, IBK e a rede neural artificial Perceptron Multicamadas proveram eficiência de 99,6667%, 99,9444% e 99,8889%, respectivamente. Dentro desta perspectiva, os resultados obtidos demonstram que o método empregado é eficiente, robusto e propicia alta velocidade de processamento, o que permite, a priori, aplicação em tempo real.

Palavras-chave: Diagnóstico de distúrbios. Sistemas de distribuição de energia elétrica. Técnicas de mineração de dados.

In this paper proposes a methodology to detect and classify five types of voltage disturbances in electric power distribution systems using signal processing and data mining techniques. This support tool for decision making provides greater security, reliability and profitability to electricity utilities. The procedure adopted for the disturbances diagnosis contemplates the extraction of voltage signals characteristics through the wavelet transform, multi-resolution analysis and the energy concept, which

are analyzed using data mining technique, more specifically, the J48, IBK algorithms and the Multilayer Perceptron artificial neural network. Within this perspective, the Matlab software was used to compose the database based on the equations representing disturbances sag, swell, outage, harmonics, flicker and the normal operation, and for extracting the signals characteristics. In this sense, the diagnostic system comprises the analysis and evaluation of a set of 1800 events relating to the electric power system operating state. The diagnosis was performed via WEKA software, in which the selection of attributes was also implemented through the InfoGainAttributeEval Ranker method, in order to optimize the process and obtain better computational performance. The classifier modules designed using the J48, IBK algorithms and the Multilayer Perceptron artificial neural network provided efficiency of 99.6667%, 99.9444% and 99.8889%, respectively. The results obtained demonstrate that the method used is efficient, robust and provides high processing speed, which allows, a priori, real-time application.]

[Keywords: Disturbances diagnosis. Electric power distribution systems. Data mining techniques.]

1 INTRODUÇÃO

Nos últimos anos têm ocorrido uma evolução notória no que concerne os sistemas de distribuição de energia elétrica. A penetração considerável de geração distribuída, o progresso da tecnologia da informação, que propicia o desenvolvimento e implementação de técnicas de processamento de sinais e de armazenamento de dados, a tecnologia digital que engloba novos materiais e componentes que possibilitam alta velocidade de comunicação, e o avanço na concepção de sistemas inteligentes, estão permitindo uma ampla transformação das redes de energia elétrica. Dentro desta perspectiva, evidencia-se o conceito de *smart grids*, *i.e.*, redes elétricas modernas, capaz de integrar de maneira inteligente e eficiente os processos de aquisição de dados, monitoramento, diagnóstico e prognóstico do sistema elétrico de potência, com objetivo precípuo de fornecer energia elétrica com eficácia, confiabilidade, segurança e economicidade. De acordo com o cenário atual identifica-se a busca do setor de energia elétrica por tecnologias *smart grids*, isso se deve, principalmente, pela exigência por parte dos consumidores, da qualidade e confiabilidade da energia elétrica e da mudança na construção dos equipamentos.

Estas mudanças estruturais e filosóficas advindas do conceito de *smart grids* demandam um maior aporte das agências governamentais e centros de pesquisa no desenvolvimento de sistemas inteligentes flexíveis e integrados que abarquem a aquisição, monitoramento, análise e processamento de sinais com intuito de propiciar os meios necessários para implementar a automação, controle e tomada de decisão.

A presença de deformidades nas oscilografias de tensão e corrente é um indicativo de possíveis distúrbios. Dependendo do tipo de anormalidade, os distúrbios de tensão são classificados como afundamento (*sag*), elevação (*swell*), interrupções (*outage*), harmônicos (*harmonics*), oscilação (*flicker*) ou até mesmo a combinação destes eventos, e podem ser de duração instantânea, momentânea ou temporária.

Esses distúrbios estão relacionados com diferentes eventos nos sistemas de potência. Alguns são gerados pela operação normal dos sistemas, tais como chaveamento de banco de capacitores e entrada e saída de grandes cargas lineares, enquanto outros estão relacionados com eventos não operacionais do sistema, como as descargas atmosféricas e as faltas na rede elétrica (FERREIRA et al., 2011).

Os métodos tradicionais utilizados para o diagnóstico de falhas em campo baseiam-se na análise visual das oscilografias dependendo diretamente do conhecimento e experiência dos operadores, o que pode tornar o processo ineficiente e inseguro (LIMA, 2016). A partir disso, o emprego de sistemas inteligentes se torna uma possível solução para o problema de diagnóstico de distúrbios em sistemas elétricos de potência proporcionando segurança, rapidez e eficiência no planejamento de ações corretivas (SILVA et al., 2014).

As detecções precisam ser realizadas por meio de sistemas de baixa complexidade computacional, permitindo eficácia no monitoramento para aplicações de compensação e proteção e a implementação em tempo real. Normalmente, esses sistemas possuem como etapas a extração de parâmetros e a aplicação de um algoritmo que realize a detecção baseada nos dados extraídos (ALCÂNTARA; BARBOSA; FERREIRA, 2019). Em um sistema de monitoramento de qualidade de energia elétrica, a detecção dos distúrbios é o primeiro passo para uma correta

análise do problema, uma vez que a certa detecção dos distúrbios facilita o emprego das técnicas de classificação, identificação e compreensão de sinais (FERREIRA et al., 2009).

Por conseguinte, destacam-se alguns trabalhos: Em Dash et al. (2000) apresentou-se uma metodologia híbrida para classificar distúrbios em sistemas elétricos, através do combinador linear de Fourier e um sistema especialista *fuzzy*. Em Kaewarsa e Attakitmongcol (2005) é proposto um classificador empregando redes neurais artificiais e transformada *wavelet* para diagnóstico de distúrbios. Decanini et al. (2011a) apresentam um método para detecção automática e classificação de distúrbios de tensão para problemas relacionados à qualidade de energia usando técnicas de processamento de sinais e sistemas inteligentes. Em Barros (2013), a detecção e classificação de distúrbios de tensão utiliza Redes Neurais ARTMAP Euclidiana com treinamento continuado. Silva et al. (2014) utilizam o algoritmo Imuno-Neural para o diagnóstico de distúrbios de tensão, isto é, a junção de um algoritmo imunológico de seleção negativa com uma rede neural artificial ARTMAP-*Fuzzy*. Em Alcântara, Barbosa e Ferreira (2019) o sistema de detecção foi baseado em um modelo neural NARX (*Nonlinear Autoregressive with exogenous inputs*), a partir da aplicação de um filtro *notch* no sinal de tensão, a rede neural estima o valor do sinal um passo à frente e, por meio da análise de resíduo deste modelo, o distúrbio é detectado. Monteiro, Teixeira e Bretas (2022) empregam uma rede neural convolucional profunda para realizar as etapas de processamento e classificação de distúrbios de qualidade de energia elétrica, e ressaltam a facilidade de implementação, a alta exatidão e a não necessidade de pré-processamento do sinal. Gonzales-Abreu et al. (2021) propõem um método inovador de diagnóstico de distúrbios de qualidade de energia elétrica por meio do aprendizado profundo de redes neurais, com alta capacidade de caracterização de padrões, o qual possibilitou uma elevada taxa de acerto.

Dentro desta perspectiva, nesta pesquisa propõem-se uma metodologia para detectar e classificar cinco distúrbios de tensão (afundamento, elevação, harmônicas, interrupção e flutuação) em sistemas de distribuição de energia elétrica empregando técnicas de processamento de sinais, *i.e.*, transformada *wavelet*, análise multirresolução e conceito energia para compor os padrões representativos do estado operativo do sistema, seleção de atributos e técnicas de mineração de dados, utilizando os algoritmos IBK, J48 e redes neurais artificiais, para diagnosticar o estado operativo do sistema, assim, proporcionando maior segurança, confiabilidade e rentabilidade às concessionárias de energia elétrica. O sistema de diagnóstico desenvolvido é de fácil implementação e propicia alta performance computacional, visto que a seleção de atributos previamente realizada identifica as características essenciais para composição do padrão representativo do estado operativo do sistema, provendo consequentemente resultados promissores e capacidade de aplicação em tempo real.

2 ESTRUTURA DA METODOLOGIA PROPOSTA

A concepção do sistema inteligente para diagnóstico de distúrbio de tensão em sistemas de distribuição de energia elétrica baseia-se na ideia de extração dos sinais de tensão na saída do alimentador da subestação. Por conseguinte, a estrutura do sistema de diagnóstico pode ser descrita sucintamente em cinco etapas: aquisição de dados; processamento dos sinais para extração da assinatura do estado operativo do sistema; seleção de atributos, classificação do

estado operativo empregando técnica de mineração de dados, e a impressão do diagnóstico, conforme mostrado na Figura 1.

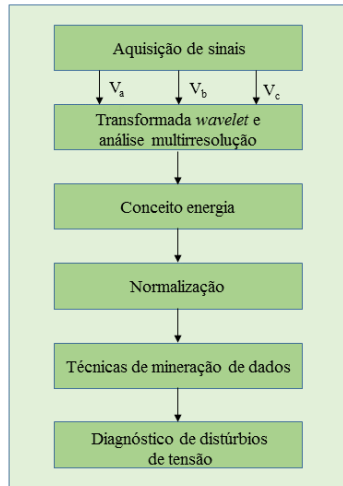


Figura 1 - Estrutura da metodologia proposta para diagnóstico de distúrbio de tensão.

Fonte: Elaborada pelos autores.

3 TRANSFORMADA WAVELET

A Transformada *Wavelet* (TW) é um conceito disponível que supera as limitações das técnicas baseadas na transformada de Fourier, uma vez que ela fornece a evolução temporal dos transitórios de frequência (DECANINI; TONELLI-NETO; MINUSSI, 2011b). O procedimento de análise por *wavelets* consiste em adotar uma função base denominada *wavelet* mãe, que serve como protótipo para todas as janelas usadas no processo. As janelas usadas são versões dilatadas ou comprimidas da *wavelet* mãe, também podendo ser transladadas no tempo (REIGOTA, 2007).

A adaptação da resolução do sinal permite processar apenas os detalhes relevantes para uma tarefa específica (MALLAT, 2009). Na Análise Multirresolução (AMR) decompõe-se o sinal empregando filtros passa-baixa (*h*) e filtros passa-alta (*g*). Os coeficientes de aproximação ($a_1[n]$, $a_2[n]$ e $a_3[n]$) obtidos na saída dos filtros passa-baixa correspondem aos componentes de alta escala e baixa frequência do sinal. Já os coeficientes de detalhe ($d_1[n]$, $d_2[n]$ e $d_3[n]$) provenientes da saída dos filtros passa-alta consistem nos componentes de baixa escala e alta frequência. A cada nível de resolução as saídas destes filtros são subamostradas por 2. A Figura 2 representa a análise multirresolução com 3 níveis de resolução, para mais níveis de resolução basta expandir o procedimento.

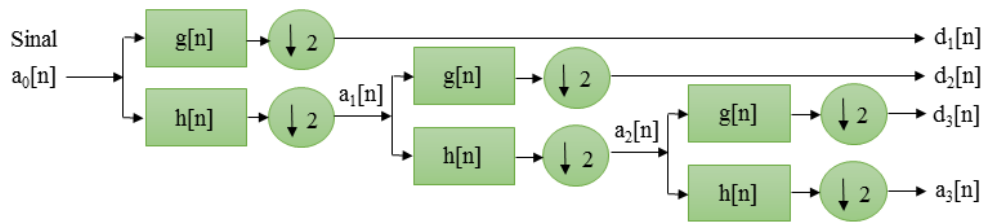


Figura 2 – Análise multirresolução com três níveis de resolução.

Fonte: Elaborada pelos autores.

A análise multirresolução é fundamentada no cálculo do coeficiente a_{m+1} por meio da convolução discreta no sinal a_m com um filtro passa-baixa, e do coeficiente d_{m+1} , com um filtro passa-alta, conforme exposto nas Equações (1) e (2), respectivamente (MALLAT, 2009).

$$a_{m+1}[n] = \sum_n h[k - 2n]a_m[k] \quad (1)$$

$$d_{m+1}[n] = \sum_n g[k - 2n]a_m[k] \quad (2)$$

3.1 Análise multirresolução aplicada para avaliação de sinais de tensão

A avaliação do estado operativo do sistema elétrico de potência aplicando a técnica de análise multirresolução em 2 níveis pode ser observada nas Figuras 3, 4 e 5. Mais especificamente, realizou-se a decomposição dos sinais da operação normal e dos distúrbios de tensão *sag* e *flicker*, a fim de se verificar o comportamento dos coeficientes de aproximação e detalhe, e assim elucidar a aplicação da AMR em sinais não estacionários. Dentro desta perspectiva, analisando as figuras supracitadas contata-se de forma límpida que ocorrendo variações bruscas do comportamento do sinal, *i.e.*, ocorrência de distúrbio, há a detecção via coeficiente de detalhe, o qual eleva repentinamente seu valor, provendo nesse sentido, uma informação de possível anormalidade no sistema e o momento de ocorrência, outro comportamento a ser observado corresponde ao comportamento consistente dos coeficientes de detalhe (com valores diferentes de zero em todo o domínio) em distúrbios com a presença da deformação durante todo o sinal. Sendo assim, em cada figura consta o sinal relativo ao estado operativo do sistema seguido dos coeficientes de aproximação de segundo nível e dos coeficientes de detalhe de segundo e primeiro nível.

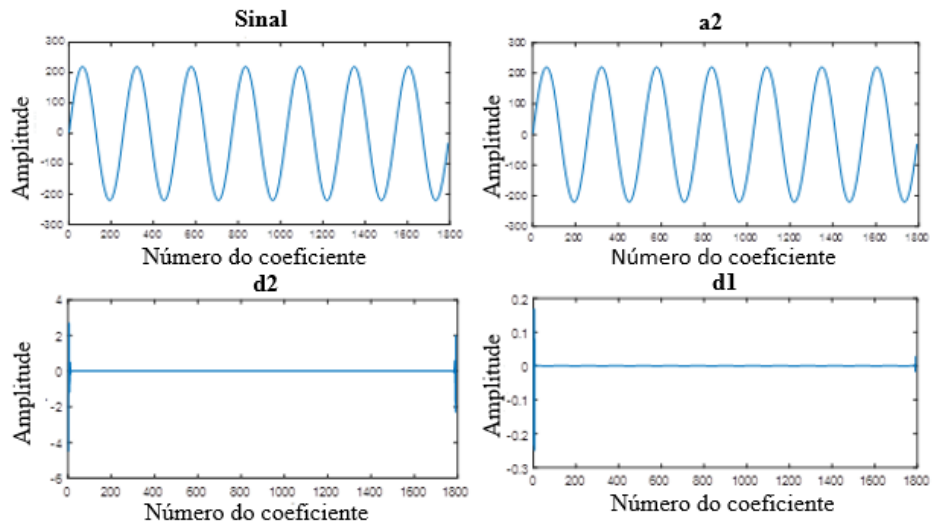


Figura 3 – Aplicação da técnica de AMR considerando a operação normal do sistema.

Fonte: Elaborada pelos autores.

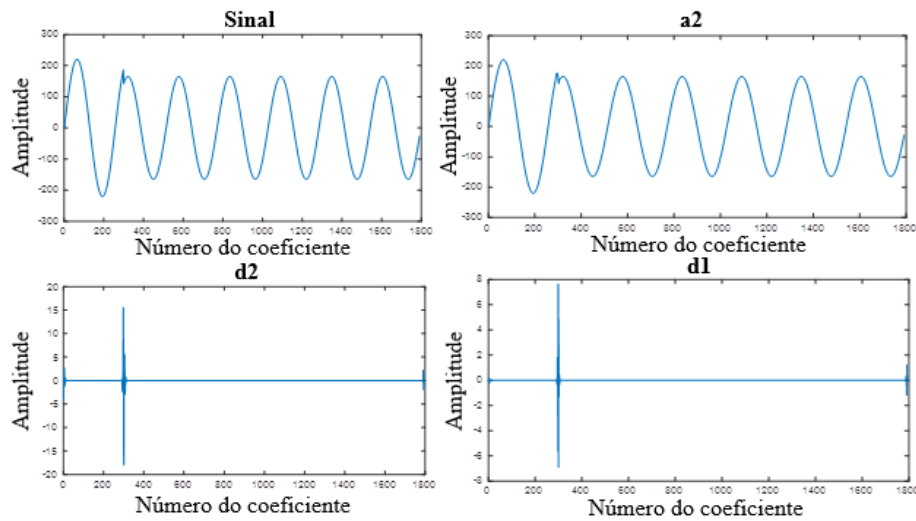


Figura 4 – Aplicação da técnica de AMR em distúrbio de tensão: *sag*.

Fonte: Elaborada pelos autores.

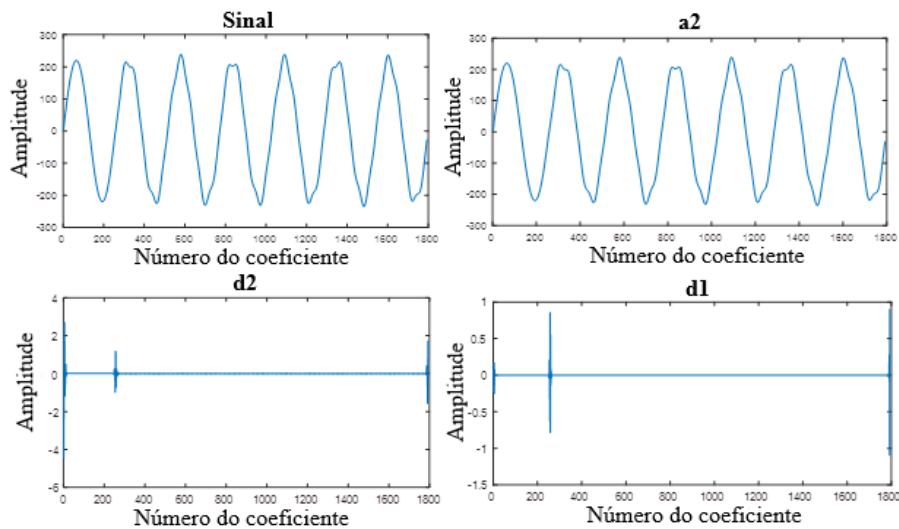


Figura 5 – Aplicação da técnica de AMR em distúrbio de tensão: *flicker*.

Fonte: Elaborada pelos autores.

A partir desse conjunto de figuras exemplificativas da aplicação da técnica de AMR na avaliação do estado operativo do sistema verifica-se de maneira clara o comportamento dos coeficientes de aproximação e detalhe, oriundos dos filtros passa-baixa (h) e dos filtros passa-alta (g).

4 TÉCNICAS DE MINERAÇÃO DE DADOS

Mineração de dados é uma área de pesquisa multidisciplinar, incluindo tecnologia de banco de dados, inteligência artificial, aprendizado de máquina, redes neurais artificiais, estatística, reconhecimento de padrões, sistemas baseados em conhecimento, recuperação da informação, computação de alto desempenho e visualização de dados (AMO, 2004) para descobrir padrões e regularidades em grandes conjuntos de dados auxiliando a tomada de decisão. A aplicação das técnicas de mineração de dados é facilitada devido a existência de ferramentas que dispõem de recursos de análise de dados e implementam algoritmos específicos.

A ferramenta de mineração de dados WEKA (HALL et al., 2009) foi a escolhida para este trabalho devido aos seguintes fatores: facilidade de aquisição e disponibilidade para *download* diretamente da página do desenvolvedor sem custo de utilização; presença de várias versões de algoritmos empregados na mineração de dados e disponibilidade de recursos estatísticos para comparar o desempenho dos algoritmos.

4.1 Seleção de atributos

De acordo com Paes, Plastino e Freitas (2013), a seleção de atributos é uma técnica muito explorada na área de mineração de dados, principalmente na tarefa de classificação. Tem como objetivo identificar os atributos mais relevantes a fim de reduzir o tempo de execução, aumentar

a capacidade preditiva do classificador retirando os atributos redundantes ou irrelevantes e obter uma representação compacta do conceito, produzindo os mesmos resultados analíticos.

4.1.1 Ranker

O método de avaliação *InfoGainAttributeEval* utiliza o ganho de informações para selecionar os atributos. Quando combinado com o método de busca *Ranker*, o qual coloca os atributos em ordem crescente utilizando o valor individual de seus ganhos. Seja A um atributo de uma base de dados e C o seu conjunto de classes, o cálculo da entropia do atributo classe antes e depois de observado o atributo A é dado pelas Equações (3) e (4), respectivamente:

$$E(c) = - \sum_{c \in C} p(c) \log_2 p(c) \quad (3)$$

em que $p(c)$ é a probabilidade da classe c ocorrer na base de dados.

$$E(C|A_i) = - \sum_{a \in A} p(a) \sum_{c \in C} p(c|a) \log_2 p(c|a) \quad (4)$$

em que $p(a)$ é a probabilidade do valor a ocorrer na base de dados e $p(c|a)$ a probabilidade da classe c ocorrer, dado, que o valor de atributo a ocorreu. A redução causada na entropia C devido a informação fornecida pelo atributo A é chamado de ganho de informação. De acordo com Merschmann (2007), cada atributo A_i da base de dados é associado a um valor correspondente ao ganho de informação GI , calculado de acordo com a Equação (5):

$$GI_i = E(C) - E(C|A_i) \quad (5)$$

Desta forma, os atributos associados aos maiores ganhos de informações serão os selecionados em ordem crescente.

4.2 Classificação

A classificação é uma função de aprendizado que mapeia conjunto de dados de entrada, o objetivo do algoritmo de classificação é encontrar alguma correlação entre os atributos e uma classe, de modo que o processo de classificação possa reconhecer outros dados desconhecidos. Os algoritmos de classificação que foram usados no trabalho são: J48, IBK e a rede neural artificial Perceptron Multicamadas.

4.2.1 Algoritmo J48

O algoritmo J48 é um tipo de algoritmo que tem como finalidade gerar uma árvore de decisão baseada no conjunto de dados de treinamento. As árvores de decisão apresentam um baixo custo computacional, e por isso, tem sido muito utilizada em problemas de classificação.

A técnica de classificação árvore de decisão, consiste na separação dos dados em subconjuntos de forma hierárquica. Em cada nó interno é indicado um teste sobre o atributo, cada ramo representa um resultado do teste, e os nós terminais correspondem às classes ou

distribuição de classes. Na árvore de decisão tem-se o nodo raiz (primeiro dado da hierarquização), os nodos internos (todos os nós que estão entre 2 outros nós), e os nodos terminais que são os mais próximos da saída.

O processo de construção da árvore de decisão começa de cima para baixo, e é chamado de indução. Em cada iteração, a partir da base de dados de treinamento, os algoritmos procuram pelo atributo que melhor separa as classes para realizar a ramificação da árvore (MERSCHMANN, 2007).

4.2.2 Algoritmo IBK

O algoritmo IBK utiliza o método de k-vizinhos, que consiste em encontrar os exemplos de treino mais similares aos atributos do exemplo de teste. Neste método, cada tupla é classificada como um ponto de dado em um espaço n-dimensional, sendo n o número de atributos (BLOMBERG, 2014). Desta forma, para classificar uma nova instância, o classificador utiliza a distância métrica para determinar as k-instâncias de treinamento mais similares à instância que deseja classificar. A função mais utilizada para cálculo da distância é a Euclidiana, no qual, a distância entre duas instâncias (X e Y) é dada a partir da Equação (6).

$$d(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (6)$$

4.2.3 Perceptron Multicamadas

A estrutura e os algoritmos relativos à Rede Neural Artificial (RNA) Perceptron Multicamadas são inspirados em um modelo de redes neurais biológicas, onde pequenas unidades de processamento chamadas neurônios são organizadas em camadas que, se bem configuradas, são capazes de se aproximar de qualquer função. Uma RNA é formada por um conjunto de unidades de entrada conectadas a unidades de saída, no qual, cada associação possui um peso associado. A fase de treinamento de uma RNA envolve o ajuste de seus pesos de modo a capacitá-la prever corretamente a classe de instâncias desconhecidas (MERSCHMANN, 2007). A rede Perceptron Multicamadas, apresenta uma ou mais camadas intermediárias (escondidas) de neurônios e uma camada de saída.

De acordo com Merschmann (2007), as redes multicamadas só foram possíveis devido ao algoritmo de treinamento *backpropagation*, que é baseado na retropropagação dos erros para realizar os ajustes dos pesos das camadas intermediárias. Um vetor de entradas $[X = (x_1, x_2, \dots, x_m)]$ formado por m elementos, no qual, cada elemento é ponderado por seu respectivo peso $[W = (w_{k1}, w_{k2}, \dots, w_{km})]$, que determina a atividade do neurônio. Por fim, aplica-se uma função de ativação para se obter a saída do neurônio.

A função sigmoide trata-se de uma função estritamente crescente exibindo um balanceamento satisfatório entre o comportamento linear e o comportamento não linear (HAYKIN, 1999). Além disso, definem-se os parâmetros a seguir:

TrainingTime : Quantidade de vezes no qual o algoritmo irá realizar a atualização dos pesos;

HiddenLayer: Quantidade de neurônios na camada escondida;

Learning: Taxa de aprendizagem, controla a velocidade de adaptação da rede;

Momentum: Pondera a variação das matrizes de pesos entre duas iterações sucessivas.

5 METODOLOGIA

O sistema inteligente para diagnóstico de distúrbios de tensão em sistemas de distribuição de energia elétrica foi concebido de forma modular o que propicia maior flexibilidade sendo basicamente desenvolvido em quatro processos: aquisição de dados, processamento de sinais, seleção de atributos e classificação.

5.1 Aquisição de dados

Conforme proposto em publicações especializadas, muitos distúrbios podem ser gerados variando alguns parâmetros (ABDEL-GALIL et al., 2004; ERISTI; UÇAR; DEMIR, 2010), proporcionando meios para a composição de uma base de dados pertinente para o treinamento e teste de sistemas inteligentes para automatização de processos. Este procedimento é extremamente necessário tendo em vista o nível de monitoramento atual de diversos alimentadores, assim, auxilia na composição de um banco de dados pertinente e suficiente para elaborar o processo de automação. Nesta perspectiva, as equações paramétricas referentes aos tipos de distúrbios e as variações em seus parâmetros são apresentadas na Tabela 1. Este conjunto de informações foi implementado no *software* Matlab.

Tabela 1 – Equações relativas ao estado operativo do sistema.

Estado Operativo	Equações	Variação dos Parâmetros
Operação normal	$v(t) = A(\text{sen}(wt))$	-
Afundamento (<i>sag</i>)	$v(t) = A(1 - \alpha(u(t - t_1) - u(t - t_2))) (\text{sen}(wt), t_1 < t_2)$ $u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$	$0,1 \leq \alpha \leq 0,9$ $T \leq t_2 - t_1 \leq 7T$
Elevação (<i>swell</i>)	$v(t) = A(1 + \alpha(u(t - t_1) - u(t - t_2))) (\text{sen}(wt), t_1 < t_2)$ $u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$	$0,1 \leq \alpha \leq 0,8$ $T \leq t_2 - t_1 \leq 7T$

Estado Operativo	Equações	Variação dos Parâmetros
Interrupção (Outage)	$v(t) = A(1 - \alpha(u(t - t_1) - u(t - t_2))) \text{sen}(wt), t_1 < t_2$ $u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$	$0,9 \leq \alpha \leq 1$ $T \leq t_2 - t_1 \leq 7T$
Harmônicas (Harmonics)	$v(t) = A \left(\begin{array}{l} \alpha_1 \text{sen}(wt) + \alpha_3 \text{sen}(3wt) \\ + \alpha_5 \text{sen}(5wt) + \alpha_7 \text{sen}(7wt) \end{array} \right)$	$\alpha_1 = 1$ $0,05 \leq \alpha_{3,5,7} \leq 0,15$
Flutuação (Flicker)	$v(t) = A(1 + \alpha \text{sen}(\beta wt)) \text{sen}(wt)$	$0,01 \leq \alpha \leq 0,1$ $2\text{Hz} \leq \beta \leq 8\text{Hz}$

Fonte: Abdel-Galil et al. (2004).

Destaca-se que os parâmetros α e β foram gerados de forma pseudoaleatória, dentro dos intervalos preestabelecidos, para composição do banco de dados para simulação.

5.2 Processamento de sinais

O processamento de sinais foi realizado via transformada *wavelet* discreta e análise multirresolução, mais especificamente, em cinco níveis de resolução, obtendo informações no domínio tempo-frequência do comportamento do sistema por meio do *software* Matlab.

Na análise multirresolução decompõe-se o sinal empregando filtros passa-baixa e filtros passa-alta. Os coeficientes de aproximação e detalhe são agregados por nível de resolução usando o conceito de energia, que corresponde à uma técnica de *data fusion* que permite caracterizar sinais não estacionários. Ainda, os dados foram normalizados com intuito de se obter maior generalização do método e simplificar o processo de análise e classificação.

5.3 Seleção de atributos

A seleção de atributos foi usada para reduzir o número de características relativas à assinatura do estado operativo do sistema, eliminando os atributos redundantes e irrelevantes, e assim, melhorando a performance do algoritmo de aprendizagem. Utilizou-se o método de seleção *Ranker*, presente no *software* WEKA, que classifica as características em ordem crescente de relevância, permitindo a simplificação do sistema e a consequente melhoria no desempenho computacional a partir do método de avaliação *InfoGainAttributeEval*.

5.4 Classificação

Neste estudo, realizaram-se três experimentos com o objetivo de comparar o desempenho dos algoritmos de mineração de dados aplicados ao domínio do problema. A análise dos

resultados contemplando o emprego de técnicas distintas propicia um diagnóstico mais efetivo da composição da ferramenta para análise do estado operativo do sistema. Os algoritmos J48, IBK e Perceptron Multicamadas (*backpropagation*) foram escolhidos visto as diferentes técnicas que cada um utiliza, sendo elas: árvore de decisão, K-vizinho e redes neurais artificiais, respectivamente. Todas as simulações foram realizadas no *software* WEKA.

O método para validação escolhido foi o *cross-validation* com 10 partições. Esta técnica de validação cruzada consiste em dividir o conjunto de dados em k partições. A cada iteração, treino e teste, um conjunto formado por k-1 subconjuntos são utilizados para treinamento e o subconjunto restante para teste. Este mesmo processo repete por k vezes, provendo maior confiabilidade aos resultados obtidos.

6 APLICAÇÕES E RESULTADOS

Neste tópico serão apresentados os resultados da metodologia proposta para detectar e classificar cinco tipos de distúrbios de tensão em sistemas de distribuição de energia elétrica usando técnicas de mineração de dados. O intuito é analisar a performance do diagnóstico de distúrbio de tensão por meio de sistemas inteligentes, mais especificamente, empregando 3 técnicas distintas, com objetivo de se obter uma melhor compreensão do processo.

6.1 Base de dados

No que tange a realidade dos sistemas de distribuição de energia elétrica, evidencia-se ainda diversas subestações com baixo nível de monitoramento e automatização, não possuindo um conjunto significativo de dados relativos à operação do sistema. Neste sentido, torna-se imprescindível o uso de métodos solidamente testados e publicados na literatura especializada para composição de banco de dados que refletem a operação do sistema. Portanto, a aquisição dos sinais de tensão foi obtida por meio de simulação no *software* Matlab, considerando as funções representativas do estado operativo do sistema apresentadas na Tabela 1. O nível de tensão de linha utilizado foi de 13800 V, frequência de 60 Hz e frequência de amostragem de 15400 Hz, abrangendo 7 ciclos e com passo de análise, *e.g.*, de 6 ciclos.

Dentro desta perspectiva, realizaram-se 300 simulações para cada estado operativo da rede, totalizando 1800 informações.

6.2 Processamento e classificação

O procedimento contínuo de detecção e classificação de anomalias é realizado empregando-se a *mother wavelet* Daubechies com filtro de quarta ordem (db4) na AMR com cinco níveis de resolução. A representação gráfica do comportamento dos coeficientes de aproximação e detalhe devidamente normalizados está exposta na Figura 6.

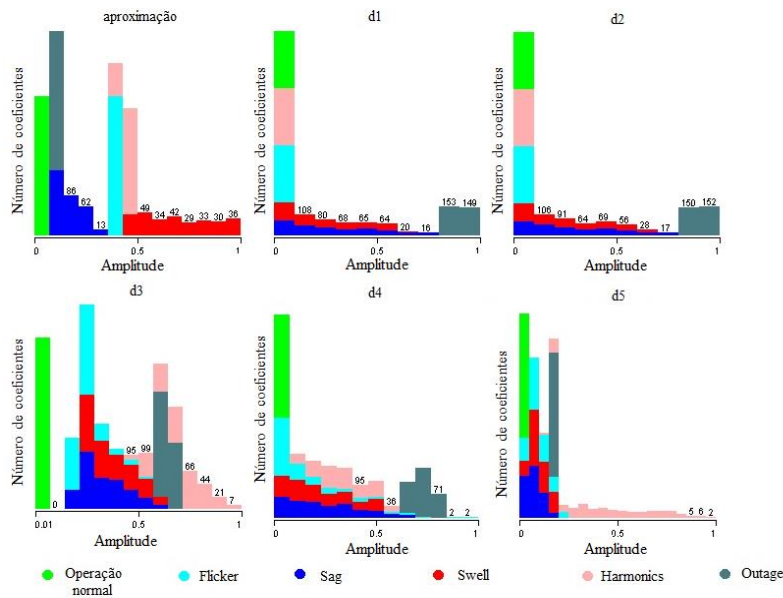


Figura 6 – Características dos estados operativos do sistema.

Fonte: Elaborada pelos autores.

A Figura 6 propicia uma análise macro do comportamento das características dos estados operativos do sistema, porém torna-se essencial o emprego de técnicas que proveem uma análise micro, *i.e.*, mais refinada sobre tais índices. Por conseguinte, empregou-se a técnica de seleção de atributos *InfoGainAttributeEval Ranker*, com objetivo precípuo de se determinar quais são as informações mais relevantes para o processo de diagnóstico do estado operativo de sistemas de distribuição de energia elétrica. Assim, apresenta-se na Tabela 2 o resultado obtido do procedimento para seleção dos principais atributos.

Tabela 2 – Seleção de atributos.

<i>Ranker</i>	<i>Atributos</i>
2,487	Aproximação
2,267	d1
2,208	d2
1,661	d3
1,63	d4
1,465	d5

Fonte: Elaborada pelos autores.

6.2.1 Diagnóstico do estado operativo do sistema empregando o algoritmo J48

A Tabela 3 apresenta a matriz confusão considerando apenas os coeficientes de aproximação e de detalhe de primeiro nível, de acordo com a seleção de atributos exposta na Tabela 2, e empregando o algoritmo J48. A eficiência foi de 99,6667%, resultando em apenas 6 erros do total de 1800 casos analisados.

Tabela 3 – Matriz confusão considerando o algoritmo J48.

<i>Sag</i>	<i>Swell</i>	<i>Flicker</i>	<i>Outage</i>	<i>Harmonics</i>	Operação normal
299	1	0	0	0	0
0	300	0	0	0	0
0	0	297	0	3	0
1	0	0	299	0	0
0	1	0	0	299	0
0	0	0	0	0	300

Fonte: Elaborada pelos autores.

A matriz confusão apresenta na diagonal principal a quantidade de acerto de cada estado operativo da rede, os demais elementos significam erros na classificação, *i.e.*, as linhas da matriz de confusão correspondem aos estados operativos a serem classificados e as colunas referem-se à classificação obtida por meio do sistema de diagnóstico. Além da matriz confusão, por meio da Figura 7 é possível visualizar os dados que não foram classificados corretamente.

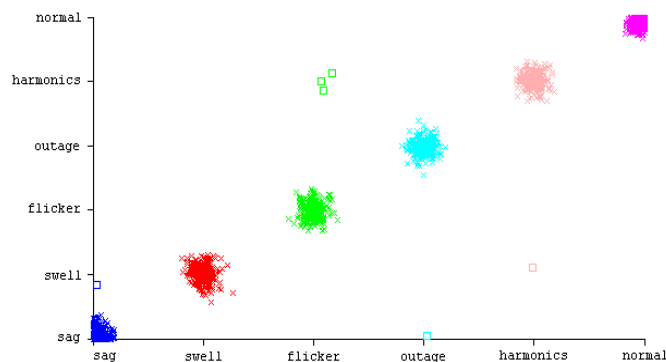


Figura 7 – Resultado da classificação do estado operativo do sistema empregando o algoritmo J48.

Fonte: Elaborada pelos autores.

6.2.2 Diagnóstico do estado operativo do sistema empregando o algoritmo IBK

A Tabela 4 apresenta a matriz confusão considerando apenas os coeficientes de aproximação e detalhe de primeiro nível, de acordo com a seleção de atributos da Tabela 2, e utilizando o algoritmo IBK. A eficiência foi de 99,9444%, resultando em apenas 1 erro do total de 1800 casos avaliados com K=1.

Tabela 4 – Matriz confusão considerando o algoritmo IBK.

<i>Sag</i>	<i>Swell</i>	<i>Flicker</i>	<i>Outage</i>	<i>Harmonics</i>	Operação normal
299	0	0	1	0	0
0	300	0	0	0	0
0	0	300	0	0	0
0	0	0	300	0	0
0	0	0	0	300	0
0	0	0	0	0	300

Fonte: Elaborada pelos autores.

Analisando a Tabela 4 observa-se um único erro na classificação do estado operativo do sistema, o qual consistiu em uma classificação como *outage* de um distúrbio *sag*. A visualização da classificação pode ser realizada por meio da Figura 8.

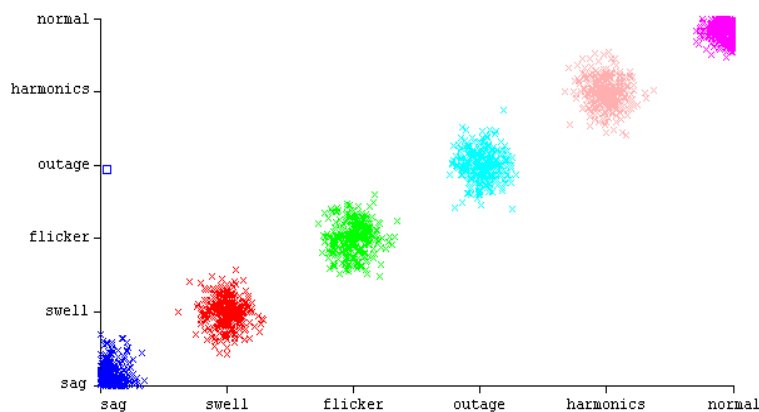


Figura 8 – Resultado da classificação do estado operativo do sistema empregando o algoritmo IBK.

Fonte: Elaborada pelos autores.

6.2.3 Diagnóstico do estado operativo do sistema empregando a rede neural artificial Perceptron Multicamadas

A estrutura da rede neural artificial empregada para composição do módulo de diagnóstico foi concebida com 1 camada oculta composta por 4 neurônios e 1 camada de saída com 6 neurônios. A função de ativação utilizada foi a sigmoide, considerando 500 épocas, taxa de aprendizagem igual a 0,3 e momentum igual a 0,2.

A Tabela 5 apresenta a matriz confusão considerando apenas os coeficientes de aproximação e detalhe de primeiro nível, de acordo com a seleção de atributos, e empregando a rede neural artificial Perceptron Multicamadas. A eficiência foi de 99,8889%, resultando em apenas 2 erros do total de 1800 casos analisados.

Tabela 5 – Matriz confusão considerando a rede neural artificial Perceptron Multicamadas.

<i>Sag</i>	<i>Swell</i>	<i>Flicker</i>	<i>Outage</i>	<i>Harmonics</i>	Operação normal
298	0	0	2	0	0
0	300	0	0	0	0
0	0	300	0	0	0
0	0	0	300	0	0
0	0	0	0	300	0
0	0	0	0	0	300

Fonte: Elaborada pelos autores.

Avaliando a Tabela 5, observa-se que 2 distúrbios *sag* foram classificados erroneamente como *outage*. A Figura 9 apresenta a classificação obtida empregando rede neural artificial.

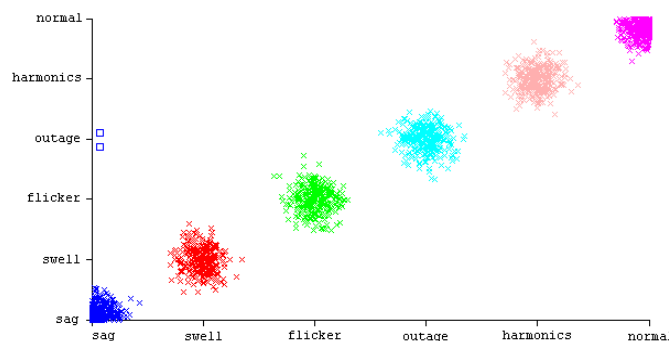


Figura 9 – Resultado da classificação do estado operativo do sistema empregando a rede neural artificial Perceptron Multicamadas.

Fonte: Elaborada pelos autores.

6.3 Performance computacional

Os diagnósticos resultantes considerando os 5 coeficientes de detalhe e o coeficiente de aproximação como entrada do sistema inteligente composto pelo algoritmo J48 ou pelo algoritmo IBK tiveram os mesmos resultados quando se utilizou apenas os coeficientes de aproximação e de detalhe de primeiro nível. Já a redução do número de entradas na rede Perceptron Multicamadas ocasionou aumento de apenas um erro. Assim, comprova-se a eficácia da seleção de atributos proposta, a qual diminuiu a quantidade de dados de entrada para o processamento e mantendo a mesma qualidade nos resultados.

Por fim, o procedimento completo de diagnóstico de distúrbios de tensão em sistemas de distribuição empregando as três técnicas de mineração de dados foi executado em um tempo na ordem de 100 ms, utilizando um notebook Intel core i7, 2,4GHz com 8GB de memória RAM.

7 CONCLUSÕES

Neste trabalho é proposta uma metodologia para detectar e classificar cinco tipos de distúrbios de tensão, sendo eles: afundamento, elevação, interrupção, harmônicas e flutuação, em sistemas de distribuição de energia elétrica empregando técnicas de processamento de sinais e de mineração de dados. Com o intuito precípua de desenvolver uma ferramenta de apoio à tomada de decisão, que provê subsídios para que os procedimentos relacionados à automação de subestação sejam rápidos, seguros e confiáveis, tornando o dia-a-dia dos operadores menos exaustivo. Dentro desta perspectiva, o sistema inteligente para diagnóstico do estado operativo do sistema foi concebido de forma flexível empregando modernas técnicas para a análise e classificação dos sinais.

Os coeficientes de detalhe e de aproximação, que representam o estado operativo do sistema, são obtidos usando a *wavelet* da família *Daubechies* com filtro de quarta ordem (db4). A utilização do conceito energia fornece os índices comportamentais do sistema. A classificação dos distúrbios foi realizada usando os algoritmos J48, IBK e a rede neural artificial Perceptron Multicamadas com treinamento via algoritmo *backpropagation*, considerando o método *cross-validation* com 10 partições. Sendo que, o sistema de diagnóstico apresentou eficiência superior a 99%. Ressalta-se que os três algoritmos empregados possuem características distintas, enquanto o algoritmo IBK, por meio da técnica dos k-vizinhos mais próximos, proveu o melhor desempenho, as redes neurais Perceptron Multicamadas possuem alta capacidade de generalização, porém não propiciam uma interpretação lógica e fácil do resultado apresentado, já o algoritmo J48 apresenta uma estrutura clara para compreensão do diagnóstico realizado, visto que corresponde à uma árvore de decisão. Portanto, performance, desempenho computacional, facilidade de entendimento do diagnóstico no que tange a técnica de mineração de dados e flexibilidade são requisitos essenciais que devem ser avaliados para a devida implementação do sistema de diagnóstico.

Os resultados obtidos evidenciam a eficácia dos métodos propostos para diagnóstico de distúrbios de tensão em sistemas elétricos de potência, propiciando informações pertinentes aos operadores que refletirá em maior confiabilidade, segurança e lucratividade às concessionárias.

AGRADECIMENTOS

Os autores agradecem ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo pelo apoio e suporte fornecidos para o desenvolvimento deste trabalho.

REFERÊNCIAS

- ABDEL-GALIL, T. K.; KAMEL, M.; YOUSSEF, A. M.; EL-SAADANY, E. F.; SALAMA, M. M. A. Power quality disturbance classification using inductive inference approach. **IEEE Transactions on Power Delivery**, v. 19, n. 4, 1812–1818, 2004.
- ALCÂNTARA, I. F.; BARBOSA, B. H. G.; FERRREIRA, Danton Diego. Detecção de Distúrbios em Qualidade de Energia usando modelos NARX Neurais. In: **SBAI - Simpósio Brasileiro de Automação Inteligente**, 14., 2019, Ouro Preto. Lavras: 2019. p. 1707-1712.
- AMO, S. **Técnicas de mineração de dados**. Universidade Federal de Uberlândia, Faculdade de Computação, 2004. Disponível em https://www.researchgate.net/profile/Sandra-Amo/publication/260300816_Tecnicas_de_Minerao_de_Dados/links/54230bd80cf290c9e3ae25e3/Tecnicas-de-Minerao-de-Dados.pdf acessado em setembro de 2021.
- BARROS, A. C. **Detecção e classificação de distúrbios de tensão usando redes neurais ARTMAP euclidiana modificada com treinamento continuado**. 2013. 99 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Universidade Estadual Paulista Júlio de Mesquita Filho, Ilha Solteira, 2013.
- BLOMBERG, L. C. **Um algoritmo evolutivo para indução de árvores de regressão robusto a valores ausentes**. 2014. 97 f. Tese (Doutorado) - Curso de Ciências da Computação, Universidade Católica do Rio Grande do Sul, Porto Alegre, 2014.
- DASH, P. K.; MISHRA, S.; SALAMA, M. M. A.; LIEW, A. C. Classification of power system disturbances using a fuzzy expert system and a fourier linear combiner. **IEEE Transactions on Power Delivery**, New York, v. 15, n. 2, p. 472–477, 2000.
- DECANINI, J. G. M. S.; TONELLI-NETO, M. S.; MALANGE, F. C. V.; MINUSSI, C. R. Detection and classification of voltage disturbances using a fuzzy-ARTMAP-wavelet network. **Electric Power Systems Research**, Lausanne, v. 81, n. 12, p. 2057-2065, 2011a.
- DECANINI, J. G. M. S.; TONELLI-NETO, M. S.; MINUSSI, C. R. Robust fault diagnosis in power distribution systems based on fuzzy ARTMAP neural network-aided evidence theory. **IET Generation, Transmission and Distribution**, Stevenage, v. 6, n. 11, p. 1112-1120, 2011b.
- Eristi, H., Uçar, A., Demir, Y. Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines. **Electric Power Systems Research**, Lausanne, v. 8, n. 7, p. 743–752, 2010.
- FERREIRA, D. D.; MARQUES, C. A. G.; CERQUEIRA, A. S.; DUQUE, C. A.; RIBEIRO, Moisés Vidal. Sistema automático de detecção e classificação de distúrbios elétricos em qualidade da energia elétrica. In: **SBA CONTROLE E AUTOMAÇÃO**, 20., 2009. **Sba: Controle & Automação Sociedade Brasileira de Automatica**. [S.L.]: Fapunifesp (SciELO), 2009. p. 53-62.

FERREIRA, D. D.; CERQUEIRA, A. S.; DUQUE, C. A.; SEIXAS, J. M.; RIBEIRO, M. V. Sistema Automático de classificação de distúrbios elétricos múltiplos e isolados. **Controle e Automação**, Minas Gerais, v. 1, n. 22, p. 39-48, fev. 2011.

GONZALES-ABREU, A.; DELGADO-PRIETO, M.; OSORNIO-RIOS, R.; SAUCEDO-DORANTES, J.; ROMERO-TRONCOSO, R. A novel deep learning-based diagnosis method applied to power quality disturbances. **Energies**, v. 14, n. 10, p. 1–17, 2021.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P. and Witten. I.H. (2009) “**The WEKA data mining software: an update**” SIGKDD Explorations, Volume 11, Issue 1.

HAYKIN, S. **Neural networks: a comprehensive foundation**. 2. ed. [S.l.]: Pearson Education, 1999.

KAEWARSA, S.; ATTAKITMONGCOL, K. Wavelet-based neural network approach to power quality disturbance recognition. In: 2005 INTERNATIONAL POWER ENGINEERING CONFERENCE, 7., 2005, N. **2005 International Power Engineering Conference**. [S.L.]: Ieee, 2005.

LIMA, F. P. A. **Diagnóstico de distúrbios de tensão em sistemas de distribuição baseado num Sistema imulológico artificial com aprendizado continuado**. 2016. 101 f. Tese (Doutorado em Engenharia Elétrica) - Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2016.

MALLAT, S. **A wavelet tour of signal processing the sparse way**. 3. ed. [S.l.]: Elsevier, 2009.

MERSCHMANN, L. H. C. **Classificação probabilística baseada em análise de padrões**. 2007. 103 f. Tese (Doutorado) - Curso de Computação, Universidade Federal Fluminense, Niterói, 2007.

MONTEIRO, R. V. A.; TEIXEIRA, R. F. S.; BRETAS, A. S. Power quality disturbances diagnosis: A 2D densely connected convolutional network framework. **Electric Power Systems Research**, Lausanne, v. 212, p. 1–13, 2022.

PAES, B. C.; PLASTINO, A.; FREITAS, A. A. **Seleção de atributos aplicada à classificação hierárquica**. Universidade Federal Fluminense, University of Kent 2013. Disponível em <<https://homepages.dcc.ufmg.br/~gfrancis/kdmile/papers/KDMiLe.pdf>> Acesso em 15 de Outubro de 2021.

REIGOTA, N. S. **Comparação da transformada wavelet discreta e da transformada do cosseno para compressão de imagens de impressão digital**. 2007. 143 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Universidade de São Paulo, São Carlos, 2007.

SILVA, J. C.; LIMA, F. P. A.; LOPES, M. L. M.; MINUSSI, C. R. Diagnóstico de Distúrbios de Tensão em Sistemas de Distribuição de Energia Elétrica Usando um Algoritmo Imuno-Neural. In: **SBSE - Simpósio Brasileiro de Sistemas Elétricos**, Foz do Iguaçu. Foz do Iguaçu: 2014.