

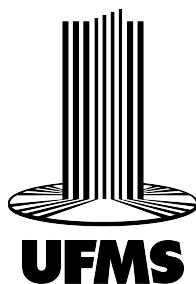
UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE FÍSICA
CURSO DE FÍSICA BACHARELADO

Classificação de Distúrbios de Qualidade de Energia Elétrica via Transformada Wavelet Packet e Aprendizado de Máquina

Rodrigo de Carvalho

Campo Grande - MS

09 de Junho de 2026



UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE FÍSICA
CURSO DE FÍSICA BACHARELADO

Classificação de Distúrbios de Qualidade de Energia Elétrica via Transformada Wavelet Packet e Aprendizado de Máquina

Rodrigo de Carvalho

Trabalho de Conclusão de Curso apresentado
como exigência para obtenção do grau de Ba-
charelado em Física da Universidade Federal de
Mato Grosso do Sul – UFMS.

Orientador: Prof. Dr. Marcus Vinícius Monteiro Varanis

Campo Grande - MS

09 de Junho de 2026

Classificação de Distúrbios de Qualidade de Energia Elétrica via Transformada Wavelet Packet e Aprendizado de Máquina

Trabalho de Conclusão de Curso apresentado como exigência para obtenção do grau de Bacharelado em Física da Universidade Federal de Mato Grosso do Sul – UFMS.

Banca Examinadora:

Prof. Dr. Marcus Vinícius Monteiro Varanis

Prof. Dr. Cicero Rafael Cena da Silva

Me. João Vitor Oliveira Ribas

Campo Grande - MS

09 de Junho de 2026

Agradecimentos

Aos meus pais, Valdez e Jonir, por sempre acreditarem em mim e me apoiarem, mesmo com todas as dificuldades, nessa escolha e em todos os aspectos da minha vida.

À minha irmã, Carolina, por todo o amor e paciência que teve comigo durante toda a minha vida para que eu pudesse aprender a ser uma pessoa melhor e seguir o seu exemplo de bondade e caráter.

Aos meus avós, José, Zalma e Anna, por todo o amor, apoio e estarem sempre com as portas abertas e dispostos a me guiarem mesmo sem entender direito as dificuldades e o que eu faço.

À minha madrinha, Alda, por ter cuidado de mim desde a minha infância e me tratar como um filho.

Aos meus tios, Clélia, Silvestre e Laerte, pelas palavras de sabedoria e orientação quando tive dúvidas em relação às provas da vida que escolhi.

A todos os servidores da Universidade Federal do Mato Grosso do Sul, especialmente os do Instituto de Física, por me darem a oportunidade de usar o espaço para meus estudos e por todo o zelo em questões administrativas, me fazendo ver a UFMS como uma segunda casa.

A todos os professores, que me deram oportunidades e ensinamentos que levarei para a vida, em especial ao professor Marcus Varanis, por me orientar com tanto zelo, levando em conta as minhas dificuldades e vantagens.

Aos amigos de curso, em especial ao Yuri, Guilherme e João por me fazerem enxergar meu potencial e por me ajudarem a aprender como se realmente estuda.

A todos os outros amigos, que me mantiveram são durante os momentos difíceis, me ajudaram a descontrair e sempre acreditaram em mim, em especial ao Gustavo e seu irmão Marcus, por sempre me motivarem e ouvirem minhas dores da vida acadêmica e me darem tantas oportunidades; ao Jorge, pelos conselhos nos momentos mais tortuosos e dividir paixões da vida comigo; à Nathália, por estar presente como uma irmã mais velha e sempre me dar conselhos e à Miyuki, pela preocupação, constante vontade de me animar e sempre estar presente para me ouvir.

“A vida do espírito e o conhecimento podem ser expressos como um jogo do amor consigo mesmo; mas essa ideia decai para a insipidez se nela faltam a seriedade, a dor, a paciência e o trabalho do negativo.”

G.W.F. Hegel

Resumo

A crescente complexidade das redes elétricas modernas e a proliferação de cargas não lineares intensificam a suscetibilidade dos sistemas a Distúrbios de Qualidade de Energia Elétrica (PQD), exigindo métodos de diagnóstico automáticos, escaláveis e robustos. O desafio central na classificação automática de PQDs reside na identificação de eventos transitórios e distúrbios combinados operando sob severos regimes de ruído e variações estocásticas. Neste contexto, este trabalho desenvolve e avalia um *pipeline* computacional para a classificação de 16 classes de PQD baseadas na norma IEEE 1159. Para garantir a robustez contra o sobreajuste, construiu-se um banco de dados massivo composto por 160.000 sinais sintéticos, submetidos a variações paramétricas aleatórias (profundidade, duração, ângulo de fase) e altos níveis de Ruído Branco Gaussiano Aditivo (AWGN). A representação no domínio tempo-frequência foi conduzida pela Transformada *Wavelet Packet* (WPT, família db8, nível 4), extraído-se um vetor de 208 atributos com foco em formulações entrópicas (*Log-Energy*, Permutação, Shannon e Rényi). O *benchmark* preditivo comparou a Análise Discriminante Linear (LDA), o *Random Forest* (RF) e a rede Perceptron Multicamadas (MLP). Os resultados provaram matematicamente a não linearidade do espaço de características sob estresse ruidoso, evidenciada pela estagnação do LDA em 85,39% de acurácia e alto *log-loss*. A superioridade foi atingida pela MLP, que atingiu 95,21% de acurácia global e 0,95 de Macro-F1. A análise de poder discriminativo pelo *Fisher Score* atestou que as métricas de entropia dominam a separabilidade das classes, preservando a assinatura topológica da desorganização da falha mesmo quando a amplitude é mascarada pelo ruído. Conclui-se que a extração *Wavelet* orientada à entropia, aliada a classificadores não lineares clássicos, entrega generalização competitiva com o estado da arte prescindindo do elevado custo computacional de arquiteturas profundas.

Palavras-chaves: Qualidade de Energia Elétrica. Transformada *Wavelet Packet*. Entropia de Sinais. Aprendizado de Máquina. Redes Neurais Artificiais.

Abstract

The increasing complexity of modern electrical power networks and the proliferation of non-linear loads intensify the susceptibility of systems to Power Quality Disturbances (PQDs), requiring automatic, scalable, and robust diagnostic methods. The central challenge in the automatic classification of PQDs lies in the identification of transient events and combined disturbances operating under severe noise regimes and stochastic variations. In this context, this work develops and evaluates a computational pipeline for the classification of 16 PQD classes based on the IEEE 1159 standard. To ensure robustness against overfitting, a massive database composed of 160,000 synthetic signals was constructed, subjected to random parametric variations (depth, duration, phase angle) and high levels of Additive White Gaussian Noise (AWGN). The time-frequency domain representation was conducted through the Wavelet Packet Transform (WPT, db8 family, level 4), extracting a vector of 208 attributes focused on entropic formulations (Log-Energy, Permutation, Shannon, and Rényi). The predictive benchmark compared Linear Discriminant Analysis (LDA), Random Forest (RF), and the Multilayer Perceptron (MLP) network. The results mathematically demonstrated the nonlinearity of the feature space under noisy stress conditions, evidenced by the stagnation of LDA at 85.39% accuracy and high log-loss. Superior performance was achieved by the MLP, which attained 95.21% global accuracy and a Macro-F1 score of 0.95. The discriminative power analysis through the Fisher Score demonstrated that entropy metrics dominate class separability, preserving the topological signature of disturbance disorder even when amplitude is masked by noise. It is concluded that entropy-oriented Wavelet feature extraction, combined with classical nonlinear classifiers, delivers competitive generalization with the state of the art while dispensing with the high computational cost of deep architectures.

Keywords: Power Quality. Wavelet Packet Transform. Signal Entropy. Machine Learning. Artificial Neural Networks.

Lista de ilustrações

Figura 1	–	Diagrama de decomposição WPT em 4 níveis.	20
Figura 2	–	Diagrama de decomposição DWT em 4 níveis.	20
Figura 3	–	Diagrama conceitual de separação e projeção linear via LDA.	27
Figura 4	–	Fluxo de processamento, amostragem e votação na arquitetura Random Forest.	28
Figura 5	–	Fluxo conceitual de treinamento sequencial e minimização de resí- duos no XGBoost.	29
Figura 6	–	Representação simplificada da arquitetura de uma rede MLP.	30
Figura 7	–	Amostras limpas das 16 classes de distúrbios de QEE modeladas no domínio do tempo.	39
Figura 8	–	Espectro FFT do sinal sem ruído, destacando o componente fun- damental em 50 Hz.	40
Figura 9	–	Espectro FFT sobreposto de amostras estocásticas para as 16 clas- ses, ilustrando as componentes harmônicas e a variação do piso de ruído AWGN.	41
Figura 10	–	Matriz de Confusão Normalizada para a Rede Neural MLP.	44
Figura 11	–	Exemplos de sinais no domínio do tempo por classe de distúrbio, ilustrando a sobreposição estocástica de amplitude, temporização e ruído AWGN.	58
Figura 12	–	Projeções espaciais do vetor de características WPT utilizando PCA (linear) e t-SNE (não linear).	59
Figura 13	–	Top 30 características de maior poder discriminativo calculadas através do <i>Fisher Score</i>	60
Figura 14	–	Matriz de Confusão Normalizada para o classificador de Análise Discriminante Linear (LDA).	61
Figura 15	–	Matriz de Confusão Normalizada para o classificador Random Fo- rest (RF).	62

Lista de tabelas

Tabela 1	–	Classes de distúrbios de QEE modeladas conforme IEEE 1159-2019.	17
Tabela 2	–	Conjunto de atributos: 13 características extraídas por sub-banda da WPT.	23
Tabela 3	–	Configuração dos classificadores utilizados na fase experimental. . .	35
Tabela 4	–	Resultados do desempenho de classificadores — WPT (db8, L = 4).	42
Tabela 5	–	Taxa de revocação (<i>Recall</i>) por classe de distúrbio para a Rede Neural MLP.	46
Tabela 6	–	Síntese das causas físicas e mecanismos dominantes por classe. . . .	70

Sumário

1	Introdução	11
1.1	Contextualização	11
1.2	Problema e Motivação	11
1.3	Revisão Bibliográfica	12
1.3.1	Decomposição de Sinais	12
1.3.2	Extração de Atributos	13
1.3.3	Classificadores	13
1.3.4	Lacunas Identificadas	14
1.4	Objetivos	14
1.4.1	Objetivo Geral	14
1.4.1.1	Objetivos Específicos	15
1.4.2	Estrutura do Trabalho	15
2	Fundamentação Teórica	16
2.1	Qualidade de Energia Elétrica	16
2.2	Transformada Wavelet Packet (WPT)	18
2.3	Atributos Estatísticos e Entropias	22
2.4	Classificadores	26
2.5	Métricas de Avaliação	31
3	Metodologia	33
3.1	Geração dos Sinais Sintéticos	33
3.2	Decomposição WPT e Extração de Atributos	34
3.3	Preparação dos Dados	34
3.4	Classificação	35
3.5	Ambiente Computacional	36
4	Resultados e Discussão	38
4.1	Verificação dos Sinais Gerados	38
4.2	Benchmark de Classificadores	42
4.3	Análise da Matriz de Confusão	43
4.4	Recall por Classe	45
4.5	Discussão Geral e Comparação com a Literatura	47
5	Conclusão	50
	Referências	52

Apêndices	56
APÊNDICE A Figuras Complementares	57
APÊNDICE B Física Associada aos Distúrbios de	
Qualidade de Energia Elétrica	63
B.1 Distúrbios Simples	63
B.1.1 C01 — Normal (NOR)	63
B.1.2 C02 — Afundamento de Tensão (V-SAG)	64
B.1.3 C03 — Elevação de Tensão (V-SWL)	65
B.1.4 C04 — Interrupção (INT)	65
B.1.5 C05 — Harmônicos (HAR)	66
B.1.6 C06 — Flicker (FLK)	67
B.1.7 C07 — Transitório Oscilatório (OSC)	67
B.1.8 C08 — Transitório Impulsivo (IMP)	68
B.1.9 C09 — Notch Periódico (PN)	68
B.1.10 C10 — Spike (SPK)	69
B.2 Distúrbios Combinados	69
B.2.1 C11–C13: Sag/Swell/Interrupção + Harmônicos	70
B.2.2 C14: Flicker + Harmônicos	70
B.2.3 C15–C16: Flicker + Sag / Flicker + Swell	70
B.3 Síntese: O Mapa Físico do Problema	70

1 Introdução

1.1 Contextualização

A qualidade da energia elétrica (QEE) é um conceito que abrange o conjunto de características da tensão e da corrente fornecidas pelo sistema elétrico, avaliando sua adequação para alimentar cargas de forma segura e eficiente. A norma IEEE 1159-2019 (IEEE, 2019) define os principais fenômenos de QEE, incluindo variações de tensão de curta duração (afundamentos, elevações e interrupções), distorções harmônicas, flicker, transitórios e notching. No contexto brasileiro, o módulo 8 do PRODIST (ANEEL, 2021) estabelece os indicadores e limites de conformidade para a qualidade da energia fornecida pelas distribuidoras.

O cenário energético contemporâneo tem ampliado a relevância desse tema. A crescente integração de fontes renováveis interconectadas por inversores, a proliferação de cargas não-lineares (conversores eletrônicos, acionamentos de velocidade variável, iluminação LED) e o aumento da sensibilidade dos equipamentos de usuários finais têm tornado os sistemas elétricos mais suscetíveis a distúrbios de QEE (MAHELA; SHAIK; GUPTA, 2015). Equipamentos críticos em ambientes hospitalares, centros de processamento de dados e processos industriais contínuos podem sofrer falhas operacionais, degradação acelerada ou paradas não programadas quando expostos a distúrbios mesmo de curta duração (CAICEDO et al., 2023). O impacto econômico global dos problemas de qualidade de energia é estimado em bilhões de dólares anuais, abrangendo perdas de produção, danos a equipamentos e custos de manutenção corretiva.

1.2 Problema e Motivação

A identificação e classificação dos distúrbios de QEE é tradicionalmente realizada por engenheiros especialistas, a partir da análise visual de formas de onda capturadas por analisadores de qualidade de energia. Essa abordagem é limitada em escalabilidade: medidores modernos instalados em subestações e pontos de entrega geram volumes de dados que tornam a análise manual impraticável. Nesse contexto, a automatização do processo de classificação por meio de técnicas de processamento digital de sinais e aprendizado de máquina tornou-se uma necessidade operacional (MISHRA, 2019).

O desafio central é duplo. Primeiro, os sinais de QEE são inerentemente não-estacionários — distúrbios como transitórios e afundamentos envolvem variações abruptas

de amplitude e conteúdo espectral que invalidam a hipótese de estacionariedade da análise de Fourier clássica (VARANIS et al., 2021). Segundo, a presença de distúrbios combinados (por exemplo, afundamento com harmônicos ou flicker com elevação) cria sobreposições de características que dificultam a separação entre as classes (IGUAL; MEDRANO, 2020). A combinação desses fatores exige ferramentas de análise tempo-frequência capazes de capturar simultaneamente fenômenos de diferentes escalas temporais e espectrais.

1.3 Revisão Bibliográfica

A literatura em classificação de distúrbios de QEE é extensa e pode ser organizada em três eixos: decomposição de sinais, extração de atributos e classificação.

1.3.1 Decomposição de Sinais

Diversas técnicas de análise tempo-frequência têm sido aplicadas ao problema. A Transformada Wavelet Discreta (DWT) e a Transformada Wavelet Packet (WPT) são as mais amplamente empregadas, por oferecerem resolução variável no domínio tempo-frequência e decomposição determinística (MALLAT, 2009). A WPT, em particular, decompõe tanto os coeficientes de aproximação ("passa-baixo") quanto os de detalhe em cada nível, produzindo 2^L sub-bandas com resolução uniforme em frequência, o que é uma característica fundamental para sinais de QEE que contêm tanto componentes harmônicas (baixa frequência) quanto transitórias (alta frequência) (VARANIS; PEDERIVA, 2018).

A Transformada de Stockwell (S-Transform) oferece resolução variável sem o compromisso de janela fixa da STFT, e tem sido aplicada com sucesso por Mahela, Shaik e Gupta (2015) em conjunto com árvores de decisão e sistemas fuzzy. A Decomposição em Modos Empíricos (EMD) e sua variante Variacional (VMD) são métodos adaptativos, porém a EMD sofre com o problema de mistura de modos e a VMD requer a pré-definição do número de modos, limitando sua aplicação (MISHRA, 2019). A Transformada de Hilbert-Huang (HHT) combina EMD com análise de Hilbert, mas herda as limitações de determinismo da EMD (RODRIGUEZ et al., 2020).

Na comparação sistemática realizada por Madgula, Veeramsetty e Durgam (2025), a WPT apresenta o melhor compromisso entre resolução em frequência, custo computacional ($O(N \log N)$ versus $O(N^2)$ da S-Transform) e reprodutibilidade dos resultados, sendo particularmente eficaz quando combinada com famílias wavelet de suporte médio como db4–db8 (KHOKHAR et al., 2015).

1.3.2 Extração de Atributos

Os coeficientes da decomposição wavelet não são utilizados diretamente como entrada dos classificadores, mas sim caracterizados por atributos estatísticos que reduzem a dimensionalidade e aumentam a robustez a ruído. O conjunto clássico inclui energia por sub-banda, energia relativa, entropia de Shannon, desvio padrão e fator de crista (SINGH; PAL; DAHIYA, 2023; VARANIS; PEDERIVA, 2018). Rosso et al. (2001) introduziram a entropia wavelet como ferramenta de análise de sinais de curta duração, demonstrando que a distribuição de energia entre sub-bandas captura informação relevante sobre a complexidade do sinal.

Contribuições mais recentes sugerem que medidas de entropia complementares podem aumentar o poder discriminativo do conjunto de atributos. A entropia de Rényi de ordem $\alpha = 2$ (RÉNYI, 1961), também chamada *collision entropy*, é sensível à concentração de energia em poucos coeficientes, sendo particularmente discriminativa para transitórios impulsivos. A entropia de permutação, proposta por Bandt e Pompe (2002), quantifica a complexidade temporal da sequência de coeficientes independentemente de sua amplitude, capturando diferenças estruturais entre distúrbios com perfis energéticos similares. A entropia de Shannon clássica (SHANNON, 1948) complementa essas medidas ao capturar a incerteza média da distribuição de energia.

Além das entropias, estatísticas de ordem superior como curtose e assimetria (*skewness*) contribuem para a detecção de eventos impulsivos e assimetrias na distribuição dos coeficientes (ERISTI; YILDIRIM; ERISTI, 2020).

1.3.3 Classificadores

Os métodos de classificação aplicados a PQD abrangem desde técnicas lineares até arquiteturas profundas. A Análise Discriminante Linear (LDA) permanece como um baseline robusto: Singh, Pal e Dahiya (2023) demonstraram que LDA sobre atributos WPT atinge acurácias superiores a 99% em bases sintéticas com parâmetros fixos, evidenciando que a decomposição wavelet pode linearizar parcialmente o espaço de classes.

Métodos de ensemble baseados em árvores, como Random Forest (BREIMAN, 2001) e XGBoost (CHEN; GUESTRIN, 2016), têm demonstrado desempenho consistentemente superior em dados tabulares. Grinsztajn, Oyallon e Varoquaux (2022) demonstraram, em uma análise sistemática no NeurIPS 2022, que modelos baseados em gradient boosting superam redes neurais profundas na maioria dos problemas com dados tabulares de escala moderada, resultado com implicações diretas para pipelines de classificação de PQD baseados em atributos extraídos.

No domínio de aprendizado profundo, Wang, Xu e Che (2019) propuseram uma

CNN profunda para classificação de PQD diretamente a partir de sinais brutos, atingindo alta acurácia na base estudada. [Machlev et al. \(2021\)](#) disponibilizaram um gerador de dados sintéticos com classificadores de referência baseados em redes neurais, tornando-se referência para benchmarking na área. Mais recentemente, [Sipai, Jadeja e Tank \(2024\)](#) aplicaram transfer learning com redes pré-treinadas (GoogleNet, ResNet-18) sobre imagens geradas pela Transformada Wavelet Contínua (CWT), representando a fronteira atual dos métodos baseados em imagem.

No entanto, essas abordagens baseadas em imagens 2D (CWT/S-Transform \rightarrow CNN/ViT) têm custo computacional significativamente maior que pipelines tabulares (WPT \rightarrow atributos \rightarrow classificador), sem ganho proporcional de acurácia em muitos cenários ([OLIVEIRA; BOLLEN, 2024](#)). O FT-Transformer, proposto por [Gorishniy et al. \(2021\)](#) como arquitetura de atenção para dados tabulares, representa uma alternativa que opera sobre atributos pré-calculados com capacidade de modelar relações inter-atributos via mecanismo de self-attention. Até o momento, sua aplicação à classificação de PQD não foi reportada na literatura consultada.

1.3.4 Lacunas Identificadas

A revisão da literatura permite identificar três lacunas relevantes. Primeiro, a maioria dos trabalhos utiliza bases sintéticas com parâmetros fixos, gerando instâncias idênticas de cada classe com variação apenas no nível de ruído, o que inflaciona artificialmente a acurácia dos classificadores ([MACHLEV et al., 2021](#)). Segundo, a comparação entre medidas de entropia complementares (Shannon, Rényi, permutação) como atributos discriminativos para PQD não foi realizada de forma sistemática. Terceiro, a avaliação de calibração dos classificadores, via métricas como log-loss (entropia cruzada no conjunto de teste), é praticamente ausente na literatura de QEE, que se limita a reportar acurácia e, ocasionalmente, F1-score.

1.4 Objetivos

1.4.1 Objetivo Geral

Desenvolver e avaliar um pipeline computacional para classificação automática de 16 tipos de distúrbios de qualidade de energia elétrica, utilizando decomposição Wavelet Packet e classificadores baseados em aprendizado de máquina, com ênfase na análise do papel de entropias complementares como atributos discriminativos.

1.4.1.1 Objetivos Específicos

1. Implementar um gerador de sinais sintéticos de distúrbios de QEE conforme a norma IEEE 1159-2019 (IEEE, 2019), com variação paramétrica estocástica que garanta diversidade intra-classe.
2. Implementar a decomposição Wavelet Packet Transform (WPT) com wavelet db8 e nível $L = 4$, extraindo 13 atributos estatísticos por sub-banda, incluindo entropias de Shannon, Rényi ($\alpha = 2$) e permutação.
3. Comparar sistematicamente o desempenho de três classificadores — Análise Discriminante Linear (LDA), Random Forest e Perceptron Multicamadas (MLP), utilizando métricas de acurácia, macro-F1, log-loss e recall por classe.
4. Analisar os padrões de confusão entre classes e interpretá-los à luz das características físicas de cada tipo de distúrbio.

1.4.2 Estrutura do Trabalho

Este trabalho está organizado da seguinte forma: o Capítulo 2 apresenta a fundamentação teórica, cobrindo os conceitos de qualidade de energia elétrica, a transformada wavelet packet, os atributos estatísticos e entropias utilizadas, os classificadores empregados e as métricas de avaliação. O Capítulo 3 descreve a metodologia, incluindo a geração dos sinais sintéticos, a extração de atributos, a preparação dos dados e as configurações de cada classificador. O Capítulo 4 apresenta e discute os resultados experimentais. O Capítulo 5 conclui o trabalho e aponta direções futuras.

2 Fundamentação Teórica

2.1 Qualidade de Energia Elétrica

O conceito de QEE abrange uma ampla variedade de fenômenos eletromagnéticos que caracterizam a tensão e corrente num determinado instante e local na rede elétrica. De acordo com a norma IEEE 1159-2019 (IEEE, 2019) e o módulo 8 do PRODIST (ANEEL, 2021), a QEE é avaliada pela conformidade do sinal elétrico com parâmetros ideais de amplitude, frequência e forma de onda. Quaisquer desvios significativos desses parâmetros constituem um distúrbio de QEE (PQD), que pode acarretar falhas de equipamentos, interrupções de processos industriais e perdas econômicas significativas, especialmente com o avanço de construções de *datacenters*, avanço da automação industrial e equipamento mais moderno (MCGRANAGHAN et al., 2012).

Tradicionalmente, os distúrbios são classificados em categorias baseadas em seu conteúdo espectral e duração temporal. As anomalias mais recorrentes incluem variações de curta duração, como afundamentos (*sags*), elevações (*swells*), e interrupções; distorções contínuas na forma da onda, como harmônicos e *notching*, flutuações periódicas de amplitude (*flicker*) e eventos de altíssima frequência (transientes oscilatórios e impulsivos, *spikes*) (IEEE, 2019; MCGRANAGHAN et al., 2012).

Com o advento das redes elétricas inteligentes (*smart grids*) e a expansão de monitoramento contínuo, a análise de QEE se tornou um desafio de grandes volumes de dados (*Big Data*). Um único analisador de rede operando com altas taxas de amostragem pode gerar gigabytes de registros transientes em poucos dias, e milhões de amostras em poucas horas (OLIVEIRA; BOLLEN, 2024). Esse volume informacional inviabiliza a inspeção e o diagnóstico manual por operadores humanos, estabelecendo a necessidade rigorosa de sistemas automatizados de detecção e classificação baseados em processamento de sinais.

Neste trabalho, foram definidas 16 classes de PQD, a fim de representar os fenômenos mais comuns no cenário de QEE. A Tabela 1 apresenta as 16 classes com suas diferentes equações matemáticas e parâmetros, além de identificar cada classe com um ID único.

Tabela 1 – Classes de distúrbios de QEE modeladas conforme IEEE 1159-2019.

ID	Sigla	Nome	Modelo	Parâmetros
C01	NOR	Normal	$v(t) = A \sin(\omega t + \varphi)$	—
C02	V-SAG	Afundamento	$v(t) = A[1 - \alpha(u_{t_1} - u_{t_2})] \sin(\omega t + \varphi)$	$\alpha \in [0,1; 0,9]$
C03	V-SWL	Elevação	$v(t) = A[1 + \alpha(u_{t_1} - u_{t_2})] \sin(\omega t + \varphi)$	$\alpha \in [0,1; 0,8]$
C04	INT	Interrupção	$v(t) = A[1 - \alpha(u_{t_1} - u_{t_2})] \sin(\omega t + \varphi)$	$\alpha \in [0,9; 1,0]$
C05	HAR	Harmônicos	$v(t) = A[\sin(\omega t) + \sum \alpha_n \sin(n\omega t)]$	$\alpha_n \in [0,05; 0,15]$
C06	FLK	Flicker	$v(t) = A[1 + \alpha_f \sin(2\pi f_f t)] \sin(\omega t + \varphi)$	$\alpha_f \in [0,1; 0,2]$
C07	OSC	Trans. oscilatório	Fundamental + transitório amortecido	$\beta \in [0,1; 0,8]$
C08	IMP	Trans. impulsivo	Fundamental + impulso exponencial	$\beta \in [0,1; 0,4]$
C09	PN	Notch periódico	Fundamental – notch nos cruzamentos	$K \in [0,1; 0,4]$
C10	SPK	Spike	Fundamental + spike nos picos	$K \in [0,1; 0,4]$
C11	SAG-H	Sag + harmônicos	Envelope sag \times (fund. + harm.)	C02 + C05
C12	SWL-H	Swell + harmônicos	Envelope swell \times (fund. + harm.)	C03 + C05
C13	INT-H	Interrup. + harm.	Envelope interrup. \times (fund. + harm.)	C04 + C05
C14	FLK-H	Flicker + harm.	Modulação flicker \times (fund. + harm.)	C06 + C05
C15	FLK-SAG	Flicker + sag	Mod. flicker \times env. sag \times fund.	C06 + C02
C16	FLK-SWL	Flicker + swell	Mod. flicker \times env. swell \times fund.	C06 + C03

A equação 2.1 da classe C01 representa o sinal ideal:

$$v(t) = A \sin(\omega t + \varphi) \quad (2.1)$$

Onde A é a amplitude, em p.u, ω é a frequência angular em rad/s, t é o tempo em segundos e φ é o ângulo de fase. A definição das 16 classes operacionais visa cobrir tanto os fenômenos isolados quanto as ocorrências simultâneas na rede elétrica. A inclusão das classes combinadas (C11 a C16) é necessária devido à natureza caótica dos sistemas de potência reais, onde distúrbios de estados estacionários (como harmônicos e *flickers*) frequentemente coexistem com falhas de curta duração (*sags*, *swells* e interrupções). Um sistema de classificação treinado puramente com sinais isolados apresenta severas restrições de acurácia e a limitação inerente a esses métodos torna-se evidente em redes de distribuição reais, onde a incidência simultânea de anomalias é frequente e referenciada como distúrbios múltiplos (FERREIRA; SEIXAS; CERQUEIRA, 2015). O apêndice [B] possui todas as 16 classes explicadas em maiores detalhes, com descrição das falhas físicas que podem causar aquele distúrbio.

Contudo, a extração de atributos capazes de discriminar eficientemente cada uma dessas 16 classes impõe um desafio metodológico complexo. Abordagens de análise restritas ao domínio do tempo são insuficientes para capturar as componentes harmônicas, enquanto ferramentas clássicas baseadas na Transformada de Fourier falham na localização temporal de eventos não-estacionários de curta duração (KHOKHAR et al., 2015).

2.2 Transformada Wavelet Packet (WPT)

Como demonstrado pela Eq.2.1, um sinal de tensão ideal em regime estacionário pode ser modelado no domínio do tempo de forma analítica. Embora a representação temporal seja intuitiva para o monitoramento contínuo, ela apresenta limitações severas na detecção e caracterização de distúrbios de QEE. No domínio do tempo, a identificação e o isolamento de transientes de alta frequência tornam-se ineficientes, pois a dinâmica desses distúrbios é dominada pela oscilação contínua da frequência fundamental. Para superar essa limitação, a abordagem clássica de processamentos de sinais recorre à transformada de Fourier (FT)(MISHRA, 2019). A FT $X(f)$ de um dado sinal estacionário $x(t)$ é dada por (ROBERTSON et al., 1996):

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (2.2)$$

E a função $X(f)$ é a representação no domínio da frequência de $x(t)$. O uso da análise de Fourier encontra um problema ao lidar com sinais não-estacionários. A FT identifica com precisão as componentes de frequência presentes no sinal, mas perde completamente a resolução referente ao instante de tempo em que o distúrbio ocorreu, uma vez que é uma representação global no domínio da frequência(DAUBECHIES, 1992; SANTOSO et al., 1996).

Uma vez que PQDs são eventos dinâmicos, localizados temporalmente e não-estacionários, tornou-se necessária a adoção da transformada de Fourier de Tempo Curto (*Short Time Fourier Transform*, STFT), para que os diagnósticos de falhas fossem eficientes para determinar as causas do distúrbio. A formulação da STFT é dada por(ROBERTSON et al., 1996):

$$STFT(f, \tau) = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j2\pi ft} dt \quad (2.3)$$

Onde τ é a translação temporal da posição da função janela (*window function*) $w(t)$. Essa translação torna possível o mapeamento do espectro do sinal em intervalos localizados. Porém, a STFT esbarra em uma limitação física intrínseca formalizada pelo Princípio de Incerteza de Heisenberg(MALLAT, 2009). O comprimento da janela temporal é fixo, portanto, uma janela estreita garante uma alta resolução no tempo, porém deteriora a resolução na frequência; inversamente, uma janela ampla otimiza a discriminação frequencial ao custo do "borramento" (*smearing*) do instante de ocorrência do evento. Não é possível obter uma resolução ótima simultaneamente em ambos os domínios, portanto a STFT é insuficiente para a análise de sinais em regimes não-estacionários(KHOKHAR et al., 2015). A Transformada Wavelet supera esse limite empregando funções base curtas e oscilatórias, denominadas *wavelets*, caracterizadas pela seguinte propriedade:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.4)$$

Essa condição fundamental estabelece que a *wavelet-mãe* $\psi(t)$ possui média nula no domínio do tempo. Geometricamente, isso assegura que a função seja estritamente oscilatória — uma "pequena onda" com suporte compacto cujas áreas positivas e negativas se cancelam perfeitamente. Matematicamente, essa propriedade é uma consequência direta da condição de admissibilidade, um critério indispensável que viabiliza a reconstrução exata do sinal original sem perda de informação. Ao garantir que a integral temporal seja nula, a representação da *wavelet* no domínio da frequência satisfaz $\Psi(0) = 0$ (MALLAT, 2009). A transformada wavelet contínua, é dada por:

$$CWT(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.5)$$

Diferente da STFT, que utiliza um passo fixo de translação no tempo, a CWT mapeia o sinal bidimensionalmente através do parâmetro de escala a (que dilata ou contrai a função, agindo de forma inversamente proporcional à frequência) e do parâmetro de translação b (que desloca a wavelet no domínio temporal para localizar os transientes). O fator $1/\sqrt{|a|}$ garante a conservação da energia da função base em diferentes escalas.

Embora a CWT ofereça um mapeamento analítico contínuo e minucioso, sua implementação direta em sistemas digitais é computacionalmente inviável. A variação contínua dos parâmetros a e b gera um volume infinito de dados e uma representação altamente redundante do plano tempo-frequência. Para solucionar essa limitação, os parâmetros de escala e translação são discretizados. O marco teórico que viabilizou essa discretização de forma exata e computacionalmente eficiente foi a Análise de Multirresolução (MRA), desenvolvida por Mallat (1989). Mallat demonstrou matematicamente que a aplicação de wavelets com escalas diádicas comporta-se de maneira estritamente equivalente ao processamento do sinal através de um banco de filtros digitais espelho em quadratura (QMF - *Quadrature Mirror Filters*).

Neste formalismo, a extração de características não requer o cálculo de integrais complexas. O sinal discreto $x[n]$ é processado iterativamente por dois filtros complementares: um filtro passa-baixa $h[n]$, associado a uma função de escala $\phi(t)$ que captura o comportamento estacionário do sinal, e um filtro passa-alta $g[n]$, associado à wavelet-mãe $\psi(t)$ que reage a transientes e descontinuidades. As aproximações (A_j), que representam os componentes de baixa frequência, e os detalhes (D_j), que representam os componentes de alta frequência, são representados pelas equações Eq.2.6 e Eq.2.7:

$$A_j[k] = \sum_n x[n] h[2k - n] \quad (2.6)$$

$$D_j[k] = \sum_n x[n]g[2k - n] \tag{2.7}$$

Onde k é o fator de translação da *wavelet* mãe, representando a localização temporal do filtro, e h e g são os filtros de passa-baixa e passa-alta, respectivamente.

Essa arquitetura em cascata introduzida por Mallat elimina a redundância da integral contínua e processa o sinal em tempo logarítmico, fornecendo a fundação matemática sobre a qual a Transformada Wavelet Packet (WPT), adotada neste trabalho para decompor simultaneamente ambos os ramos de filtro, é construída. As Figuras 1 e 2 mostram diagramas da decomposição de um determinado sinal $x(t)$, de acordo com a decomposição WPT e com a decomposição DWT:

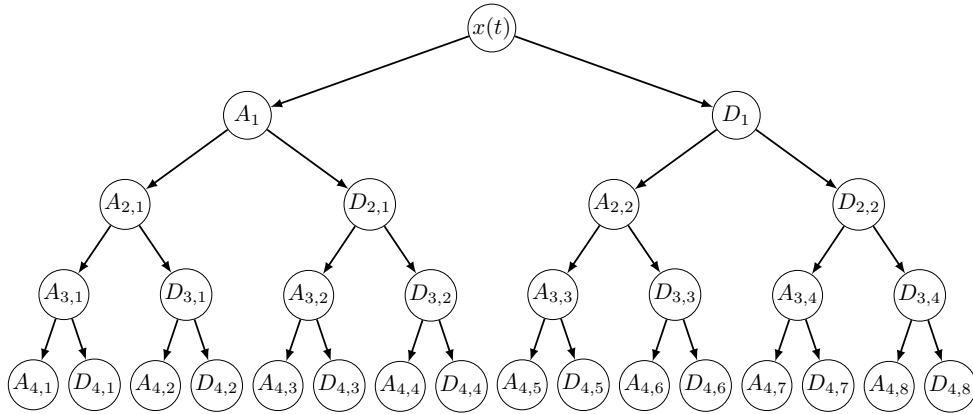


Figura 1 – Diagrama de decomposição WPT em 4 níveis.

Fonte: Autoria própria.

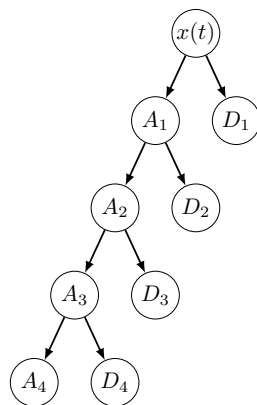


Figura 2 – Diagrama de decomposição DWT em 4 níveis.

Fonte: Autoria própria.

Para determinar quais os filtros que serão utilizados, é preciso escolher uma *wavelet*-mãe $\psi(t)$. Esta escolha dita a eficácia da WPT. Na literatura, existem muitos tipos de famílias de *wavelet*-mãe (GUO et al., 2022), tornando a escolha aparentemente impossível. Entre as famílias, existem as Daubechies, nomeadas em homenagem à sua grande

descobridora Daubechies (1992). As *wavelets* da família Daubechies seguem a convenção de nomenclatura dbN , onde $N \in \mathbb{N}$. Outra família, são as *symlets*, denotadas por $symN$. O número N denota a ordem ou "suporte" da *wavelet*, que é, fisicamente, a sua "largura" ou duração no domínio do tempo. Um suporte compacto (pouca duração que a *wavelet* "vive" no tempo, ou seja, é diferente de 0) permite que o filtro responda instantaneamente a transientes e garante excelente resolução temporal. Em contrapartida, suportes mais longos suavizam o sinal e aumentam a capacidade matemática de discriminar frequências próximas uma das outras, diminuindo a ocorrência de vazamento entre as sub-bandas de frequência.

Para este trabalho, foi escolhida a *wavelet* da família Daubechies com ordem 8 (db8). A escolha da *wavelet*-mãe Daubechies 8 (db8) justifica-se pelo compromisso estabelecido entre a resolução temporal e a resolução frequencial. Filtros de resposta ao impulso finita (FIR) com baixo suporte, como a *wavelet* Haar, apresentam uma resposta em frequência distante da ideal, ocasionando vazamento de energia (aliasing) entre as sub-bandas de decomposição (VARANIS; PEDERIVA, 2018). Por outro lado, filtros com suporte excessivamente elevado elevam o custo computacional e causam atenuação temporal de transientes rápidos. Desse modo, a db8 oferece uma resposta em frequência com seletividade adequada para mitigar o vazamento espectral entre as 16 bandas da WPT, preservando simultaneamente a localização temporal de distúrbios de curta duração, tais como afundamentos transitórios e transientes impulsivos.

Por fim, o nível de detalhe dessa análise de PQD é determinada pelo nível de decomposição (L) da árvore WPT. O nível define a profundidade da filtragem e, conseqüentemente, a resolução em frequência. Para este estudo, a frequência fundamental da rede elétrica é de 50 Hz e a taxa de aquisição é de 100 amostras por ciclo, resultando em uma frequência de amostragem $F_s = 5000$ Hz. Pelo Teorema da Amostragem de Nyquist-Shannon (SHANNON, 1949), a máxima frequência útil que pode ser analisada sem ocorrência de distorções por *aliasing* corresponde à metade de F_s (ou seja, 2500 Hz).

Ao adotar $L = 4$, o sinal é processado através de quatro estágios sucessivos, dividindo o espectro útil disponível (0 a 2500 Hz) em $2^4 = 16$ sub-bandas de energia independentes. Do ponto de vista prático e interpretativo, essa decomposição divide o sinal em 16 sub-bandas com largura de banda equivalente a $F_s / (2^{L+1}) = 5000 / 32 = 156,25$ Hz. Essa configuração provou-se ideal para o estudo, pois segmenta o sinal de forma fina o suficiente para isolar as assinaturas de frequência de cada distúrbio em sua respectiva sub-banda, sem gerar uma complexidade computacional desnecessária para a etapa posterior de classificação.

2.3 Atributos Estatísticos e Entropias

A aplicação da Transformada Wavelet Packet (WPT), detalhada na seção anterior, resulta em um conjunto de coeficientes discretos distribuídos ao longo de 16 sub-bandas de frequência. Embora esses coeficientes brutos carreguem toda a informação dinâmica do sinal original, sua utilização direta como entrada para algoritmos de classificação inteligente apresenta severas limitações práticas.

A primeira limitação reside na alta dimensionalidade dos dados. Um sinal amostrado gera milhares de coeficientes por janela temporal. Esse cria um espaço amostral massivo, o que eleva drasticamente o custo computacional e induz os classificadores ao problema da "maldição da dimensionalidade" (SINGH; PAL; DAHIYA, 2023). A segunda limitação é a sensibilidade ao ruído. Sinais elétricos reais de redes de distribuição estão inerentemente corrompidos por ruídos de fundo de alta frequência e pequenas flutuações estocásticas. Os coeficientes brutos refletem diretamente essas perturbações pontuais, o que pode mascarar os padrões reais das PQD (OLIVEIRA; BOLLEN, 2024).

Para contornar esses problemas, recorre-se ao processo de extração de atributos (*feature extraction*). Esse procedimento atua como uma técnica de redução de dimensionalidade e filtragem estatística, condensando a série temporal de coeficientes de cada sub-banda em métricas escalares compactas que retêm o comportamento físico subjacente do distúrbio. Atributos estatísticos de primeira e segunda ordem são capazes de mapear a dispersão, a assimetria e a energia do sinal de forma compacta, tornando possível a classificação dos diferentes eventos PQD (MISHRA, 2019).

Além das métricas estatísticas tradicionais, o diagnóstico preciso de fenômenos dinâmicos e não-lineares exige indicadores baseados na Teoria da Informação. Nesse cenário, as medidas de entropia destacam-se pela capacidade de quantificar a complexidade, a irregularidade e as mudanças de comportamento na distribuição de energia e na dinâmica temporal das sub-bandas da WPT. Seguindo a abordagem consolidada na literatura para assinaturas energético-entrópicas (VARANIS; PEDERIVA, 2018), este trabalho adota quatro formulações complementares de entropia: Shannon, Log-Energy, Rényi e Permutação, cujas definições matemáticas e interpretações físicas desempenham um papel crítico na discriminação e na estabilidade numérica das assinaturas de falha. A Tabela 2 apresenta os 13 atributos adotados para este trabalho, com suas diferentes equações e significado físico.

Tabela 2 – Conjunto de atributos: 13 características extraídas por sub-banda da WPT.

#	Atributo	Equação	Informação Capturada
<i>Energia e Entropia</i>			
01	Energia	$\sum c_i^2$	Energia total do sinal na sub-banda
02	Energia Relativa	$\sum c_i^2/N$	Energia média por amostra
03	Entropia de Shannon	$-\sum p_i \log p_i$	Incerteza média da distribuição de energia
04	Entropia Log-Energy	$\sum \log(c_i^2)$	Formulação de (ROSSO et al., 2001); estabilidade numérica
05	Entropia de Rényi ($\alpha = 2$)	$-\log_2(\sum p_i^2)$	Concentração de energia; sensível a picos (transitórios)
06	Entropia de Permutação	Dist. padrão ordinal	Complexidade temporal; independente da amplitude
<i>Estatísticas</i>			
07	Desvio Padrão	$\sigma(c)$	Espalhamento dos valores dos coeficientes
08	RMS	$\sqrt{\text{média}(c^2)}$	Amplitude eficaz (Root Mean Square)
09	Máximo Absoluto	$\max(c)$	Valor de pico do coeficiente
10	Fator de Crista	$\max(c)/\text{RMS}$	Indicador de impulsividade do sinal
11	Média	\bar{c}	Deslocamento DC (nível contínuo) na sub-banda
<i>Estatísticas de Ordem Superior (HOS)</i>			
12	Curtose	$E[(c - \mu)^4]/\sigma^4 - 3$	Peso das caudas; alta para eventos impulsivos
13	Assimetria	$E[(c - \mu)^3]/\sigma^3$	Grau de assimetria da distribuição

Conforme sintetizado na Tabela 2, o vetor de características consolidado neste estudo engloba três domínios analíticos distintos: métricas de energia pura, indicadores estatísticos espaciais (incluindo primeira ordem, segunda ordem e *Higher Order Statistics*, também chamada de HOS) e medidas entrópicas de complexidade. Essa diversidade é projetada intencionalmente para garantir robustez analítica.

Ao aplicar a transformada Wavelet Packet (WPT), a informação armazenada dentro de cada sub-banda de frequência é representada por um vetor de coeficientes de amplitude c_i . A energia do sinal, dada pela formulação clássica (OPPENHEIM; SCHAFER, 2010), é calculada por:

$$E_{total} = \sum_{n=1}^N |x[n]|^2 \quad (2.8)$$

Como a WPT projeta o sinal em uma base ortonormal, o Teorema de Parseval garante a conservação da energia total entre os domínios. Isso significa que a energia total do sinal é rigorosamente igual à soma das energias de todos os coeficientes gerados pela transformada (MALLAT, 2009).

Aproveitando-se desta propriedade, a energia concentrada em uma única sub-banda de frequência específica (E_{sub}) pode ser isolada e calculada diretamente pela soma quadrática apenas de seus respectivos coeficientes:

$$E_{sub} = \sum_{i=1}^K c_i^2 \quad (2.9)$$

onde K representa o número total de coeficientes contidos na respectiva sub-banda.

Para evitar distorções causadas por variações no comprimento do vetor de coeficientes, utiliza-se também a Energia Relativa (E_{rel}), que fornece a densidade média de energia por coeficiente na sub-banda:

$$E_{rel} = \frac{1}{K} \sum_{i=1}^K c_i^2 \quad (2.10)$$

Fisicamente, a Energia Relativa atua como um indicativo da densidade de energia alocada na respectiva faixa de frequência. Ao dividir a energia absoluta pelo número de coeficientes da sub-banda, a métrica torna-se invariante a flutuações no comprimento da janela temporal de amostragem. No contexto do diagnóstico de Qualidade de Energia, essa normalização é crucial, pois fornece aos algoritmos de aprendizado de máquina uma avaliação padronizada da "severidade" do distúrbio, independentemente da duração absoluta do evento analisado.

Contudo, embora métricas escalares como E_{sub} e E_{rel} sejam eficientes para detectar a magnitude de uma anomalia, como o aumento de amplitude inerente a um fenômeno de elevação de tensão (*Swell*), elas tratam a sub-banda como uma entidade uniforme. Tais atributos são incapazes de descrever a distribuição probabilística e a complexidade estrutural dessa energia. Um transitório impulsivo curtíssimo e um afundamento de tensão prolongado podem, teoricamente, depositar a mesma quantidade de energia em uma banda, tornando-os indistinguíveis sob essa ótica puramente energética.

É para suprir esta exata lacuna analítica que as métricas baseadas na Teoria da Informação são integradas ao vetor de características. A Entropia de Shannon (SHANNON, 1948), por exemplo atua como um mapeamento da incerteza global, quantificando quão uniformemente a energia do sinal está distribuída ao longo dos coeficientes da sub-banda. A Entropia de Shannon pode ser calculada através de:

$$H_{Shannon} = - \sum_{i=1}^K p_i \log_2 p_i \quad (2.11)$$

Onde, no contexto de energia do sinal por sub-banda, a probabilidade p_i é igual a distribuição de probabilidade normalizada, ou seja, a "quantidade de energia relativa" por coeficiente:

$$p_i = \frac{c_i^2}{\sum c_j^2} \quad (2.12)$$

Contudo, por sua natureza logarítmica, ela apresenta limitações para isolar variações abruptas e de curtíssima duração. Para suprir essa lacuna, a Entropia de Rényi com ordem $\alpha = 2$ é introduzida: ao indexar probabilisticamente os estados de maior energia, essa métrica exibe alta sensibilidade a picos de amplitude e descontinuidades localizadas (RÉNYI, 1961), tornando-se o indicador ideal para discriminar fenômenos transitórios e impulsivos, como *spikes* e *notches*. A equação da Entropia de Rényi de ordem $\alpha = 2$ (também conhecida na literatura como *Collision Entropy*) é definida por:

$$H_{Rényi} = -\log_2 \left(\sum_{i=1}^K p_i^2 \right) \quad (2.13)$$

Ao elevar as probabilidades ao quadrado antes do somatório logarítmico, a métrica penaliza coeficientes de baixa energia e amplifica a dominância dos picos, atuando como um sensível detector de transientes de alta frequência.

Em complemento às métricas baseadas na distribuição probabilística, a Entropia Log-Energy (ROSSO et al., 2001) é incorporada para lidar com a vasta gama dinâmica dos coeficientes wavelet. Em vez de utilizar as probabilidades p_i , sua formulação atua diretamente sobre o valor absoluto da energia dos coeficientes:

$$H_{log} = \sum_{i=1}^K \log(c_i^2) \quad (2.14)$$

Fisicamente, a operação logarítmica amplifica a contribuição dos coeficientes de baixa amplitude e comprime grandes picos. Essa característica a torna um indicador robusto para mensurar o piso de energia da sub-banda e garantir a estabilidade do sinal.

Por fim, a Entropia de Permutação (BANDT; POMPE, 2002) introduz uma quebra de paradigma na extração de atributos, deslocando o foco da magnitude para a dinâmica estrutural. Diferente de Shannon e Rényi, esta métrica não mensura energia, mas sim a complexidade da ordenação temporal da série de coeficientes. Avaliando padrões ordinais de comprimento m , a entropia é calculada como:

$$H_{perm} = -\frac{1}{\log_2(m!)} \sum_{\pi} p(\pi) \log_2 p(\pi) \quad (2.15)$$

onde π representa os possíveis padrões ordinais e $p(\pi)$ a sua respectiva frequência relativa. O termo $\log_2(m!)$ atua como um fator de normalização. Por ser estritamente independente das amplitudes do sinal, a Entropia de Permutação é a métrica fundamental para distinguir fenômenos que coabitam a mesma faixa espectral. É esta propriedade que permite ao modelo de classificação isolar, por exemplo, a componente fundamental (50 Hz) de seu terceiro harmônico (150 Hz) inseridos na mesma sub-banda da WPT (0 a 156,25 Hz), uma vez que a taxa de oscilação temporal intrínseca e, conseqüentemente, a

entropia de permutação do sinal de 150 Hz é intrinsecamente superior, independente das equivalências em suas energias totais.

Além desses atributos de Energia e Entropia, também é necessário utilizar medidas estatísticas "tradicionais" para a classificação dos sinais, uma vez que, com ruído introduzido ao sinal, é possível que alguns distúrbios possuam entropia similar ao sinal "ideal", ou até mesmo, possuam a mesma entropia mas representem distúrbios completamente diferentes (distúrbios com naturezas físicas completamente opostas, como um afundamento (*sag*) e uma elevação (*swell*) de tensão que possuam perfis temporais simétricos).

Para solucionar esta ambiguidade, atributos como o Valor Eficaz (RMS), a Média e o Máximo Absoluto se fazem essenciais, pois atuam como âncoras de magnitude, capturando o deslocamento macroscópico da amplitude do sinal (PAPOULIS, 1991). Adicionalmente, em cenários reais onde o sinal é corrompido por ruído estocástico, métricas puramente energéticas podem se aproximar do sinal nominal ("ideal"), exigindo indicadores como o Desvio Padrão e o Fator de Crista para rastrear a verdadeira dispersão e impulsividade contidas na sub-banda (PAPOULIS, 1991).

Por fim, para mapear o comportamento estocástico do sinal além de sua dispersão básica, o vetor de características é complementado com Estatísticas de Ordem Superior (HOS - *Higher-Order Statistics*), representadas pela Assimetria (*Skewness*) e pela Curtose (*Kurtosis*) (MENDEL, 1991). A Assimetria quantifica o grau de distorção ou assimetria da distribuição de probabilidade dos coeficientes em relação à sua média, sendo útil para identificar distúrbios que afetam a onda de forma desigual. Já a Curtose avalia o "peso" das caudas da distribuição. Sub-bandas com alta ocorrência de eventos transitórios abruptos e impulsivos, típicos de descargas ou falhas de comutação, geram um excesso de curtose acentuado, permitindo que os classificadores isolem anomalias de alta frequência que dificilmente alterariam o valor RMS da janela. Deste modo, o conjunto de 13 atributos extraídos estabelece uma representação comprimida, fisicamente embasada e matematicamente robusta para a etapa subsequente de classificação.

2.4 Classificadores

A classificação automatizada de distúrbios de QEE a partir do vetor de características x configura um problema de aprendizado supervisionado multi-classe. O objetivo consiste em mapear o vetor x a um rótulo de classe $y \in C_{01}, C_{02}, \dots, C_{16}$, representando as 16 classes de distúrbios. Para garantir a robustez da análise e avaliar diferentes superfícies de decisão, este trabalho adota três paradigmas distintos de classificação: discriminantes lineares, métodos de comitê baseados em árvores e redes neurais artificiais, além de referenciar o estado da arte para dados tabulares.

A Análise Discriminante Linear (LDA) é um método clássico de classificação supervisionada fundamentado na hipótese de que as observações pertencentes a cada classe seguem uma distribuição Gaussiana multivariada e compartilham uma mesma matriz de covariância (FISHER, 1936). O método busca determinar combinações lineares das variáveis de entrada que maximizem a separação entre as diferentes classes, ao mesmo tempo em que minimizam a dispersão dos dados dentro de cada classe. A partir dessas projeções, são construídas funções discriminantes lineares utilizadas para a atribuição de novos exemplos às classes mais prováveis. Por possuir solução analítica de forma fechada e não exigir processos iterativos extensivos de treinamento nem ajuste significativo de hiperparâmetros, o LDA é amplamente empregado como modelo de referência (*baseline*) para avaliar a separabilidade linear do problema e servir de comparação para métodos de maior complexidade.

A Figura 3 ilustra o funcionamento do algoritmo para um caso bidimensional envolvendo duas classes, no qual a discriminação ocorre por meio da projeção dos dados em um único eixo discriminante. Para problemas multiclasse, como as 16 categorias de distúrbios consideradas neste trabalho, o LDA generaliza essa formulação ao projetar os dados em um subespaço discriminante de dimensão máxima $C - 1$, onde C representa o número de classes. Nesse espaço, são preservadas as direções que maximizam a separação entre classes em relação à dispersão intraclasse. Como consequência da hipótese de matriz de covariância compartilhada entre as classes, as funções discriminantes obtidas permanecem lineares, resultando em fronteiras de decisão definidas por hiperplanos capazes de separar simultaneamente as diferentes categorias de eventos.

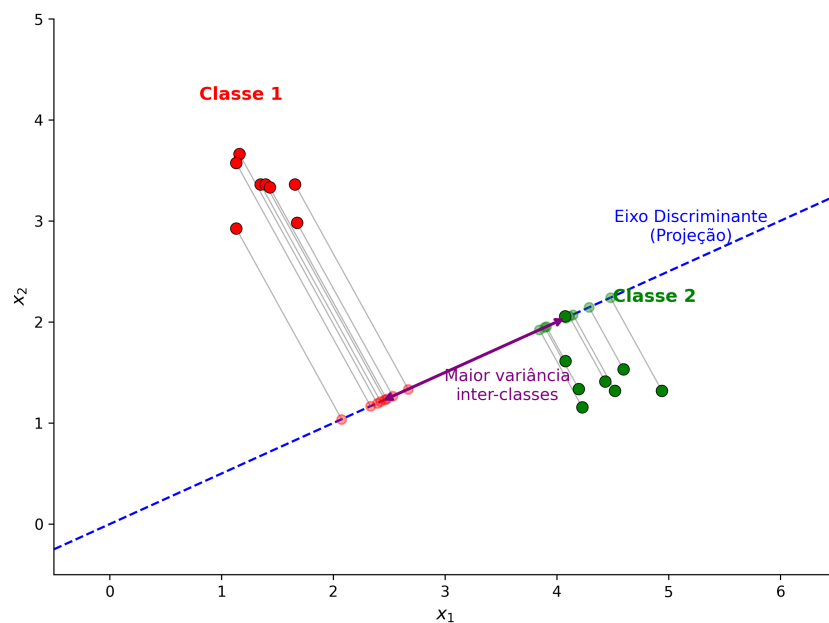


Figura 3 – Diagrama conceitual de separação e projeção linear via LDA.

Fonte: Autoria própria.

O algoritmo Random Forest (RF) é um método de aprendizado supervisionado baseado no paradigma de *ensemble learning*. Nesse paradigma, múltiplos modelos são combinados com o objetivo de produzir um classificador mais robusto e preciso do que qualquer modelo individual. Formalizado por Breiman (2001), o Random Forest constrói uma floresta composta por diversas árvores de decisão, explorando o fato de que a agregação de classificadores pouco correlacionados reduz a variância do modelo e melhora sua capacidade de generalização.

A baixa correlação entre as árvores é promovida por duas fontes de aleatoriedade. A primeira consiste na técnica de *bagging* (*Bootstrap Aggregating*), na qual cada árvore é treinada a partir de uma amostra aleatória obtida com reposição do conjunto de treinamento original. A segunda ocorre durante a construção da árvore, quando apenas um subconjunto aleatório dos atributos é considerado em cada divisão (*split*) do nó.

Como consequência, as árvores tornam-se menos correlacionadas entre si, reduzindo a tendência ao sobreajuste observada em árvores individuais (BREIMAN, 2001; HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Na etapa de inferência, o vetor de características extraído da WPT é submetido a todas as árvores da floresta, sendo a classe final determinada por votação majoritária (*majority voting*). Essa estratégia confere elevada robustez a ruídos e variações dos sinais, característica particularmente desejável na classificação de distúrbios de Qualidade de Energia. O fluxo geral do algoritmo é ilustrado na Figura 4.

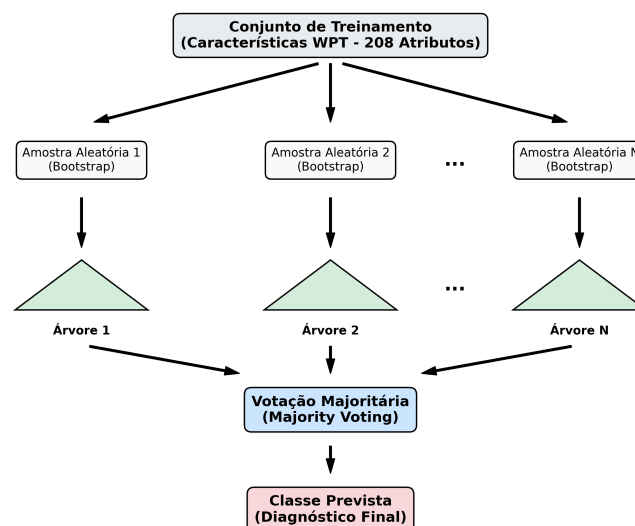


Figura 4 – Fluxo de processamento, amostragem e votação na arquitetura Random Forest.

Fonte: Autoria própria.

O Extreme Gradient Boosting (XGBoost) fundamenta-se no princípio de *gradient boosting*, construindo um *ensemble* de árvores de decisão de maneira sequencial. Diferentemente da Random Forest, que treina múltiplas árvores de forma independente e paralela,

o XGBoost adiciona novas árvores iterativamente ao modelo. Em cada etapa, a árvore recém-criada busca modelar os resíduos produzidos pelo conjunto das árvores anteriores, equivalentes ao gradiente da função de perda utilizada na otimização (CHEN; GUESTRIN, 2016). Dessa forma, cada nova árvore atua como um mecanismo de correção dos erros acumulados ao longo do treinamento. A predição final é obtida pela soma ponderada das contribuições de todas as árvores do ensemble. A Figura 5 ilustra o fluxo sequencial de correção de resíduos característico do algoritmo.

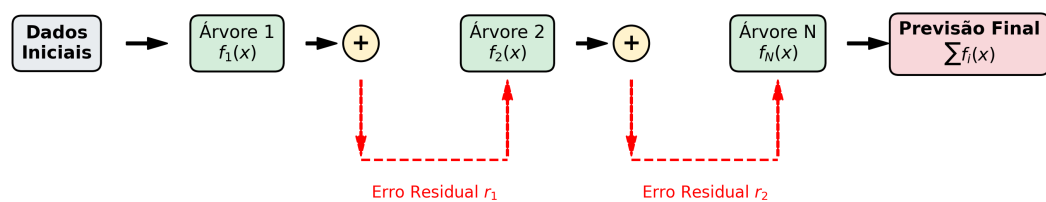


Figura 5 – Fluxo conceitual de treinamento sequencial e minimização de resíduos no XGBoost.

Fonte: Autoria própria.

O XGBoost é amplamente reconhecido como uma das abordagens mais eficazes para tarefas de classificação e regressão em dados tabulares. Segundo Grinsztajn, Oyallon e Varoquaux (2022), métodos baseados em árvores com *gradient boosting* frequentemente apresentam desempenho superior ao de arquiteturas de *deep learning* em conjuntos de dados tabulares convencionais. Essa vantagem decorre da capacidade das árvores de modelar relações não lineares e interações complexas entre atributos, aliada aos mecanismos de regularização incorporados ao algoritmo, incluindo penalizações sobre a complexidade das árvores e os pesos das folhas (CHEN; GUESTRIN, 2016). Entretanto, o presente trabalho concentra-se na avaliação da capacidade discriminativa dos atributos extraídos pela WPT para a identificação de distúrbios de qualidade da energia elétrica, e não na realização de um benchmarking abrangente entre algoritmos de aprendizado de máquina. Dessa forma, o XGBoost foi incluído na fundamentação teórica por sua relevância na literatura, mas não integrou a etapa experimental deste trabalho.

A Multi-Layer Perceptron (MLP) é uma arquitetura de rede neural artificial *feedforward*, capaz de modelar relações não lineares complexas entre variáveis de entrada e saída (HAYKIN, 2009). A rede é organizada em camadas compostas por neurônios artificiais interconectados, sendo geralmente formada por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. As camadas ocultas são responsáveis pela extração de representações intermediárias dos dados, permitindo a identificação de padrões relevantes para a tarefa de classificação.

A capacidade de modelagem da MLP decorre da utilização de funções de ativação não lineares nos neurônios ocultos. Sem essas funções, a rede se reduziria a uma única transformação linear, limitando sua capacidade de representar problemas com fronteiras de decisão complexas. Dessa forma, as funções de ativação permitem que a rede aprenda relações não lineares e realize a separação de classes não linearmente separáveis.

O treinamento da rede é realizado por meio do algoritmo de retropropagação do erro (*backpropagation*) (RUMELHART; HINTON; WILLIAMS, 1986). Inicialmente, os dados são propagados da entrada até a saída da rede (*forward propagation*), produzindo uma predição. Em seguida, uma função de custo quantifica o erro entre a saída prevista e o valor esperado. Esse erro é então propagado em sentido inverso (*backward propagation*), permitindo o cálculo dos gradientes dos parâmetros por meio da regra da cadeia. Os pesos da rede são atualizados iterativamente por algoritmos de otimização, como o Adam (KINGMA; BA, 2017), buscando minimizar a função de custo.

Devido ao elevado número de parâmetros ajustáveis, redes neurais podem apresentar sobreajuste (*overfitting*), situação em que o modelo apresenta bom desempenho sobre os dados de treinamento, mas baixa capacidade de generalização para amostras inéditas. Para reduzir esse efeito, é comum empregar técnicas de regularização, como o *early stopping*, que monitora o desempenho em um conjunto de validação e interrompe o treinamento quando não são observadas melhorias significativas na capacidade de generalização do modelo.

A Figura 6 ilustra a estrutura conceitual de uma rede MLP composta por múltiplas camadas densamente conectadas.

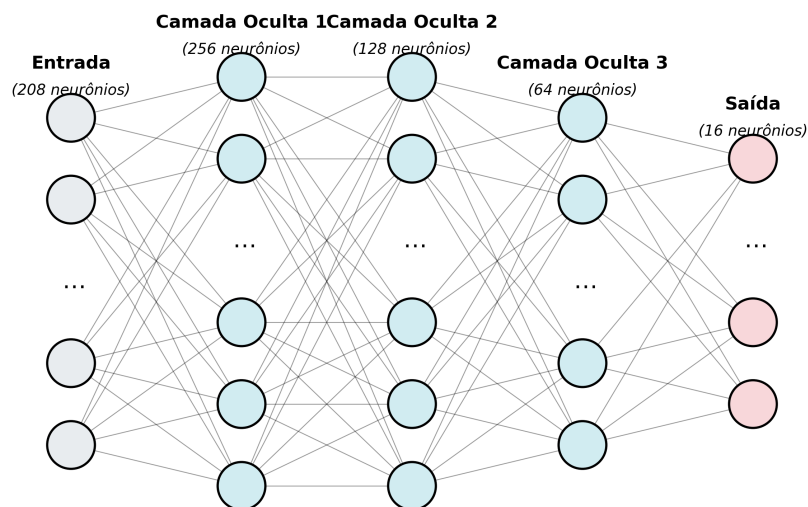


Figura 6 – Representação simplificada da arquitetura de uma rede MLP.

Fonte: Autoria própria.

2.5 Métricas de Avaliação

A validação rigorosa de classificadores em problemas multiclasse, como a identificação de 16 categorias distintas de distúrbios de QEE, exige a adoção de métricas que superem a simples proporção global de acertos. O comportamento estatístico do modelo frente a cada anomalia é extraído da Matriz de Confusão, a partir da qual são computados os índices de Verdadeiros Positivos (TP), Verdadeiros Negativos (TN), Falsos Positivos (FP) e Falsos Negativos (FN). Para mensurar o desempenho dos algoritmos investigados, este trabalho adota os critérios de avaliação amplamente empregados na literatura de aprendizado de máquina e reconhecimento de padrões (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

A Acurácia Global define-se como a razão entre o número total de predições corretas e o número total de amostras avaliadas pelo classificador.

$$Acurácia = \frac{\sum_{i=1}^C TP_i}{N_{total}} \quad (2.16)$$

Apesar de ser a métrica mais intuitiva, a acurácia apresenta limitações em conjuntos de dados desbalanceados. Em cenários reais de redes elétricas, o estado nominal (Normal) pode ocorrer com frequência significativamente superior à de falhas transitórias. Nessa situação, um classificador que previsse predominantemente a classe majoritária poderia alcançar uma elevada acurácia global, mesmo apresentando baixo desempenho na identificação dos distúrbios de interesse.

O índice F1 corresponde à média harmônica entre a Precisão (*Precision*) e o Recall, combinando simultaneamente a capacidade do modelo de evitar falsos positivos e falsos negativos. Em problemas multiclasse, utiliza-se frequentemente o Macro-F1, definido como a média aritmética não ponderada dos valores de F1 calculados individualmente para cada uma das C classes:

$$Macro-F1 = \frac{1}{C} \sum_{i=1}^C F1_i \quad (2.17)$$

O Macro-F1 fornece uma avaliação mais equilibrada do desempenho do classificador, pois atribui o mesmo peso a todas as classes, independentemente de sua frequência no conjunto de teste. Dessa forma, um desempenho insatisfatório em classes minoritárias impacta diretamente o valor final da métrica.

O Recall exprime numericamente a fração de ocorrências reais pertencentes a uma classe específica i que o modelo foi capaz de identificar corretamente.

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (2.18)$$

Na identificação de distúrbios de qualidade da energia, a análise individual do Recall por classe é particularmente relevante, pois permite identificar quais categorias apresentam maiores dificuldades de reconhecimento. Valores reduzidos de Recall indicam uma elevada taxa de Falsos Negativos, situação crítica em aplicações de monitoramento e diagnóstico de falhas.

O Log-Loss avalia a qualidade das probabilidades produzidas pelo classificador, medindo o grau de calibração das previsões probabilísticas. Para um problema contendo N amostras e C classes, essa métrica é definida por:

$$Log-loss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \ln(p_{i,j}) \quad (2.19)$$

onde $y_{i,j}$ representa uma variável indicadora que assume valor 1 quando a classe j corresponde à classe verdadeira da amostra i , e 0 caso contrário. Já $p_{i,j}$ corresponde à probabilidade estimada pelo modelo para essa mesma classe.

O Log-Loss penaliza previsões incorretas realizadas com elevada confiança. Consequentemente, dois classificadores podem apresentar a mesma acurácia global e, ainda assim, exibir valores distintos de Log-Loss caso um deles produza probabilidades excessivamente confiantes para previsões incorretas.

A Matriz de Confusão Normalizada consiste no mapeamento cruzado entre as classes reais e as classes previstas pelo classificador, no qual cada elemento da matriz é dividido pelo número total de amostras da respectiva classe verdadeira. Esse procedimento converte contagens absolutas em frequências relativas, facilitando a comparação entre classes com diferentes quantidades de amostras.

Após a normalização, os elementos da diagonal principal passam a representar o Recall de cada classe. Dessa forma, a matriz permite identificar padrões de confusão entre categorias específicas de distúrbios, evidenciando quais fenômenos apresentam maior sobreposição do ponto de vista do processo de classificação.

3 Metodologia

Esta seção do trabalho detalha os procedimentos práticos e computacionais adotados para o desenvolvimento do sistema de classificação de Distúrbios de Qualidade de Energia Elétrica (PQD). A metodologia foi estruturada para garantir a total reprodutibilidade do estudo, descrevendo as etapas de geração de dados, processamento de sinais, preparação do conjunto de atributos e configuração dos algoritmos de aprendizado de máquina.

3.1 Geração dos Sinais Sintéticos

Os sinais elétricos foram gerados computacionalmente através da implementação das equações paramétricas correspondentes a 16 classes de distúrbios, conforme as diretrizes da norma IEEE 1159-2019 (IEEE, 2019). A configuração base para a síntese das formas de onda adotou uma frequência fundamental $f_0 = 50$ Hz, uma frequência de amostragem $F_s = 5000$ Hz e uma amplitude nominal ideal $A = 1,0$ p.u.. Cada sinal gerado compreendeu uma janela de observação temporal de exatos 10 ciclos da componente fundamental.

Para assegurar a generalização do modelo, aplicou-se uma variação paramétrica estocástica rigorosa na geração dos dados. Para cada instância, os parâmetros de falha (como profundidade do afundamento, amplitude harmônica ou transientes) foram amostrados aleatoriamente a partir de distribuições uniformes dentro das faixas limites estabelecidas pela norma. Adicionalmente, o ângulo de fase inicial φ de cada sinal foi fixado inicialmente em $\varphi = 0$.

A introdução dessa variação paramétrica contínua é um ponto metodológico crítico deste estudo. Sem ela, os algoritmos de classificação tenderiam a memorizar formas de onda ou transições de fase específicas presentes no conjunto de treinamento, em vez de aprenderem as características dinâmicas fundamentais que definem a classe do distúrbio. Esse sobreajuste inflacionaria artificialmente a acurácia do modelo em testes isolados, mas colapsaria a sua utilidade em aplicações no mundo real.

Para simular o ruído de redes de distribuição práticas, cada sinal sintético puro gerado foi replicado sob diferentes níveis de Ruído Branco Gaussiano Aditivo (AWGN). Foram produzidas 5 variantes para cada instância: o sinal original limpo e quatro versões corrompidas com Relação Sinal-Ruído (SNR) de 20 dB, 30 dB, 40 dB e 50 dB. O banco de dados final foi constituído por 2000 instâncias base por classe, que multiplicadas pelas 16 classes e pelas 5 variantes de ruído, totalizaram uma base de 160000 sinais analisáveis. Todo o processo de geração utilizou sementes aleatórias fixas (*random seed*) para garantir

a reprodutibilidade exata dos tensores gerados.

3.2 Decomposição WPT e Extração de Atributos

O processamento digital dos sinais foi realizado mediante a aplicação da Transformada Wavelet Packet (WPT), utilizando a linguagem Python e a biblioteca *PyWavelets*. A função base selecionada foi a Daubechies de ordem 8 (db8), de acordo com a literatura apresentada no trabalho de [Varanis e Pederiva \(2018\)](#). A decomposição foi executada até o nível $L = 4$, resultando na divisão do espectro útil do sinal (0 a 2500 Hz) em 16 sub-bandas de frequência independentes e de igual resolução espectral igual a 156,25 Hz.

Para solucionar o problema da alta dimensionalidade dos vetores de coeficientes no domínio wavelet, aplicou-se um protocolo de extração de atributos focado em redução de dados. Em cada uma das 16 sub-bandas, foram calculadas 13 características estatísticas e entrópicas, apresentadas na Tabela 2 da sub-seção 2.3. O processo condensou cada janela temporal de 10 ciclos em um vetor de características global de dimensionalidade estática, correspondendo a $13 \times 16 = 208$ atributos por sinal.

3.3 Preparação dos Dados

Para a etapa de validação cruzada geométrica, o banco de atributos extraído foi submetido a uma divisão espacial estratificada, garantindo a preservação da proporção balanceada das 16 classes. Os dados foram particionados nas proporções de 70% para o conjunto de treinamento, 10% para o conjunto de validação interna e 20% para o teste final.

O tratamento do vetor de características exigiu a padronização das escalas numéricas para evitar que métricas de magnitudes distintas (como Energia Total e Assimetria) enviesassem os cálculos de distância dos classificadores. Utilizou-se a normalização padronizada (*StandardScaler*) disponibilizada pela biblioteca *scikit-learn* ([PEDREGOSA et al., 2011](#)), ajustando (*fit*) as médias e variâncias exclusivamente sobre os 70% do conjunto de treinamento. Os parâmetros extraídos foram então aplicados de forma linear sobre os conjuntos de validação e teste. Esse protocolo de restrição impediu o vazamento de informações estatísticas (*data leakage*) das amostras de teste para a fase de aprendizado.

3.4 Classificação

Os vetores de 208 atributos foram submetidos a diferentes paradigmas de algoritmos preditivos, todos instanciados e treinados através do ecossistema *scikit-learn*. Conforme delineado na fundamentação teórica, a escolha da arquitetura dos modelos privilegiou métodos de robustez comprovada. A configuração definitiva dos hiperparâmetros utilizados no treinamento de cada classificador está sintetizada na Tabela 3.

Tabela 3 – Configuração dos classificadores utilizados na fase experimental.

Classificador	Parâmetros	Justificativa
LDA	<code>solver='lsqr'</code> <code>shrinkage='auto'</code>	Avaliação do baseline de separabilidade linear.
Random Forest	<code>n_{est} = 300</code> <code>min_leaf=2</code>	Comitê robusto à variação estocástica e ruído.
MLP	Ocultas: (256,128,64) <code>solver='adam'</code> <code>early stopping</code> ativo	Baseline não linear de rede neural profunda com parada antecipada.

A definição dos hiperparâmetros detalhados na Tabela 3 não ocorreu de forma arbitrária, sendo estritamente fundamentada na literatura de aprendizado estatístico para lidar com a alta dimensionalidade (208 atributos simultâneos) e com a presença de ruído estocástico no conjunto de dados.

Para a Análise Discriminante Linear (LDA), em espaços de características multidimensionais, a matriz de covariância tende a se tornar mal condicionada, prejudicando a inversão matricial necessária para traçar as fronteiras de decisão. O método de *shrinkage* implementado no ecossistema *scikit-learn* regulariza a estimativa da covariância de forma analítica (PEDREGOSA et al., 2011), garantindo estabilidade numérica e produzindo estimativas mais estáveis da matriz de covariância e aumentando a robustez do classificador diante da presença de ruído e correlações entre atributos.

No modelo Random Forest, a escolha de 300 árvores (`n_estimators=300`) baseia-se no princípio formalizado por Breiman (2001) de que a adição sucessiva de classificadores fracos independentes em um comitê não agrava o sobreajuste, fazendo com que o erro de generalização convirja para um limite assintótico. O número de árvores foi definido como 300, valor suficientemente elevado para garantir a estabilização do desempenho do comitê sem aumento significativo do custo computacional. Adicionalmente, a restrição mínima de amostras por folha (`min_samples_leaf=2`) atuou como um mecanismo de regularização geométrica; permitir nós com apenas uma amostra forçaria a árvore a memorizar o ruído estocástico (decoração do erro), enquanto a exigência de agrupamento mínimo força o particionamento a capturar características estruturais reais da falha.

Para a rede neural MLP, a arquitetura adotada consiste em três camadas ocultas contendo 256, 128 e 64 neurônios, respectivamente. A configuração foi definida de forma empírica visando fornecer capacidade representacional suficiente para modelar as relações não lineares presentes nos 208 atributos extraídos pela WPT, ao mesmo tempo em que mantém uma redução progressiva da dimensionalidade interna da rede. Essa compressão arquitetural força a rede a extrair padrões latentes dos 208 atributos brutos, reduzindo a dimensionalidade para abstrações matemáticas mais densas antes do diagnóstico na camada de saída. O otimizador estocástico Adam (*Adaptive Moment Estimation*) foi adotado por sua eficácia empírica em dados com distribuições complexas, ajustando taxas de aprendizado individuais e dinâmicas para cada parâmetro sináptico (KINGMA; BA, 2017). A regularização principal do modelo deu-se pelo mecanismo de parada antecipada (*early stopping*): ao invés de iterar por um número cego de épocas, o processo foi compulsoriamente encerrado no exato instante em que a função de custo sobre os 10% do conjunto de validação estagnou, prevenindo o colapso da generalização e a memorização de anomalias isoladas (GOODFELLOW; BENGIO; COURVILLE, 2016).

3.5 Ambiente Computacional

Todos os experimentos de geração de sinais, processamento digital e treinamento dos algoritmos de aprendizado de máquina foram executados de forma local. O hardware utilizado para a condução do estudo consistiu em uma estação de trabalho equipada com um processador de alto desempenho AMD Ryzen 5 5600X com 6 núcleos e 12 *threads*, 16 GB de memória RAM primária, armazenamento de alta velocidade em unidade de estado sólido (SSD) e uma placa de vídeo dedicada AMD Radeon RX 7600 com 12 GB de memória de vídeo (VRAM).

O ecossistema de *software* foi estruturado sobre o sistema operacional Bazzite (uma distribuição Linux de natureza atômica, fundamentada no Fedora), garantindo um ambiente de execução isolado e reproduzível. O código-fonte foi desenvolvido inteiramente em linguagem Python (versão 3.14.4). As rotinas matemáticas e o processamento de tensores contaram com a biblioteca *NumPy*; as extrações de sub-bandas operaram sobre a biblioteca especializada *PyWavelets*; e as rotinas de aprendizado de máquina, particionamento e extração de métricas de avaliação foram invocadas a partir do ecossistema *scikit-learn*.

O tempo de execução computacional caracterizou-se por um intenso gargalo na fase de extração analítica das 208 características WPT para a totalidade da base de dados (160.000 instâncias), demandando processamento contínuo de CPU, totalizando tempo de execução em torno da ordem de 30 minutos. Em contraste, a fase de treinamento, inferência e classificação estática através dos modelos propostos ocorreu de maneira expressiva e rápida na mesma máquina, com o classificador LDA tendo tempo de execução na ordem

de poucos segundos, enquanto os classificadores RF e MLP possuíram tempo de execução na ordem de 600 segundos.

4 Resultados e Discussão

Esta é a seção central do presente trabalho, destinada à apresentação e interpretação da avaliação empírica do sistema de classificação de PQD. Os resultados foram estruturados de forma progressiva: inicialmente, valida-se a integridade física do banco de dados sintético; em seguida, realiza-se o *benchmark* de desempenho global dos classificadores; e, por fim, aprofunda-se a análise sobre as matrizes de confusão e as taxas de revocação, correlacionando os acertos e falhas dos modelos com a natureza eletromagnética dos fenômenos estudados.

4.1 Verificação dos Sinais Gerados

Antes da extração de características no domínio tempo-frequência, foi fundamental atestar a conformidade física e matemática dos sinais sintetizados em relação aos parâmetros estipulados pela norma [IEEE \(2019\)](#).

A Figura 7 ilustra uma amostra temporal limpa (sem adição de ruído estocástico) de cada uma das 16 classes de distúrbios modeladas neste estudo. É possível observar visualmente as assinaturas características de cada falha, tais como a atenuação de amplitude no afundamento de tensão (*Voltage sag*), a distorção periódica no *Flicker* e a ocorrência de descontinuidades abruptas nos transientes oscilatórios e impulsivos. O isolamento dessas formas de onda em regime ideal confirma o correto funcionamento das equações geradoras paramétricas. O apêndice A apresenta os gráficos das 16 classes de PQD com sobreposições de cada instância gerada, de forma aleatória.

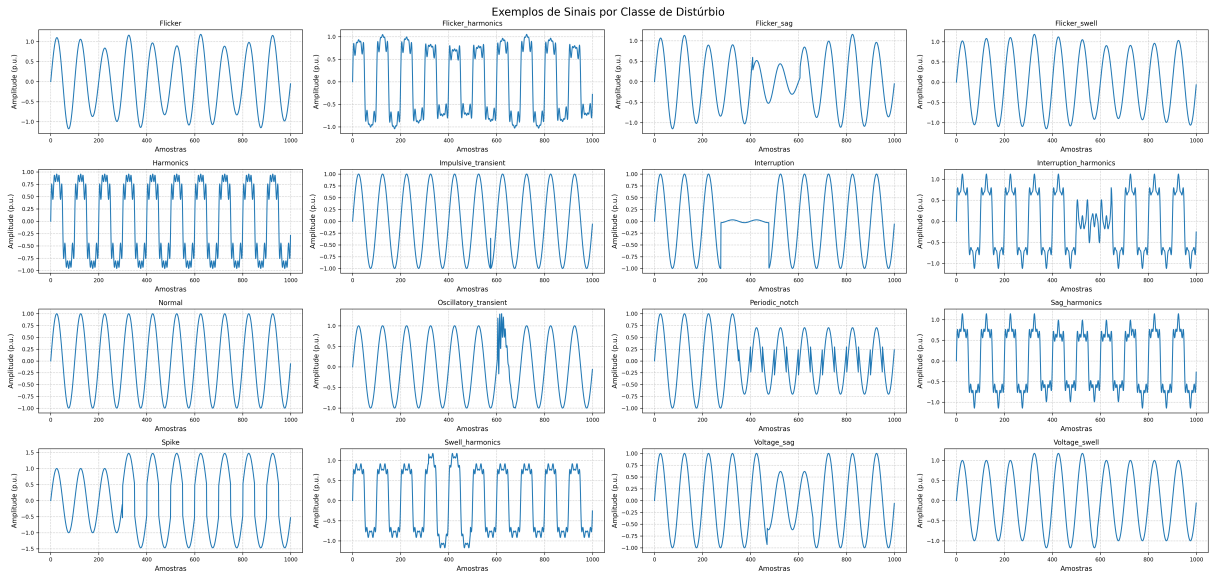


Figura 7 – Amostras limpas das 16 classes de distúrbios de QEE modeladas no domínio do tempo.

Fonte: Autoria própria.

Para validar o domínio da frequência do gerador paramétrico, os sinais foram submetidos à Transformada Rápida de Fourier (FFT), cujo espectro encontra-se plotado na Figura 8. Conforme exigido pelas premissas do sistema elétrico de potência padrão simulado, o espectro da classe normal apresenta um único e pronunciado pico de densidade de energia cravado exatamente em 50 Hz. A ausência de vazamento espectral ou em altas frequências atesta a pureza do sinal fundamental antes da injeção das perturbações.

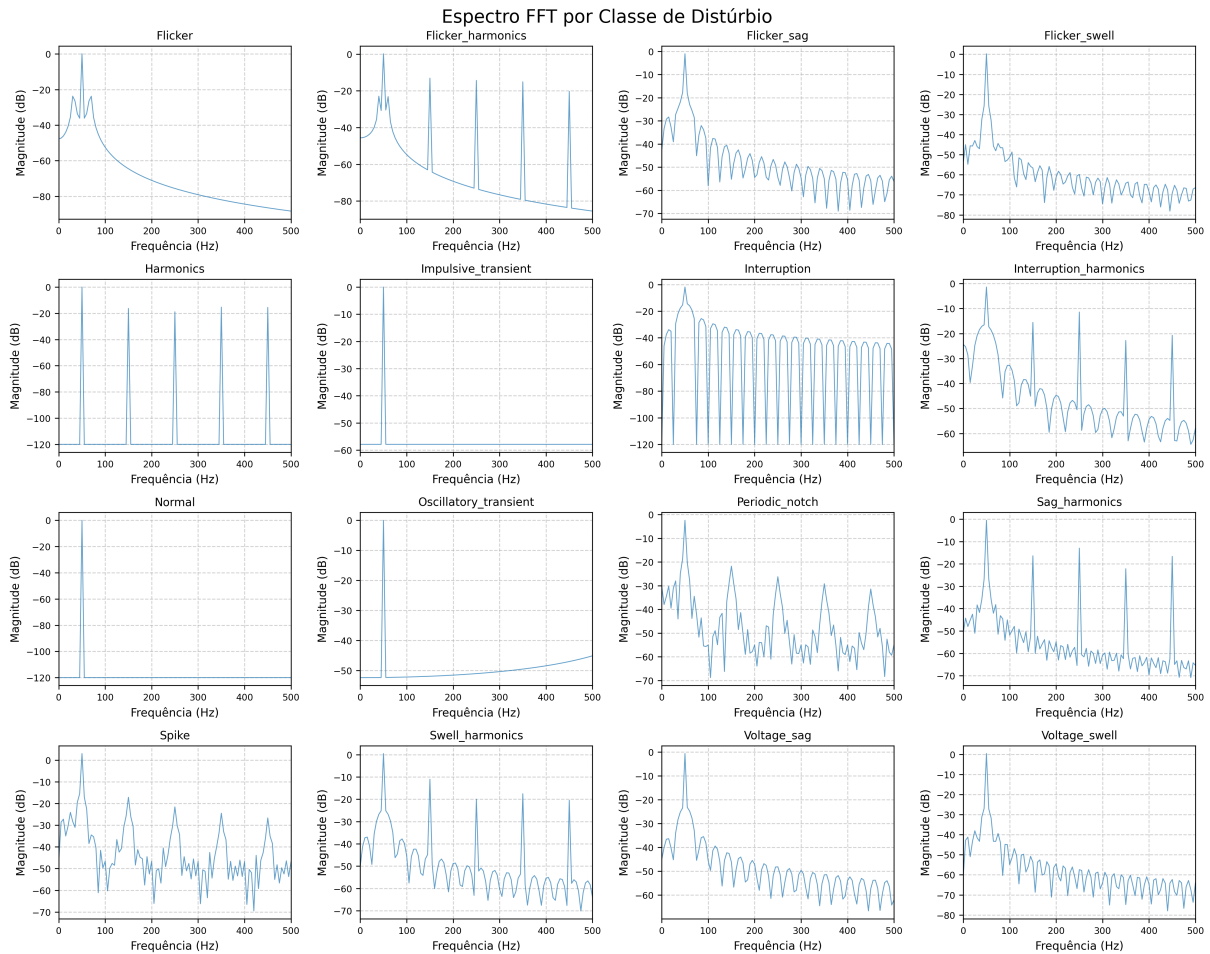


Figura 8 – Espectro FFT do sinal sem ruído, destacando o componente fundamental em 50 Hz.

Fonte: Autoria própria.

Por fim, a diversidade paramétrica e a inserção das variações estocásticas, essenciais para evitar o sobreajuste dos classificadores, são demonstradas no domínio da frequência através da Figura 9. O gráfico exibe a sobreposição dos espectros de magnitude (FFT, em decibéis) de múltiplas instâncias sorteadas aleatoriamente para cada uma das 16 classes, já submetidas aos diferentes níveis de AWGN.

A análise visual destes espectros revela assinaturas intrínsecas às falhas. O componente fundamental de 50 Hz domina a energia na maioria das classes, enquanto os distúrbios de natureza harmônica (*Harmonics*, *Sag_harmonics*, *Swell_harmonics*, entre outros) exibem picos secundários evidentes e precisos nas frequências múltiplas (como 150 Hz, 250 Hz e 350 Hz). Na classe *Oscillatory_transient*, nota-se a elevação de um amplo lóbulo de energia em frequências mais altas (tipicamente entre 300 Hz e 400 Hz), característico do decaimento amortecido do transiente.

Além disso, o impacto da variação de ruído é evidente na flutuação do piso de energia de fundo (*noise floor*) entre as linhas de uma mesma classe. Amostras com maior SNR (ex: 50 dB) apresentam atenuação acentuada nas altas frequências, enquanto os cenários

mais severos (ex: 20 dB) elevam o ruído de fundo global, mascarando os componentes espectrais de menor amplitude.

Esta representação estocástica ilustra a complexidade do problema de classificação. A Transformada de Fourier, embora útil para a verificação estacionária de harmônicos, perde completamente a localização temporal do evento (por exemplo, a identificação do exato instante em que o afundamento iniciou ou o transiente ocorreu). Essa limitação teórica inerente à FFT, somada ao mascaramento das assinaturas provocado pelo ruído variável, justifica a escolha metodológica pela Transformada Wavelet Packet (WPT) no presente trabalho. A WPT permite segmentar a energia sem perder a resolução de tempo-frequência, etapa imprescindível para alimentar os classificadores de Inteligência Artificial.

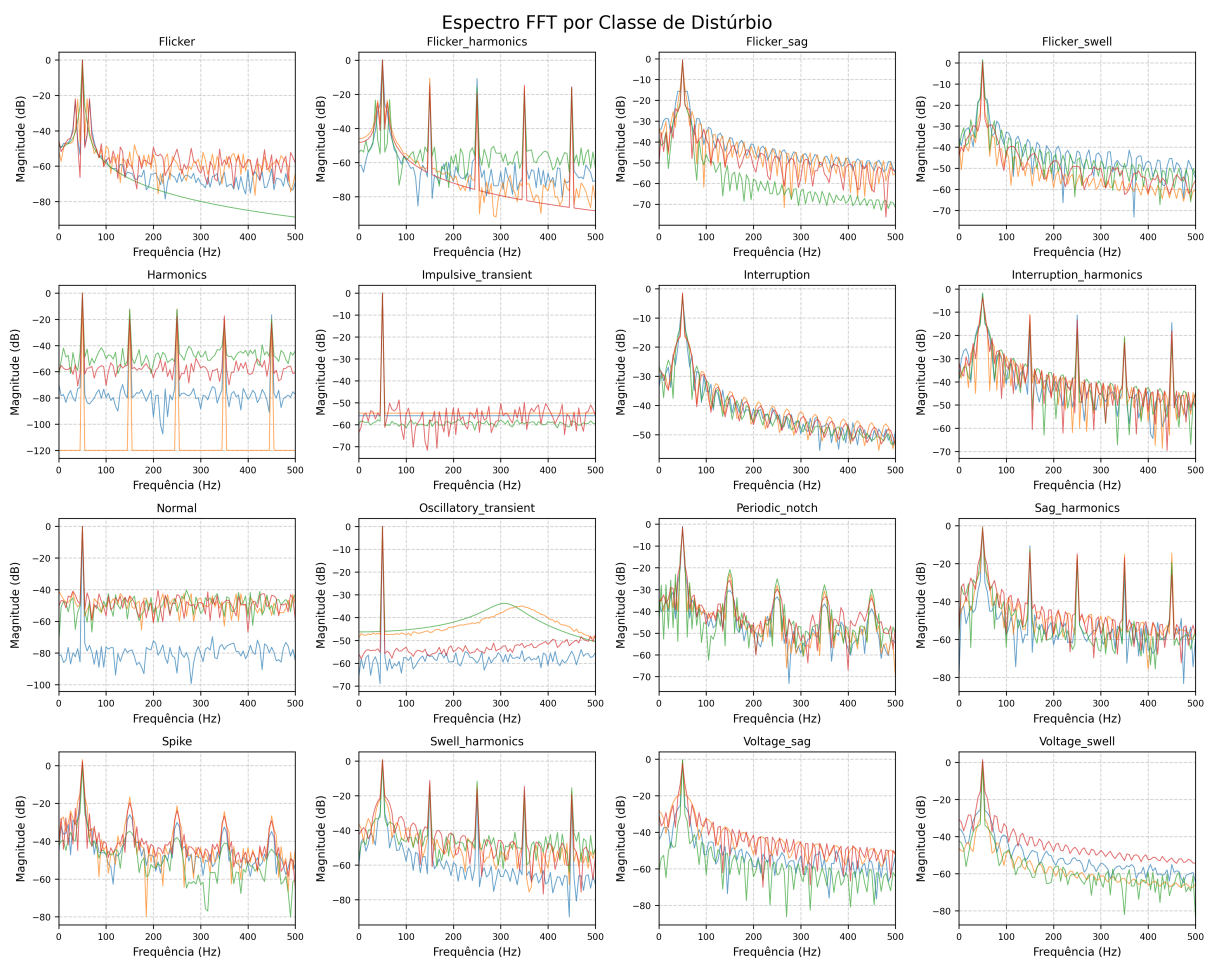


Figura 9 – Espectro FFT sobreposto de amostras estocásticas para as 16 classes, ilustrando as componentes harmônicas e a variação do piso de ruído AWGN.

Fonte: Autoria própria.

Para além da verificação espectral, avaliou-se a qualidade do espaço de características obtido pela WPT, através da projeção PCA e t-SNE. Os dois primeiros componentes PCA explicam 46,7% da variância total, indicando uma distribuição não-degenerada dos dados em alta dimensão. O t-SNE revela aglomerados bem separados para a maioria dos distúrbios, com sobreposição apenas entre classes de assinatura semelhante (ex: sag e sag

com harmônicos). Complementarmente, o cálculo do F-score de Fisher identificou as 30 características mais discriminantes, a maioria delas correspondendo a nós da WPT nas bandas de frequência críticas. Essas validações confirmam que o conjunto de características é adequado para a classificação supervisionada que se segue. O apêndice A apresenta estes gráficos e explicações em maiores detalhes.

4.2 Benchmark de Classificadores

A avaliação quantitativa do desempenho de generalização dos modelos foi conduzida sobre o conjunto de teste isolado, composto por 32.000 instâncias inéditas. Para garantir uma comparação justa entre diferentes famílias algorítmicas, os modelos foram submetidos às mesmas matrizes de atributos normalizadas. Os resultados globais de Acurácia, F1-Score (macro), Perda Logarítmica (*Log-Loss*) e o tempo computacional de treinamento estão consolidados na Tabela 4.

Tabela 4 – Resultados do desempenho de classificadores — WPT (db8, L = 4).

Método	Acurácia	Macro-F1	Log-Loss	Tempo (s)
LDA (Shrinkage)	0,8539	0,8500	0,7128	~ 3,91
Random Forest	0,9444	0,9400	0,1749	~ 538
MLP	0,9521	0,9500	0,1366	~ 557

A análise da Tabela 4 revela um salto de desempenho substancial entre a abordagem linear e as arquiteturas baseadas em Inteligência Artificial. O classificador de referência, Análise Discriminante Linear (LDA), atingiu a menor acurácia do *benchmark* (85,39%), acompanhada de um *Log-Loss* expressivamente elevado (0,7128). O *Log-Loss* penaliza predições divergentes baseadas na probabilidade de confiança geométrica; O valor elevado de *Log-Loss* obtido pelo LDA indica menor qualidade na calibração probabilística das predições quando comparado aos demais modelos. Embora o classificador alcance acertos em uma parcela significativa das amostras, suas distribuições de probabilidade apresentam maior incerteza, refletindo limitações da hipótese de separabilidade linear imposta pelo modelo.

Em contrapartida, o comitê de árvores Random Forest apresentou uma calibração substancialmente superior, reduzindo o *Log-Loss* para 0,1749 e elevando a acurácia global para 94,44%. Por se tratar de um modelo de particionamento ortogonal que não assume distribuições normais predefinidas para os dados, o Random Forest foi capaz de delimitar fronteiras rígidas ao redor das classes mais afetadas pela dispersão do ruído harmônico.

O desempenho ótimo foi alcançado pela Rede Neural Multi-Layer Perceptron (MLP), que atingiu 95,21% de acurácia global e 0,95 de Macro-F1. A superioridade da

MLP comprova que a topologia do problema exige projeções não lineares. As funções de ativação nas camadas ocultas da rede atuam distorcendo o hiperplano de características, permitindo o mapeamento de assinaturas complexas — como a fusão entre transientes e distorções harmônicas — que se apresentam topologicamente emaranhadas no espaço linear. O custo computacional associado a esta capacidade analítica refletiu-se apenas no tempo de treinamento (557 segundos para a convergência completa), um valor amplamente aceitável para aplicações de engenharia, especialmente considerando que o tempo de inferência (*feedforward*) de novos sinais permaneceu na ordem de milissegundos. Vale destacar que a rede atingiu a convergência em apenas 46 épocas, indicando que o espaço de características da WPT, embora não linearmente separável, possui alta estruturação geométrica, facilitando o aprendizado das fronteiras de decisão pela MLP.

A comparação com a literatura recente, como o trabalho de [Singh, Pal e Dahiya \(2023\)](#), evidencia diferenças metodológicas relevantes na construção dos sistemas de classificação. Em seu estudo, os autores reportaram acurácias superiores a 98% utilizando uma estratégia baseada em LDA para redução e seleção de características, seguida da aplicação de classificadores como SVM, kNN, Naive Bayes e Random Forest. No presente trabalho, por outro lado, optou-se pela Transformada Wavelet Packet (WPT) como técnica principal de extração de atributos, explorando simultaneamente informações temporais e espectrais dos sinais de qualidade de energia.

Embora a WPT apresente maior custo computacional durante a etapa de processamento, sua capacidade de representar fenômenos transitórios e não estacionários torna-a particularmente adequada para a caracterização de distúrbios de qualidade de energia elétrica. Dessa forma, as diferenças observadas entre os resultados não devem ser interpretadas apenas em termos de desempenho numérico, mas também à luz das distintas estratégias de representação dos sinais adotadas em cada estudo.

4.3 Análise da Matriz de Confusão

Uma avaliação aprofundada da acurácia de 95,21% atingida pela MLP requer a inspeção das taxas de acerto e erro por categoria. A Figura 10 apresenta a Matriz de Confusão Normalizada gerada sobre o conjunto de teste, onde a diagonal principal reflete a taxa de revocação (*Recall*) de cada classe. A normalização linha a linha permite evidenciar diretamente as tendências de falso-negativo.

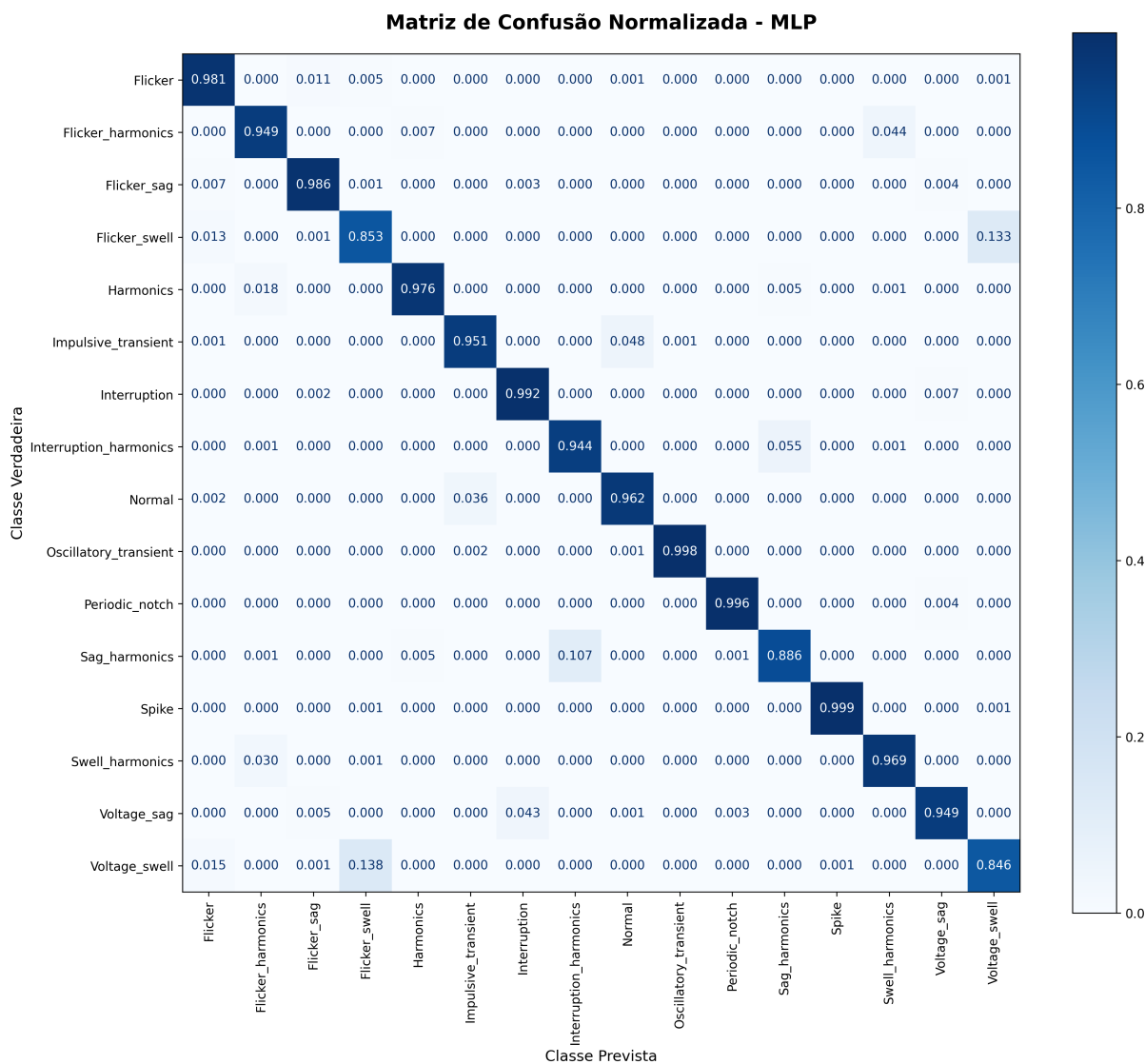


Figura 10 – Matriz de Confusão Normalizada para a Rede Neural MLP.
 Fonte: Autoria própria.

A inspeção da matriz revela que o classificador não comete erros aleatórios: as confusões ocorrem majoritariamente entre distúrbios que compartilham similaridades físicas e morfológicas, as quais são borradas pela presença do ruído AWGN. O caso de maior desafio discriminativo reside na fronteira entre as classes *Voltage_swell* e *Flicker_swell*. O modelo confunde instâncias verdadeiras de *Voltage_swell* prevendo-as como *Flicker_swell* em 13,8% das vezes, e comete o erro inverso em 13,3% dos casos. Este comportamento é fisicamente consistente: ambos os fenômenos envolvem uma elevação sustentada da amplitude de tensão. Quando um degrau de *swell* é severamente mascarado por ruído de alta potência, a ondulação aparente do sinal aproxima-se da oscilação de baixa frequência característica da modulação de *flicker*, estreitando a margem de separabilidade na WPT.

Outra confusão esperada com coerência fenomenológica é a classificação errônea de *Interruption_harmonics* como *Sag_harmonics* em 6% das amostras. A distinção norma-

tiva entre um afundamento profundo e uma interrupção é o limiar de tensão remanescente (tipicamente 10% da nominal). A adição simultânea de distorção harmônica severa e ruído estocástico atenua a nitidez deste limiar de amplitude, empurrando as características de entropia de uma falha para a zona de decisão da outra.

Apesar destas fronteiras nebulosas, a robustez do extrator WPT destaca-se nas classes de variação rápida e descontinuidade temporal. Fenômenos como *Impulsive_transient*, *Oscillatory_transient*, *Spike* e *Periodic_notch* atingiram taxas de acerto perfeitas ou quase perfeitas (acima de 95%). Estes distúrbios possuem assinaturas de alta energia muito abruptas e localizadas em faixas espectrais elevadas, características que as *wavelets-mãe* de suporte compacto, como a Daubechies (db8), conseguem isolar com exímia precisão, preservando boa capacidade discriminativa mesmo na presença de ruído. As Matrizes de Confusão dos outros dois classificadores, RF e LDA, estão apresentadas no Apêndice A.

De maneira geral, observa-se que os erros concentram-se em pares de classes que compartilham características físicas semelhantes, enquanto os distúrbios com assinaturas espectrais mais distintas apresentam elevadas taxas de *recall*. Esse comportamento sugere que a principal limitação do sistema não está na identificação de eventos abruptos ou altamente energéticos, mas na diferenciação de fenômenos que produzem alterações correlacionadas na amplitude e no conteúdo harmônico do sinal. Ainda assim, a predominância dos elementos na diagonal principal da matriz confirma a elevada capacidade discriminativa da combinação WPT-MLP, corroborando os resultados globais de acurácia e Macro-F1 apresentados na Tabela 4.

4.4 Recall por Classe

Para complementar a análise da matriz de confusão, a Tabela 5 detalha a métrica de revocação (*Recall*) individual de cada classe obtida pelo modelo definitivo (MLP). Conforme estabelecido na metodologia, esta métrica quantifica a proporção de instâncias verdadeiramente pertencentes a um distúrbio que foram corretamente isoladas pelo classificador. Classes que apresentaram desempenho inferior ao limiar de excelência estabelecido para este estudo ($\text{Recall} < 0,95$) encontram-se sinalizadas.

Tabela 5 – Taxa de revocação (*Recall*) por classe de distúrbio para a Rede Neural MLP.

Classe de Distúrbio	Recall
Spike	0,999
Oscillatory_transient	0,998
Periodic_notch	0,996
Interruption	0,992
Flicker_sag	0,986
Flicker	0,981
Harmonics	0,976
Swell_harmonics	0,969
Normal	0,962
Impulsive_transient	0,951
Flicker_harmonics*	0,949
Voltage_sag*	0,949
Interruption_harmonics*	0,944
Sag_harmonics*	0,886
Flicker_swell*	0,853
Voltage_swell*	0,846

* Classes com recall inferior a 0,95.

A análise da Tabela 5 evidencia uma dicotomia clara na complexidade de classificação estrutural. Observa-se que os maiores índices de revocação (*recall*) concentram-se em distúrbios que apresentam assinaturas energéticas marcantes, seja por eventos transitórios de alta frequência, seja por alterações expressivas de amplitude ao longo do sinal. A WPT extrai as assinaturas destas anomalias de alta frequência com perdas praticamente nulas, tornando-as altamente distinguíveis para a rede neural.

Por outro lado, o grupo de seis classes que não atingiu o limiar de 0,95 compartilha uma raiz física comum: a sobreposição contínua de envelope de amplitude com distorções de estado plano. As menores taxas de acerto (*Voltage_swell* com 0,846 e *Flicker_swell* com 0,853) refletem o mascaramento estocástico discutido na seção anterior. De forma análoga, a queda de desempenho em *Sag_harmonics* (0,886) e *Interruption_harmonics* (0,944) demonstra que a injeção severa de ruído compromete a estimação da profundidade do afundamento, dificultando a diferenciação entre um Sag severo e uma interrupção parcial da rede. Esta redução de desempenho sugere uma região de sobreposição entre as classes, decorrente tanto das características físicas dos fenômenos quanto da degradação provocada pelo ruído, indicando uma limitação inerente ao problema de classificação e não apenas ao algoritmo empregado.

4.5 Discussão Geral e Comparação com a Literatura

A avaliação global do sistema proposto, que culminou na acurácia de 95,21% utilizando a combinação WPT e MLP, consolida-se de forma expressiva quando posta em perspectiva com o estado da arte. A literatura recente apresenta abordagens diversificadas para o diagnóstico de PQD, frequentemente reportando acurácias na faixa de 95% a 99%. Contudo, uma análise crítica revela assimetrias metodológicas, especialmente no que tange à complexidade geométrica do banco de dados e à modelagem estocástica do ruído.

A diferença observada entre o desempenho da LDA (85,39%) e dos classificadores não lineares (94,44% para RF e 95,21% para MLP) fornece evidência de que a Transformada Wavelet Packet, embora altamente eficaz na extração de informações discriminantes, não lineariza completamente o problema de classificação dos distúrbios de qualidade de energia. Caso o espaço de características produzido pela WPT fosse linearmente separável, seria esperado que a LDA apresentasse desempenho próximo ao dos demais modelos. A superioridade consistente da RF e da MLP sugere a permanência de relações não lineares entre os atributos extraídos, especialmente nas classes compostas por múltiplos fenômenos simultâneos e sujeitas à degradação por ruído AWGN. Dessa forma, os resultados indicam que a principal contribuição da WPT não consiste em tornar as classes linearmente separáveis, mas em concentrar informações relevantes em um conjunto compacto de atributos, simplificando o trabalho dos classificadores subsequentes.

Outro ponto de discussão central estabelecido neste estudo refere-se à real contribuição dos atributos entrópicos para a separabilidade geométrica do problema. A análise de poder discriminativo, quantificada estatisticamente pela Razão de Fisher (detalhada no Apêndice A), revelou que as métricas de entropia, notadamente *Log-Energy*, Permutação, Shannon e Rényi, dominam o ranqueamento das 30 características mais relevantes entre as 208 extraídas do pacote Wavelet. Esta predominância possui uma justificativa física direta: atributos estatísticos clássicos, como valor RMS, desvio padrão ou máximo absoluto, baseiam-se na amplitude do sinal, sendo severamente corrompidos e mascarados pela injeção do ruído estocástico AWGN e pelas variações paramétricas de profundidade das falhas. Em contrapartida, as formulações de entropia quantificam a complexidade, a desorganização e a distribuição espectral da energia dentro das sub-bandas de alta frequência. Dessa forma, mesmo quando a topologia temporal de um distúrbio é borrada por forte ruído de fundo, a sua assinatura de desorganização permanece estável. Conclui-se, portanto, que a inclusão das entropias não atua como um mero complemento numérico, mas como o vetor de informações fundamental que permite às redes neurais sustentar acurácias superiores a 95% em cenários de alta distorção. Apesar do amplo uso de atributos energéticos e estatísticos na literatura de PQD, a combinação simultânea das quatro

métricas de entropia (Log-Energy, Permutação, Shannon e Rényi) como vetor de características extraídas da WPT não foi encontrada nos trabalhos revisados, sugerindo um caráter inovador para a abordagem proposta.

O trabalho de [Mishra \(2019\)](#), um abrangente artigo de revisão bibliográfica (*survey*), compila diversas pesquisas que atestam a eficácia da Transformada WPT e suas variantes, reportando acurácias consistentemente superiores a 95% utilizando classificadores Support Vector Machine (SVM). No entanto, a maioria dos trabalhos revisados na literatura restringe-se à classificação de, no máximo, 8 categorias de distúrbios elétricos. O salto metodológico de 8 para 16 classes, incorporando distúrbios compostos complexos, como afundamentos e elevações conjugados com harmônicos (*Sag_harmonics*, *Swell_harmonics*) ou modulações de envelope (*Flicker_swell*), aumenta significativamente a sobreposição das fronteiras de decisão. Assim, a manutenção do patamar de excelência (95%) em um problema envolvendo o dobro de classes e uma maior diversidade de distúrbios compostos corrobora a robustez do vetor de 208 características estruturado neste estudo.

A comparação com metodologias específicas de extração de características, como a proposta por [Singh, Pal e Dahiya \(2023\)](#), evidencia diferenças relevantes de escopo experimental. Os autores reportam acurácias superiores a 98% utilizando uma etapa de redução de dimensionalidade baseada em Análise Discriminante Linear (LDA), combinada com classificadores tradicionais. Em contraste, o presente trabalho emprega a Transformada Wavelet Packet (WPT) para representar os sinais no domínio tempo-frequência, preservando informações espectrais e temporais associadas a fenômenos transitórios e não estacionários.

Resultados igualmente elevados também são reportados por [Sipai, Jadeja e Tank \(2024\)](#) e [Wang, Xu e Che \(2019\)](#), com acurácias superiores a 98%. Entretanto, os cenários experimentais adotados nesses estudos diferem substancialmente do presente trabalho. Enquanto tais pesquisas utilizam bases de dados da ordem de milhares de amostras, o banco de dados aqui construído foi expandido para 160.000 sinais a partir da combinação de 16 classes de distúrbios, múltiplos níveis de ruído AWGN (20 dB a 50 dB) e variações aleatórias dos parâmetros físicos das falhas, incluindo profundidade, duração e instante de ocorrência.

Essa estratégia aumenta significativamente a diversidade estatística do conjunto de dados e reduz a probabilidade de que os modelos aprendam apenas padrões específicos do processo de geração. Consequentemente, a comparação direta entre valores absolutos de acurácia deve ser realizada com cautela, uma vez que diferenças na quantidade de classes, na modelagem do ruído e na variabilidade paramétrica alteram substancialmente a complexidade do problema de classificação. Nesse contexto, a acurácia de 95,21% obtida pela combinação WPT-MLP demonstra elevada capacidade de generalização mesmo sob

condições de treinamento e teste caracterizadas por forte variabilidade estocástica.

Por fim, abordagens pautadas integralmente em Aprendizado Profundo (*Deep Learning*), como a arquitetura proposta por [Machlev et al. \(2021\)](#), alcançam resultados quase perfeitos delegando a extração de atributos diretamente aos dados, através de Redes Neurais Convolucionais (CNN) combinadas com redes bidirecionais de memória curta (Bi-LSTM). Embora notável, o elevado desempenho destas arquiteturas deve ser interpretado à luz da síntese de dados. O processo de geração de dados descrito pelos autores apresenta características distintas da metodologia adotada neste trabalho, particularmente quanto à modelagem dos distúrbios e à incorporação de ruído.

Os resultados obtidos indicam que a combinação entre WPT e uma MLP convencional é capaz de atingir elevado desempenho mesmo em um cenário caracterizado por ampla variabilidade paramétrica e múltiplos níveis de ruído. Tal resultado sugere que uma etapa robusta de extração de características pode reduzir a necessidade de arquiteturas profundas mais complexas, mantendo níveis competitivos de desempenho para aplicações de monitoramento da QEE.

5 Conclusão

Este trabalho apresentou o desenvolvimento e a avaliação empírica de um *pipeline* computacional voltado à classificação de 16 tipos de Distúrbios de Qualidade de Energia Elétrica (PQD). O objetivo central da pesquisa consistiu em explorar a capacidade de extração de características da Transformada *Wavelet Packet* (WPT) e responder a uma questão fundamental pertinente à topologia dos dados: a aplicação da WPT é suficiente para linearizar o problema de diagnóstico de falhas elétricas sob regimes de ruído severo e variações estocásticas?

Os resultados obtidos indicam de forma consistente que a WPT não lineariza completamente o espaço de características. A análise comparativa revelou uma estagnação no desempenho do modelo linear de referência (LDA), que atingiu 85,39% de acurácia, contrastando fortemente com os modelos não lineares, como o *Random Forest* (94,44%) e, principalmente, a Rede Neural MLP, que se consolidou como o melhor classificador do *benchmark* ao atingir 95,21% de acurácia global e 0,95 de *Macro-F1*. Em cenários realistas, onde distúrbios compostos (como *Flicker_swell* ou *Sag_harmonics*) coexistem com altos níveis de AWGN, as fronteiras de decisão permanecem topologicamente emaranhadas, exigindo arquiteturas baseadas em Inteligência Artificial para mapear as complexas correlações não lineares.

Apesar de não linearizar o espaço, a principal contribuição metodológica deste trabalho foi demonstrar a eficácia da WPT na compactação de informações discriminantes, especificamente através da aplicação de métricas de entropia. A análise baseada na Razão de Fisher indicou que as métricas entrópicas figuram entre os atributos mais discriminantes do conjunto de características. Enquanto a amplitude do sinal é severamente mascarada pelo AWGN, a assinatura de "desorganização" e a complexidade energética extraídas nas sub-bandas da Wavelet permanecem estáveis, atuando como o motor fundamental para a alta taxa de revocação do sistema.

Ao posicionar estes achados frente à literatura contemporânea da área, a abordagem proposta mostra-se altamente competitiva. Embora trabalhos do estado da arte frequentemente reportem acurácias globais entre 95% e 99% (SINGH; PAL; DAHIYA, 2023; MISHRA, 2019; WANG; XU; CHE, 2019; MACHLEV et al., 2021; SIPAI; JADEJA; TANK, 2024), grande parte dessas pesquisas fundamenta-se em bases de dados com menor diversidade de classes, cenários de ruído brandos ou no uso de arquiteturas de Aprendizado Profundo (*Deep Learning*) de extremo custo computacional. A capacidade de sustentar uma acurácia de 95,21% utilizando uma rede clássica (MLP) sobre uma base de 160.000 amostras estocásticas sugere que a extração otimizada pela WPT oferece um

equilíbrio saudável entre precisão preditiva e eficiência algorítmica.

Como limitações da presente pesquisa, destaca-se a utilização exclusiva de dados sintéticos, gerados a partir das formulações paramétricas da norma [IEEE \(2019\)](#). Adicionalmente, a análise do domínio tempo-frequência ficou restrita a uma única família wavelet (Daubechies de ordem 8, com decomposição de nível 4), e o sistema não foi submetido a uma etapa de validação com sinais elétricos reais provenientes de oscilógrafos em campo. O reconhecimento destas restrições reflete a compreensão de que os algoritmos, embora altamente generalistas no ambiente simulado, ainda operam sobre representações idealizadas da rede de distribuição.

Essas limitações naturais, contudo, abrem um vasto escopo para desenvolvimentos futuros, os quais já estão previstos para o **SciML**, programa de pesquisa mais amplo conduzido pelo grupo. Como desdobramentos diretos deste TCC, propõe-se: a substituição das arquiteturas clássicas por modelos de atenção voltados a dados tabulares (*FT-Transformer*); a avaliação de generalização *cross-noise* (treinar o modelo em um nível de SNR e testar cegamente em outro); a realização de estudos comparativos de escala com diferentes *wavelets-mãe*; e, fundamentalmente, a aplicação de técnicas de adaptação de domínio para transferir o aprendizado do banco de dados sintético para a classificação em tempo real de dados experimentais.

Referências

- ANEEL. *Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional — PRODIST, Módulo 8: Qualidade da Energia Elétrica*. [S.l.]: Agência Nacional de Energia Elétrica, 2021. Revisão 12.
- BANDT, C.; POMPE, B. Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, v. 88, n. 17, p. 174102, 2002.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer, 2006.
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.
- CAICEDO, J. E. et al. A systematic review of real-time detection and classification of power quality disturbances. *Protection and Control of Modern Power Systems*, v. 8, n. 1, p. 1–37, 2023.
- CHEN, T.; GUESTRIN, C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S.l.: s.n.], 2016. p. 785–794.
- DAUBECHIES, I. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992. Disponível em: <<https://epubs.siam.org/doi/abs/10.1137/1-9781611970104>><https://epubs.siam.org/doi/abs/10.1137/1.9781611970104>.
- ERISTI, B.; YILDIRIM, O.; ERISTI, H. A new embedded power quality event classification system based on the wavelet transform. *International Journal of Electrical Power & Energy Systems*, v. 118, p. 105760, 2020.
- FERREIRA, D. D.; SEIXAS, J. M. de; CERQUEIRA, A. S. A method based on independent component analysis for single and multiple power quality disturbance classification. *Electric Power Systems Research*, v. 119, p. 425–431, 2015. ISSN 0378-7796. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S037877961400399X>><https://www.sciencedirect.com/science/article/pii/S037877961400399X>.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, v. 7, n. 2, p. 179–188, 1936.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.
- GORISHNIY, Y. et al. Revisiting deep learning models for tabular data. In: *Advances in Neural Information Processing Systems (NeurIPS)*. [S.l.: s.n.], 2021. v. 34.
- GRINSZTAJN, L.; OYALLON, E.; VAROQUAUX, G. Why do tree-based models still outperform deep learning on typical tabular data? In: *Advances in Neural Information Processing Systems (NeurIPS)*. [S.l.: s.n.], 2022. v. 35.
- GUO, T. et al. A review of wavelet analysis and its applications: Challenges and opportunities. *IEEE Access*, v. 10, p. 58869–58903, 2022.

- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009. (Springer series in statistics). ISBN 9780387848846. Disponível em: <<https://books.google.com.br/books?id=eBSgoAEACAAJ>><https://books.google.com.br/books?id=eBSgoAEACAAJ>.
- HAYKIN, S. *Neural Networks and Learning Machines*. 3. ed. [S.l.]: Pearson, 2009.
- IEEE. *IEEE Recommended Practice for Monitoring Electric Power Quality*. [S.l.]: IEEE, 2019. IEEE Std 1159-2019.
- IGUAL, R.; MEDRANO, C. Research challenges in real-time classification of power quality disturbances applicable to microgrids: A systematic review. *Renewable and Sustainable Energy Reviews*, v. 132, p. 110050, 2020.
- KHOKHAR, S. et al. A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances. *Renewable and Sustainable Energy Reviews*, v. 51, p. 1650–1663, 2015.
- KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Disponível em: <<https://arxiv.org/abs/1412.6980>><https://arxiv.org/abs/1412.6980>.
- MACHLEV, R. et al. Open source dataset generator for power quality disturbances with deep-learning reference classifiers. *Electric Power Systems Research*, v. 195, p. 107152, 2021.
- MADGULA, S.; VEERAMSETTY, V.; DURGAM, R. Signal processing approaches for power quality disturbance classification: A comprehensive review. *Results in Engineering*, v. 25, p. 104569, 2025.
- MAHELA, O. P.; SHAIK, A. G.; GUPTA, N. A critical review of detection and classification of power quality events. *Renewable and Sustainable Energy Reviews*, v. 41, p. 495–505, 2015.
- MALLAT, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, n. 7, p. 674–693, 1989.
- MALLAT, S. *A Wavelet Tour of Signal Processing: The Sparse Way*. 3. ed. [S.l.]: Academic Press, 2009.
- MCGRANAGHAN, M. et al. *Electrical Power Systems Quality, Third Edition*. McGraw-Hill Education, 2012. ISBN 9780071761550. Disponível em: <<https://books.google.com.br/books?id=69QjygAACAAJ>><https://books.google.com.br/books?id=69QjygAACAAJ>.
- MENDEL, J. Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *Proceedings of the IEEE*, v. 49, p. 278 – 305, 04 1991.
- MISHRA, M. Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review. *International Transactions on Electrical Energy Systems*, v. 29, n. 8, p. e12008, 2019.

- OLIVEIRA, R. A. S. de; BOLLEN, M. H. J. Deep learning for power quality. *Electric Power Systems Research*, v. 214, p. 108887, 2024.
- OPPENHEIM, A.; SCHAFER, R. *Discrete-time Signal Processing*. Pearson, 2010. (Prentice-Hall signal processing series). ISBN 9780131988422. Disponível em: <<https://books.google.com.br/books?id=mYsoAQAAMAAJ>><https://books.google.com.br/books?id=mYsoAQAAMAAJ>.
- PAPOULIS, A. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1991. (Communications and signal processing). ISBN 9780070484771. Disponível em: <<https://books.google.com.br/books?id=4IwQAQAIAAJ>><https://books.google.com.br/books?id=4IwQAQAIAAJ>.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.
- RÉNYI, A. On measures of entropy and information. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. [S.l.]: University of California Press, 1961. v. 1, p. 547–561.
- ROBERTSON, D. et al. Wavelets and electromagnetic power system transients. *IEEE Transactions on Power Delivery*, v. 11, n. 2, p. 1050–1058, 1996.
- RODRIGUEZ, M. A. et al. Classification of power quality disturbances using hilbert huang transform and a multilayer perceptron neural network model. *International Journal of Electrical Power & Energy Systems*, v. 121, p. 106105, 2020.
- ROSSO, O. A. et al. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *Journal of Neuroscience Methods*, v. 105, n. 1, p. 65–75, 2001.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, v. 323, n. 6088, p. 533–536, Oct 1986. ISSN 1476-4687. Disponível em: <<https://doi.org/10.1038/323533a0>><https://doi.org/10.1038/323533a0>.
- SANTOSO, S. et al. Power quality assessment via wavelet transform analysis. *IEEE Transactions on Power Delivery*, v. 11, n. 2, p. 924–930, 1996.
- SHANNON, C. Communication in the presence of noise. *Proceedings of the IRE*, v. 37, n. 1, p. 10–21, 1949.
- SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, v. 27, n. 3, p. 379–423, 1948.
- SINGH, G.; PAL, Y.; DAHIYA, A. K. Classification of power quality disturbances using linear discriminant analysis. *Applied Soft Computing*, v. 138, p. 110181, 2023.
- SIPAI, U.; JADEJA, R.; TANK, T. Power quality disturbance classification using transfer learning with deep convolutional neural networks. *International Journal of Electrical Power & Energy Systems*, v. 155, p. 109535, 2024.
- VARANIS, M.; PEDERIVA, R. Statements on wavelet packet energy–entropy signatures and filter influence in fault diagnosis of induction motor in non-stationary operations. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, v. 40, n. 2, p. 98, 2018.

VARANIS, M. et al. A tutorial review on time-frequency analysis of non-stationary vibration signals with nonlinear dynamics applications. *Brazilian Journal of Physics*, v. 51, n. 3, p. 859–877, 2021.

WANG, J.; XU, Z.; CHE, Y. Power quality disturbance classification based on dwt and multilayer perceptron extreme learning machine. *Applied Sciences*, v. 9, n. 11, 2019. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/9/11/2315>><https://www.mdpi.com/2076-3417/9/11/2315>.

Apêndices

APÊNDICE A – Figuras Complementares

Este apêndice apresenta representações visuais complementares do conjunto de dados sintético gerado, detalhando o comportamento das formas de onda no domínio do tempo sob o efeito das variações paramétricas.

A Figura 11 exibe um painel completo com as 16 classes de Distúrbios de Qualidade de Energia Elétrica (PQD) abordadas neste trabalho. Diferentemente de uma representação ideal unitária, cada subgráfico sobrepõe múltiplas instâncias geradas aleatoriamente para a mesma classe.

O objetivo desta visualização é atestar graficamente a aplicação da variação estocástica descrita na Metodologia. Esta representação corrobora o fato de que os algoritmos de aprendizado de máquina foram treinados para generalizar o fenômeno físico, e não apenas para memorizar formas de onda estáticas.

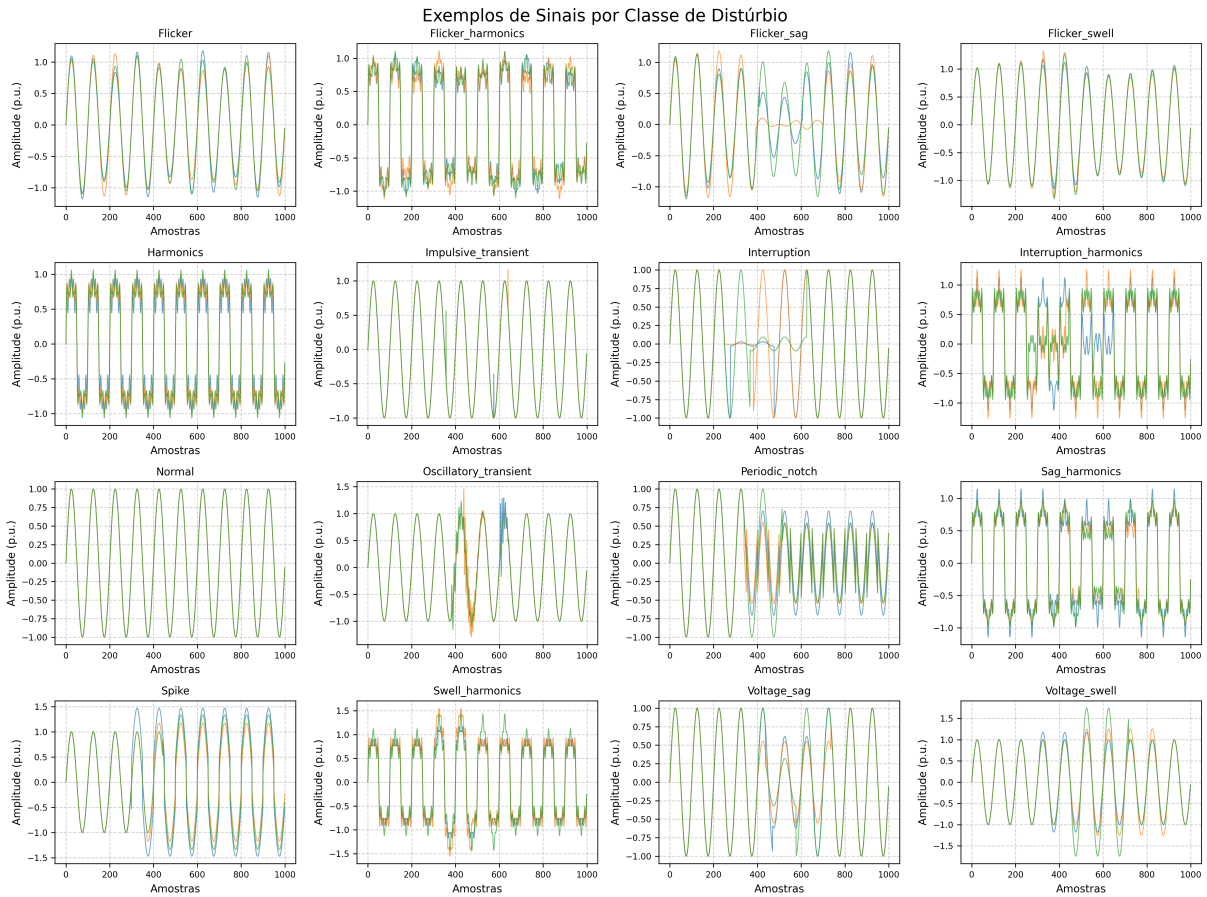


Figura 11 – Exemplos de sinais no domínio do tempo por classe de distúrbio, ilustrando a sobreposição estocástica de amplitude, temporização e ruído AWGN.

Fonte: Autoria própria.

Para aprofundar a compreensão sobre os mecanismos de decisão dos classificadores propostos, realizou-se uma avaliação da topologia do espaço de características gerado pela Transformada Wavelet Packet (WPT), utilizando técnicas de redução de dimensionalidade e a métrica de pontuação de Fisher (*Fisher Score*).

A visualização do espaço de atributos projetado bidimensionalmente corrobora a análise de desempenho do classificador linear (LDA) discutida anteriormente. Conforme ilustrado na projeção PCA (Figura 12, à esquerda), a transformação ortogonal linear das 208 dimensões resulta em um espaço altamente sobreposto, retendo apenas 46,7% da variância explicada nas duas componentes principais. Este alto grau de sobreposição central prova, geometricamente, que os atributos extraídos não são linearmente separáveis sob a influência de variação paramétrica e ruído estocástico AWGN.

Em contraste, a projeção t-SNE (Figura 12, à direita) — uma técnica de preservação não linear de vizinhanças estocásticas — revela a formação de agrupamentos (*clusters*) muito mais isolados e densos. Esta separabilidade latente, que só se manifesta sob transformações não lineares do espaço, justifica matematicamente a superioridade de acurácia apresentada pela Rede Neural MLP e pelo comitê Random Forest.

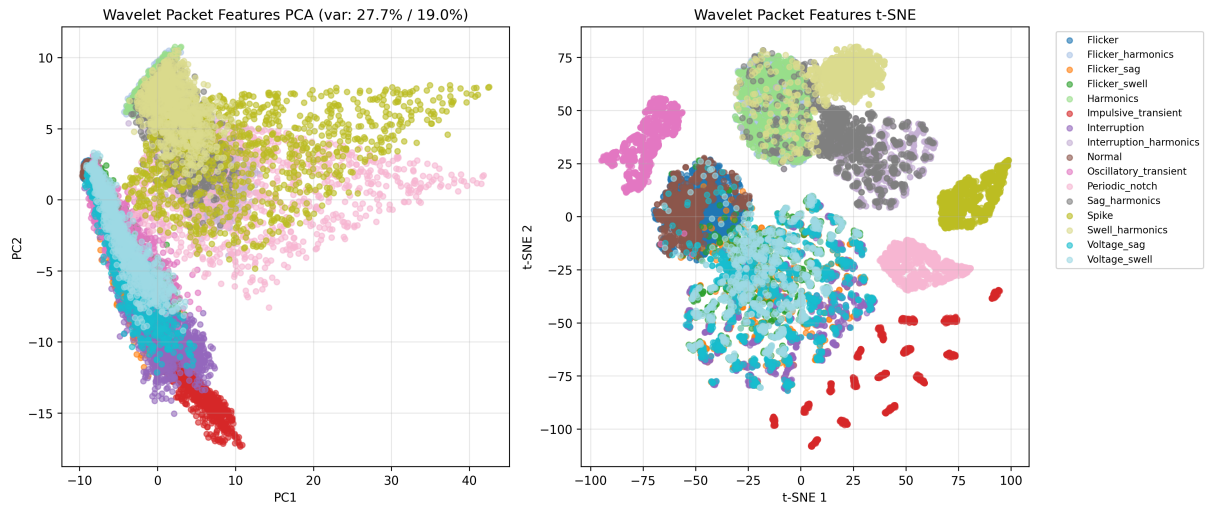


Figura 12 – Projeções espaciais do vetor de características WPT utilizando PCA (linear) e t-SNE (não linear).

Fonte: Autoria própria.

Adicionalmente, investigou-se a contribuição individual de cada um dos 208 atributos para o poder discriminativo global através da Razão de Fisher (Figura 13), que quantifica a razão entre a variância inter-classes e a variância intra-classes.

O ranqueamento estabeleceu o domínio absoluto das formulações entrópicas. As 15 características de maior poder diagnóstico consistem quase integralmente em variações de Entropia (*Log-Energy*, Permutação, Shannon e Rényi) extraídas de nós específicos da árvore Wavelet (como as sub-bandas *aaad* e *aadd*). Em oposição, atributos estatísticos básicos no domínio temporal, como valor eficaz (RMS), desvio padrão e máximo absoluto, apresentaram F-scores residuais. Esta evidência empírica responde de forma contundente à hipótese da pesquisa: enquanto o ruído AWGN macula as referências puras de amplitude e energia, a complexidade estrutural e a desorganização da informação dentro de bandas de alta frequência (capturadas pela Entropia) preservam a assinatura intrínseca do distúrbio de Qualidade de Energia, consolidando-se como o motor fundamental da alta precisão do sistema.

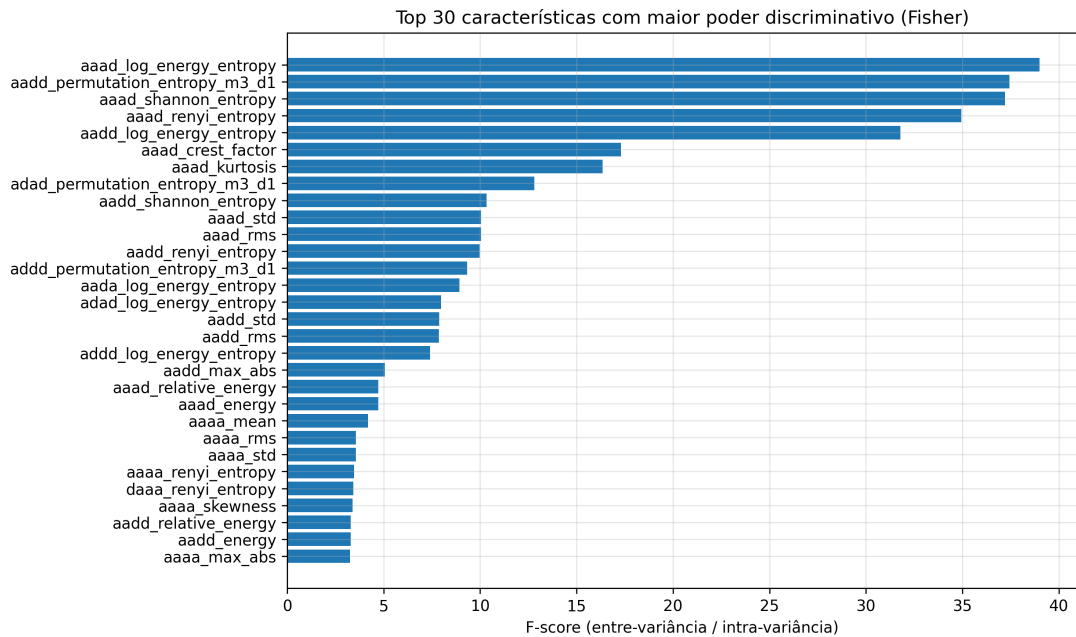


Figura 13 – Top 30 características de maior poder discriminativo calculadas através do *Fisher Score*.

Fonte: Autoria própria.

Para o classificador de referência linear (LDA), observa-se uma dispersão significativa das predições fora da diagonal principal da matriz de confusão, resultados ilustrados na Figura 14. Estes resultados indicam limitações na separação das classes no espaço de atributos construído a partir da WPT. As menores taxas de revocação foram observadas para as classes *Voltage-swell* (57%), *Flicker-swell* (63%) e *Sag-harmonics* (67%), evidenciando maior dificuldade na distinção entre fenômenos que compartilham características relacionadas à variação de amplitude e à presença de componentes harmônicas. Esses resultados sugerem que as fronteiras de decisão lineares impostas pelo LDA não são suficientes para representar integralmente a estrutura dos dados extraídos, especialmente diante da variabilidade paramétrica e da presença de ruído introduzidas no banco de dados. Como consequência, observa-se uma distribuição mais dispersa dos erros de classificação e uma acurácia global limitada a 85,39

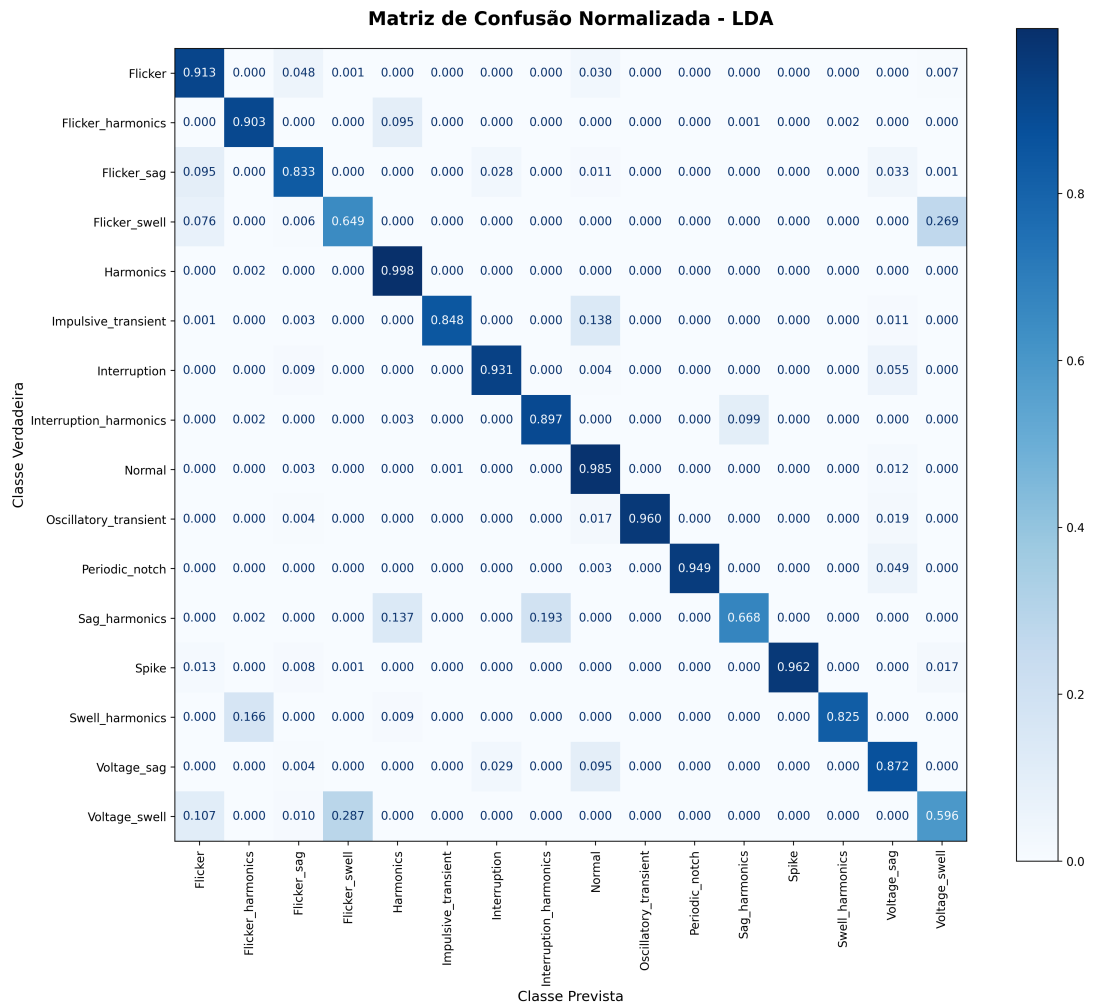


Figura 14 – Matriz de Confusão Normalizada para o classificador de Análise Discriminante Linear (LDA).

Fonte: Autoria própria.

Em contraste, o classificador Random Forest apresenta uma distribuição de erros significativamente mais concentrada, ilustrada pela Figura ???. Essa distribuição de erros reflete sua capacidade de modelar relações não lineares entre os atributos. Com acurácia global de 94,44%, o modelo preserva elevadas taxas de revocação para a maior parte das classes analisadas. As confusões remanescentes concentram-se principalmente entre categorias com forte similaridade física, destacando-se os pares *Voltage-swell* (81%) e *Flicker-swell* (83,3%), além da classe *Sag-harmonics* (81,4%). Ainda que essas categorias permaneçam como as mais desafiadoras do conjunto de dados, a redução expressiva dos erros em comparação ao LDA evidencia o benefício da utilização de modelos capazes de construir fronteiras de decisão não lineares.

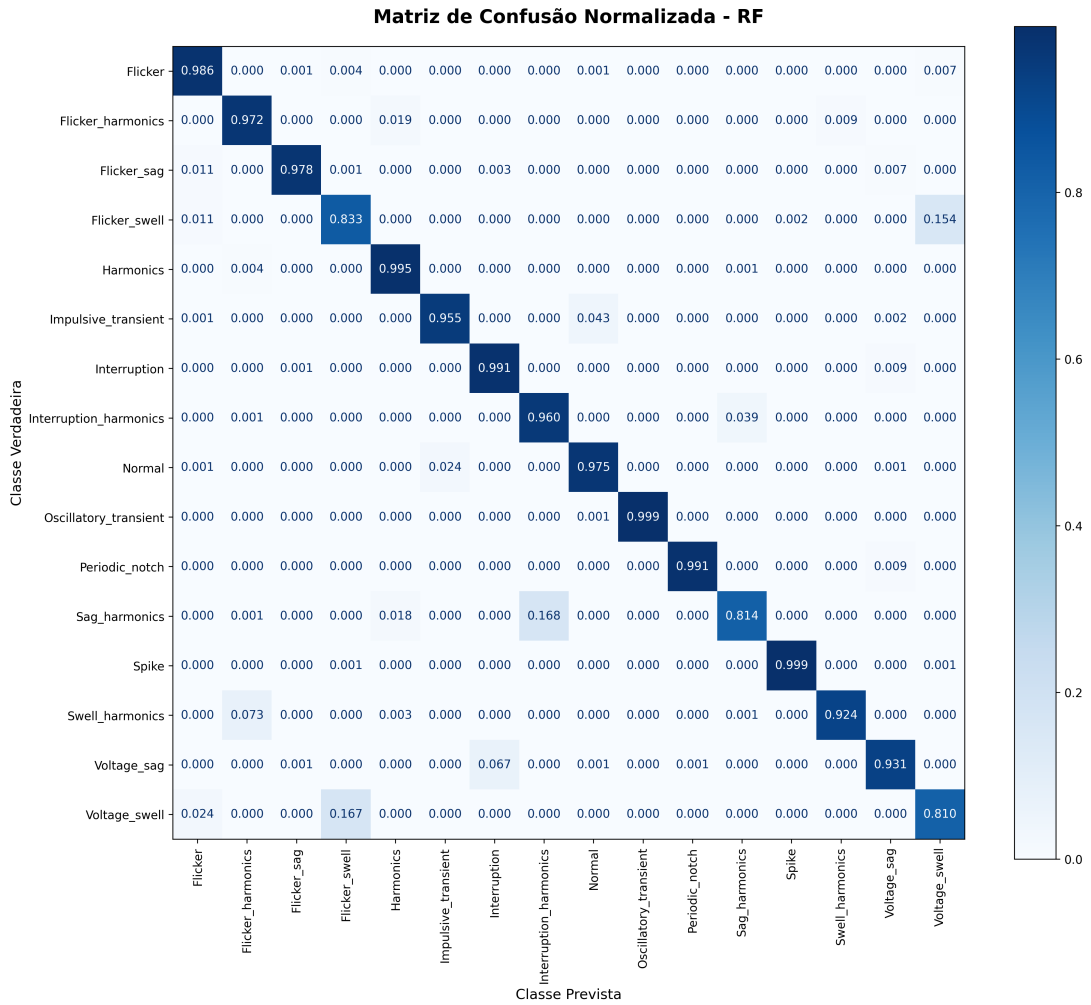


Figura 15 – Matriz de Confusão Normalizada para o classificador Random Forest (RF).

Fonte: Autoria própria.

Em conjunto, as matrizes de confusão corroboram os resultados quantitativos apresentados anteriormente. Enquanto o LDA apresenta dificuldades na separação de classes com padrões semelhantes, os modelos não lineares demonstram maior capacidade de explorar a informação discriminativa contida nos atributos extraídos pela WPT. Esses resultados são consistentes com a hipótese de que o problema de classificação de distúrbios de qualidade de energia não é completamente separável por fronteiras lineares, beneficiando-se de abordagens capazes de modelar relações mais complexas entre as características extraídas dos sinais.

APÊNDICE B – Física Associada aos Distúrbios de Qualidade de Energia Elétrica

Um classificador de distúrbios de qualidade de energia não opera no vácuo. Cada classe do nosso problema corresponde a um fenômeno físico real, com causas específicas, mecanismos distintos e impactos mensuráveis sobre equipamentos e processos. Quando os modelos confundem V-SWL com FLK-SWL, isso não é um “erro do algoritmo” — é um reflexo do fato de que esses dois fenômenos compartilham características físicas reais. Este apêndice cobre as 16 classes do projeto, organizadas em distúrbios simples (Seção B.1) e combinados (Seção B.2).

B.1 Distúrbios Simples

B.1.1 C01 — Normal (NOR)

O que é: O sinal normal é a referência: uma senoide pura a 50 Hz (ou 60 Hz), com amplitude nominal $A = 1,0$ p.u., sem distorções, modulações ou transitórios. Na prática, nenhum sistema elétrico real é perfeitamente “normal” — sempre há níveis residuais de distorção harmônica e ruído. O sinal NOR representa a condição ideal de referência.

Modelo:

$$v(t) = A \sin(\omega t + \varphi), \quad \omega = 2\pi f_0 = 2\pi \times 50$$

Hz

(B.1)

Relevância para Classificação: O sinal normal é a classe mais “fácil” conceitualmente, mas na prática é uma das mais problemáticas. Qualquer distúrbio sutil (notch

estreito, impulso fraco, harmônico de baixa amplitude) pode ser confundido com NOR sob ruído elevado. Nos resultados obtidos, NOR apresentou recall de 0,962 — 3,6% das instâncias foram confundidas principalmente com Impulsive Transient.

B.1.2 C02 — Afundamento de Tensão (V-SAG)

O que é: Uma redução momentânea da amplitude da tensão para valores entre 0,1 e 0,9 p.u., com duração típica de 0,5 ciclo a 1 minuto. É o distúrbio mais frequente em sistemas de distribuição.

Causas:

- **Curto-circuitos em alimentadores adjacentes:** Um curto-circuito fase-terra em um alimentador vizinho provoca queda de tensão no barramento comum. O sag dura até o religador ou disjuntor isolar a falta.
- **Partida de motores de indução:** Motores de grande porte (compressores, bombas, elevadores) drenam 5–8 vezes a corrente nominal durante a partida, causando queda de tensão no ponto de acoplamento.
- **Energização de transformadores:** O *inrush current* pode atingir 8–12 vezes a corrente nominal, com decaimento em 0,5–2 segundos.
- **Faltas em linhas de transmissão:** Curtos distantes propagam-se como sags atenuados ao longo da rede.

Mecanismo físico: A queda de tensão ocorre porque a impedância do sistema não é nula. Quando uma corrente elevada é drenada (por falta ou partida de motor), a queda de tensão na impedância da fonte (Z_s) reduz a tensão no ponto de entrega:

$$V_{\text{carga}} = V_{\text{fonte}} - Z_s \cdot I_{\text{falta}} \quad (\text{B.2})$$

Quanto maior Z_s (sistema “fraco”) ou I_{falta} , mais profundo o sag.

Impactos:

- Desligamento de contadores e relés (tensão abaixo do limiar de sustentação)
- Falha em fontes chaveadas de computadores e CLPs
- Desarme de variadores de frequência (VFDs) com proteção de subtensão
- Flicker visível em iluminação
- Perda de sincronismo em motores síncronos

Modelo:

$$v(t) = A [1 - \alpha (u(t - t_1) - u(t - t_2))] \sin(\omega t + \varphi), \quad \alpha \in [0,1; 0,9] \quad (\text{B.3})$$

A profundidade α define a severidade: $\alpha = 0,1$ é um sag de 10% (tensão cai para 0,9 p.u.); $\alpha = 0,9$ é quase uma interrupção (tensão a 0,1 p.u.).

B.1.3 C03 — Elevação de Tensão (V-SWL)

O que é: Um aumento momentâneo da amplitude para valores entre 1,1 e 1,8 p.u., com duração de 0,5 ciclo a 1 minuto. Menos frequente que o sag, mas potencialmente mais destrutivo.

Causas:

- **Desligamento súbito de carga pesada:** Quando uma carga de grande porte é desconectada, a tensão no barramento sobe até o regulador compensar.
- **Faltas fase-terra em sistemas não-aterrados:** As fases sãs podem experimentar elevações de até $\sqrt{3} \approx 1,73$ p.u.
- **Chaveamento de bancos de capacitores:** Energização rápida eleva a tensão temporariamente.
- **Erros de tap em transformadores com comutação automática.**

Impactos: Degradação acelerada de isolamento, falha de varistores (MOVs), queima de fontes de alimentação, desmagnetização de núcleos de transformadores.

Modelo:

$$v(t) = A [1 + \alpha (u(t - t_1) - u(t - t_2))] \sin(\omega t + \varphi), \quad \alpha \in [0,1;0,8] \quad (\text{B.4})$$

B.1.4 C04 — Interrupção (INT)

O que é: Redução da tensão para menos de 0,1 p.u. (essencialmente zero), com duração de 0,5 ciclo a 1 minuto (interrupção momentânea) ou superior a 1 minuto (sustentada).

Causa:

- **Atuação de dispositivos de proteção:** Disjuntores, fusíveis e religadores isolam faltas permanentes ou temporárias.
- **Falha em equipamentos de manobra:** Chaves seccionadoras, disjuntores com defeito.
- **Interrupções programadas:** Manutenção preventiva da rede.
- **Descargas atmosféricas:** Raios em linhas de distribuição causam faltas que disparam a proteção.

Modelo:

$$v(t) = A [1 - \alpha (u(t - t_1) - u(t - t_2))] \sin(\omega t + \varphi), \quad \alpha \in [0,9; 1,0] \quad (\text{B.5})$$

Note que a interrupção usa o **mesmo modelo** do sag — a diferença é apenas a faixa de α . Fisicamente, a interrupção é o caso extremo do sag.

B.1.5 C05 — Harmônicos (HAR)

O que é: Presença de componentes senoidais em frequências múltiplas inteiras da fundamental (150, 250, 350 Hz para 3^a, 5^a, 7^a harmônicos a 50 Hz). O sinal não é mais senoidal — a forma de onda é distorcida permanentemente.

Causas:

- **Cargas não-lineares:** Retificadores (fontes chaveadas, computadores, carregadores), variadores de frequência (VFDs), fornos a arco, iluminação fluorescente e LED.
- **Transformação magnética sob saturação:** Transformadores operando próximos à saturação geram harmônicos ímpares (principalmente 3^a e 5^a).
- **Conversores eletrônicos de potência:** Pontes retificadoras de 6 pulsos geram harmônicos de ordem $6k \pm 1$ ($k = 1, 2, \dots$): 5^a, 7^a, 11^a, 13^a...

Mecanismo físico: Cargas não-lineares drenam corrente em pulsos (não senoidais). Essa corrente não-senoidal, ao fluir pelas impedâncias do sistema, produz quedas de tensão não-senoidais que distorcem a forma de onda de tensão no ponto de acoplamento comum (PAC). A Distorção Harmônica Total (THD) quantifica essa distorção:

$$\text{THD} = \frac{\sqrt{\sum_{n=2}^{\infty} V_n^2}}{V_1} \times 100\% \quad (\text{B.6})$$

Impactos: Sobreaquecimento de transformadores e motores (perdas adicionais no cobre e no ferro), falha em capacitores (ressonância harmônica), erro em medidores de energia, interferência em sistemas de comunicação, aumento do condutor neutro em sistemas trifásicos (harmônicos de sequência zero).

Modelo:

$$v(t) = A \left[\sin(\omega t + \varphi) + \sum_{n \in \{3,5,7\}} \alpha_n \sin(n\omega t + \varphi_n) \right], \quad \alpha_n \in [0,05; 0,15] \quad (\text{B.7})$$

Modelo **aditivo:** harmônicos somam-se à fundamental. Cada harmônico tem fase independente φ_n .

B.1.6 C06 — Flicker (FLK)

O que é: Flutuação periódica da amplitude da tensão, tipicamente na faixa de 5–20 Hz, que causa variação perceptível no brilho de lâmpadas. O olho humano é mais sensível a flicker em torno de 8,8 Hz.

Causas:

- **Fornos a arco elétrico:** Variações aleatórias na impedância do arco causam oscilações de potência na faixa de flicker. É a causa clássica e mais severa.
- **Soldadores de resistência:** Operação intermitente com ciclo rápido.
- **Motores com carga variável:** Compressores alternativos, serras, britadores — cargas que variam ciclicamente geram flicker proporcional à variação de corrente.
- **Turbinas eólicas:** Passagem das pás pelo vento gera flutuações de potência a $3\times$ a frequência de rotação (típico 1–3 Hz em turbinas de 3 pás).

Mecanismo físico: O flicker é essencialmente uma **modulação de amplitude** (AM) do sinal de tensão. A envoltória do sinal oscila com a frequência do flicker f_f :

$$v(t) = A [1 + \alpha_f \sin(2\pi f_f t)] \sin(\omega t + \varphi) \quad (\text{B.8})$$

No domínio espectral, isso produz duas bandas laterais em $f_0 \pm f_f$ (e.g., $50 \pm 10 = 40$ e 60 Hz). Essas bandas estão muito próximas da fundamental, exigindo alta resolução espectral para detectá-las — exatamente onde a STFT com janela curta falha.

Impactos: Desconforto visual (lâmpadas), fadiga em ambientes industriais, problemas de qualidade em processos dependentes de tensão estável (semicondutores, lab. analíticos).

B.1.7 C07 — Transitório Oscilatório (OSC)

O que é: Oscilação amortecida de alta frequência (300–900 Hz, podendo chegar a kHz) sobreposta ao sinal fundamental, com duração de 0,5 a 3 ciclos.

Causas:

- **Chaveamento de bancos de capacitores:** Causa mais comum. A energização de um banco cria um circuito LC com a indutância do sistema, gerando oscilação amortecida. Frequência típica: $f_{tr} = 1/(2\pi\sqrt{LC})$, geralmente 300–800 Hz.
- **Eliminação de faltas:** A recuperação de tensão após abertura de disjuntor pode incluir anel oscilatório (TRV — Transient Recovery Voltage).

- **Descargas atmosféricas indiretas:** Raios que não atingem diretamente a linha, mas induzem tensões oscilatórias por acoplamento eletromagnético.

Modelo:

$$v(t) = A \sin(\omega t + \varphi) + \beta e^{-\gamma(t-t_1)} \sin(\omega_{tr} t + \varphi_{tr}) \cdot [u(t - t_1) - u(t - t_2)] \quad (\text{B.9})$$

O decaimento γ ($50\text{--}200\text{ s}^{-1}$) é inversamente proporcional ao fator de qualidade Q do circuito LC .

B.1.8 C08 — Transitório Impulsivo (IMP)

O que é: Alteração abrupta e unidirecional (não-oscilatória) da tensão, com tempo de subida muito rápido (nanossegundos a microssegundos) e duração de $T/20$ a $T/10$.

Causas:

- **Descargas atmosféricas diretas:** Raios atingindo a linha produzem impulsos de tensão com frente de onda íngreme ($1,2/50\mu\text{s}$ é a forma padrão de ensaio).
- **Chaveamento de cargas indutivas:** Abertura de circuitos com indução ($L dI/dt$) gera picos de tensão.
- **Descargas eletrostáticas (ESD)** em ambientes industriais.
- **Operação de fusíveis:** A fusão do elemento gera um impulso transitório.

Impactos: Perfuração de isolamento, queima de componentes eletrônicos sensíveis, corrupção de dados em memórias e processadores, degradação acumulativa de isolantes (envelhecimento elétrico).

Por quê IMP confunde com NOR O impulso é muito curto ($\sim 1\text{ ms}$ a 50 Hz) e de amplitude moderada ($0,1\text{--}0,4\text{ p.u.}$). Com ruído a 20 dB , a energia do impulso é comparável à do ruído em muitas sub-bandas. O classificador vê features estatísticas quase idênticas às do sinal normal ruidoso. A entropia de Rényi ($\alpha = 2$) deve ajudar: ela é sensível à concentração de energia em poucos coeficientes, que é a assinatura do impulso.

B.1.9 C09 — Notch Periódico (PN)

O que é: Entalhes (reduções abruptas) na forma de onda de tensão, ocorrendo periodicamente nos cruzamentos por zero ou próximos a eles, com largura de $T/40$ a $T/20$.

Causas:

- **Comutação natural de retificadores controlados (SCRs/tiristores):** Durante o período de comutação, duas fases do retificador ficam momentaneamente em curto-circuito através da impedância de comutação. Isso cria um “entalhe” na tensão de linha.
- **Conversores AC-DC de 6 e 12 pulsos:** Geram 6 ou 12 notches por ciclo, respectivamente.

Mecanismo físico: O notch ocorre porque, durante a comutação do SCR, a tensão de linha colapsa brevemente enquanto a corrente se transfere de um tiristor para o próximo. A largura do notch é proporcional à indutância de comutação: quanto maior L_c , mais largo o notch.

Por quê PN confunde com NOR: O notch é extremamente estreito (0,5–1

ms) e ocorre nos cruzamentos por zero, onde a amplitude do sinal já é próxima de zero. Isso significa que a sequência temporal de coeficientes muda localmente, mesmo que a energia não mude.

B.1.10 C10 — Spike (SPK)

O que é: Adições abruptas e periódicas nos picos da forma de onda, simétricas em relação ao sinal fundamental. Pode-se pensar como o “oposto” do notch: em vez de remover, adiciona energia nos picos.

Causas:

- **Arcos elétricos periódicos:** Soldadores a arco, fornos a arco em operação contínua.
- **Chaveamento repetitivo de cargas indutivas:** Motores com comutação mecânica (escovas).
- **Corona e descargas parciais periódicas** em isoladores degradados.

B.2 Distúrbios Combinados

Na realidade, os distúrbios raramente ocorrem isoladamente. Um sistema alimentando um forno a arco (harmônicos + flicker) que sofre um curto-circuito em alimentador adjacente (sag) experimenta simultaneamente **sag + harmônicos + flicker**. Os modelos combinados capturam essa coexistência. Do ponto de vista de classificação, são as classes mais desafiadoras — porque as características de cada componente se sobrepõem.

B.2.1 C11–C13: Sag/Swell/Interrupção + Harmônicos

Cenário físico: Uma planta industrial com cargas não-lineares (retificadores, VFDs) gera distorção harmônica permanente. Quando ocorre um curto-circuito externo (sag) ou o desligamento de carga pesada (swell), o sinal resultante contém **simultaneamente** a variação de amplitude e os harmônicos.

O modelo é multiplicativo: o envelope (sag/swell/interrupção) modula o sinal distorcido:

$$v(t) = A [1 \pm \alpha (u_{t_1} - u_{t_2})] \cdot \left[\sin(\omega t + \varphi) + \sum \alpha_n \sin(n\omega t + \varphi_n) \right] \quad (\text{B.10})$$

B.2.2 C14: Flicker + Harmônicos

Cenário físico: Forno a arco elétrico. O arco é simultaneamente uma fonte de harmônicos (corrente não-linear) e de flicker (variação cíclica da impedância do arco). Ambos fenômenos coexistem na mesma forma de onda.

B.2.3 C15–C16: Flicker + Sag / Flicker + Swell

Cenário físico: Uma fábrica com fornos a arco (flicker permanente) sofre um evento externo (sag ou swell). A modulação do flicker se sobrepõe à variação de amplitude do sag/swell:

$$v(t) = A [1 + \alpha_f \sin(2\pi f_f t)] \cdot [1 \pm \alpha (u_{t_1} - u_{t_2})] \cdot \sin(\omega t + \varphi) \quad (\text{B.11})$$

B.3 Síntese: O Mapa Físico do Problema

Tabela 6 – Síntese das causas físicas e mecanismos dominantes por classe.

ID	Classe	Causa principal	Mecanismo	Assinatura no sinal
C01	NOR	—	—	Senóide pura
C02	V-SAG	Curto-circuito, partida motor	Queda em Z_s	Redução amplitude
C03	V-SWL	Rejeição de carga, falta f-t	Alívio de $Z_s \cdot I$	Aumento amplitude
C04	INT	Proteção, fusível	Isolamento total	Amplitude ≈ 0
C05	HAR	Cargas não-lineares	Corrente não-senoidal	Distorção permanente
C06	FLK	Forno a arco, carga variável	Modulação AM	Envoltória oscilante
C07	OSC	Chaveamento capacitores	Circuito LC	Oscilação amortecida HF
C08	IMP	Raio, chaveamento indutivo	$L \, dI/dt$	Pulso unidirecional
C09	PN	Comutação SCR/tiristor	Curto momentâneo	Entalhes periódicos
C10	SPK	Arco, descargas parciais	Adição periódica	Picos nos máximos