



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE ENGENHARIAS
CURSO DE BACHARELADO EM ENGENHARIA ELÉTRICA

LINO ALESSANDRO MORAIS ALVES

**CLASSIFICAÇÃO DE DISTÚRBIOS DE QUALIDADE DA ENERGIA UTILIZANDO
TÉCNICAS DE APRENDIZADO DE MÁQUINA**

CARAÚBAS – RN

2025

LINO ALESSANDRO MORAIS ALVES

**CLASSIFICAÇÃO DE DISTÚRBIOS DE QUALIDADE DA ENERGIA UTILIZANDO
TÉCNICAS DE APRENDIZADO DE MÁQUINA**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Engenharia Elétrica.

Orientadora: Érica Manguiera Lima, Profa. Dra.

Coorientadora: Naara Shiva de Oliveira Melo, B. ela.

CARAÚBAS – RN

2025

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

A474c Alves, Lino Alessandro Moraes .
Classificação de distúrbios de qualidade da
energia utilizando técnicas de aprendizado de
máquina / Lino Alessandro Moraes Alves. - 2025.
57 f. : il.

Orientador: Érica Manguiera Lima.
Coorientador: Naara Shiva de Oliveira Melo.
Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Engenharia
Elétrica, 2025.

1. Qualidade de energia elétrica. 2.
Aprendizado de máquina. 3. Distúrbios. 4.
Analisador. I. Lima, Érica Manguiera, orient. II.
Melo, Naara Shiva de Oliveira, co-orient. III.
Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo autor(a).
Biblioteca Campus Caraúbas
Bibliotecária: Sarah Freire Bezerra
CRB: 15/1052

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

LINO ALESSANDRO MORAIS ALVES

**CLASSIFICAÇÃO DE DISTÚRBIOS DE QUALIDADE DA ENERGIA UTILIZANDO
TÉCNICAS DE APRENDIZAGEM DE MÁQUINAS**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Engenharia Elétrica.

Defendida em: 27/03/2025.

BANCA EXAMINADORA

Documento assinado digitalmente
 **ERICA MANGUEIRA LIMA**
Data: 27/03/2025 10:55:43-0300
Verifique em <https://validar.iti.gov.br>

Érica Manguiera Lima, Profa. Dra.
Presidente

Documento assinado digitalmente
 **NAARA SHIVA DE OLIVEIRA MELO**
Data: 27/03/2025 16:32:24-0300
Verifique em <https://validar.iti.gov.br>

Naara Shiva de Oliveira Melo, Bac.
Membro Examinador

Documento assinado digitalmente
 **JAMILE PINHEIRO NASCIMENTO AMOAH**
Data: 27/03/2025 14:03:23-0300
Verifique em <https://validar.iti.gov.br>

Jamile Pinheiro Nascimento Amoah, Profa. Dra.
Membro Examinador

Documento assinado digitalmente
 **RODRIGO DE ALMEIDA COELHO**
Data: 27/03/2025 11:03:11-0300
Verifique em <https://validar.iti.gov.br>

Rodrigo de Almeida Coelho, Prof. Dr.
Membro Examinador

*Dedico a Deus, à minha noiva, Marília
Fernanda, aos meus avós, Vilmaci e João de
Deus, à minha mãe, Miria Micheli, e à minha
sogra, Gilvaneide, pelo amor, pela ajuda e por
acreditarem em mim.*

AGRADECIMENTOS

Dedico este momento de gratidão a todas as pessoas que, de alguma forma, me apoiaram e me ajudaram a realizar este sonho. Cada incentivo e gesto de apoio foram fundamentais para a conclusão deste trabalho.

A Deus, pela vida e pela força que me permitiram chegar até aqui. À minha esposa, Marília Fernanda Gurgel de Góis, minha companheira em todos os momentos, que me incentivou, auxiliou e, acima de tudo, acreditou em mim.

À minha mãe, Míria Micheli de Moraes, que esteve ao meu lado desde o meu primeiro dia de vida, me educou da melhor forma possível e sempre apoiou minhas escolhas. À minha avó materna, Maria Vilmaci de Moraes, e ao meu avô materno, João de Deus de Moraes, que, assim como minha mãe, sempre me deram todo o apoio e fizeram tudo o que estava ao seu alcance por mim.

Aos meus amigos Dulciel Marcos Abel da Silva, Erick Aragão Gadelha de Oliveira, Werly Renan Fernandes Lucena, Luiz Paulo Gomes da Silva, Adriel Batista Soares e Ednardo Édson Lucena e Lima, entre tantos outros que fizeram parte desta jornada.

Às minhas professoras e orientadoras, Érica Manguiera Lima e Naara Shiva de Oliveira Melo, pela orientação, conselhos e dedicação durante a elaboração deste trabalho.

Aos demais professores que contribuíram para a minha formação acadêmica, bem como aos servidores e técnicos que, direta ou indiretamente, me auxiliaram ao longo desta trajetória.

“É muito melhor lançar-se em busca de conquistas grandiosas, mesmo expondo-se ao fracasso, do que alinhar-se com os pobres de espírito, que nem gozam muito nem sofrem muito, porque vivem numa penumbra cinzenta, onde não conhecem nem vitória, nem derrota.”

(Theodore Roosevelt)

RESUMO

O aumento da demanda por energia elétrica e a crescente presença de dispositivos eletrônicos conectados à rede elétrica tornaram a qualidade da energia elétrica (QEE) um tema de grande relevância. Dispositivos eletrônicos, especialmente aqueles baseados em eletrônica de potência, introduzem distorções nos sinais de tensão e corrente, podendo causar diversos problemas, como superaquecimento de equipamentos, interferências em sistemas de comunicação e redução da vida útil dos dispositivos. A identificação precisa dos distúrbios associados a esses problemas permite a tomada de medidas proativas para mitigar impactos negativos, prevenindo danos a equipamentos e prejuízos financeiros. Diante desse contexto, este trabalho propõe o estudo dos principais distúrbios de QEE, a identificação de parâmetros de referência e o desenvolvimento de um classificador capaz de diferenciá-los com base em suas características mais relevantes. Para isso, foram utilizados conceitos de aprendizado de máquina, treinando o modelo com uma base de dados sintética fornecida pelo IEEE *Dataport* e com o auxílio do *software* MATLAB®. Os resultados evidenciaram a possibilidade de classificar diversos distúrbios, tanto isolados quanto combinados, utilizando árvores de decisão, com acurácia superior a 99%. Ademais, a análise realizada enfatizou que a escolha correta dos parâmetros identificadores é essencial para a eficácia do classificador.

Palavras-chave: qualidade da energia elétrica; aprendizado de máquinas; classificação de distúrbios.

ABSTRACT

The increasing demand for electricity and the growing presence of electronic devices connected to the electrical grid have made power quality a highly relevant topic. Electronic devices, especially those based on power electronics, introduce distortions in voltage and current signals, which can cause various problems, such as overheating of equipment, interference in communication systems, and reduced lifespan of devices. The precise identification of disturbances associated with these problems allows for proactive measures to mitigate negative impacts, preventing damage to equipment and financial losses. In light of this context, this work proposes the study of the main power quality disturbances, the identification of reference parameters, and the development of a classifier capable of differentiating them based on their most relevant characteristics. To achieve this, concepts of machine learning are used, training the model with a synthetic dataset provided by the IEEE Dataport and with the aid of MATLAB[®] software. The results show that it is possible to classify several isolated and combined disorders using decision trees, presenting an accuracy higher than 99%. Furthermore, the analysis performed emphasizes that the correct choice of identification parameters is essential for the effectiveness of the classifier.

Keywords: power quality; machine learning; disturbance classification.

LISTA DE FIGURAS

Figura 1: Distúrbios que alteram a QEE.....	18
Figura 2: Exemplo de árvore de decisão	26
Figura 3: Exemplo de RNA.....	26
Figura 4: Matriz confusão.....	28
Figura 6: Configuração do <i>Classification Learner</i>	33
Figura 7: (a) Sinal ideal, (b) Sinal com afundamento.	35
Figura 8: (a) Elevação de tensão, (b) interrupção de tensão.....	36
Figura 9: (a) Transitório impulsivo, (b) Transitório oscilatório.	36
Figura 10:(a) Flutuação de tensão; (b) Harmônicos	37
Figura 11: <i>Notching</i> ou <i>Spike</i>	38
Figura 12: Resultados para todas as técnicas analisadas com os 11 atributos.....	40
Figura 13: Matriz de confusão para árvore de decisão (<i>fine tree</i>).	41
Figura 14: Matriz de confusão para árvore de decisão (<i>Boosted Trees</i>).	41
Figura 15: Matriz de confusão para RNA média (<i>Medium Neural Network</i>).....	42
Figura 16: Resultados do algoritmo MRMR.	44
Figura 17: Matriz de confusão para árvore de decisão (Grupo A8).	45
Figura 18: Matriz de confusão para <i>boosted trees</i> (Grupo A8).	46
Figura 19: Matriz de confusão para RNA média (Grupo A8).	46
Figura 20: Matriz de confusão para o treino da <i>fine tree</i> (Grupo A5).	48
Figura 21: Matriz de confusão para o treino da <i>Boosted trees</i> (Grupo A5).	49
Figura 22: Matriz de confusão para o treino da RNA (Grupo A5).....	49
Figura 23: Comparativo de resultados para os três grupos de teste.....	51

LISTA DE TABELAS

Tabela 1: Características dos distúrbios de QEE.....	18
Tabela 2: Modelo matemático dos distúrbios.....	31
Tabela 3: Especificação dos sinais presentes na base de dados sintética.	32
Tabela 4: Classificação dos distúrbios na base de dados sintética.	32
Tabela 5: Parâmetros para simulação dos distúrbios por equações.....	34
Tabela 6: Atributos escolhidos e sua posição na matriz PQparam (Grupo A11).....	38
Tabela 7: Resultados para o treinamento com todos os parâmetros.....	42
Tabela 8: Resultados para os testes com todos os parâmetros.	43
Tabela 9: Resultados para o treinamento do Grupo A8.....	47
Tabela 10: Resultados para os testes do Grupo A8.	47
Tabela 11: Resultados para o treinamento e testes do Grupo A5.....	50

LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
FFT	<i>Fast Fourier Transform</i>
FN	Falsos Negativos
FP	Falsos Positivos
IEC	<i>International Electrotechnical Commission</i>
IEEE	Instituto de Engenheiros Eletricistas e Eletrônicos
IA	Inteligência Artificial
MRMR	Máxima relevância e mínima redundância
PRODIST	Procedimentos de Distribuição de Energia Elétrica no Sistema Eletrônico Nacional.
QEE	Qualidade da Energia Elétrica
RNA	Redes Neurais Artificiais
RAM	<i>Random Access Memory</i>
RMS	<i>Root Mean Square</i>
SVM	Máquinas de Vetores de Suporte
TCC	Trabalho de Conclusão de Curso
TF	Transformada de Fourier
UFERSA	Universidade Federal Rural do Semi-Árido
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS.....	15
1.2	ORGANIZAÇÃO DO TEXTO.....	15
2	REFERENCIAL TEÓRICO.....	17
2.1	QUALIDADE DA ENERGIA ELÉTRICA.....	17
2.1.1	Transitórios	18
2.1.2	Variações de tensão de curta duração	20
2.1.3	Variações de tensão de longa duração	20
2.1.4	Desequilíbrio de tensão	21
2.1.5	Distorções harmônicas	21
2.1.6	Flutuação de tensão	22
2.1.7	<i>Notching e Spike</i>	22
2.2	TRANSFORMADA DE FOURIER	23
2.3	APRENDIZADO DE MÁQUINA	24
2.3.1	Técnicas para validação e testes	27
2.3.2	Escolha dos atributos caracterizadores	28
3	METODOLOGIA	30
4	RESULTADOS E DISCUSSÕES	34
4.1	SELEÇÃO DOS ATRIBUTOS CLASSIFICADORES.....	34
4.2	CLASSIFICAÇÃO POR APRENDIZADO DE MÁQUINA.....	39
4.2.1	Redução dos parâmetros classificadores	43
4.2.1.1	Treinamento e testes para o Grupo A8	44
4.2.1.2	Treinamento e testes para o Grupo A5	47
4.2.2	Escolha do modelo de classificador	50
5	CONCLUSÕES	52
	REFERÊNCIAS BIBLIOGRÁFICAS	53

1 INTRODUÇÃO

O aumento gradual da demanda por energia elétrica, impulsionado principalmente pelos avanços tecnológicos das últimas décadas, trouxe desafios relevantes para os sistemas elétricos atuais. A constante diversificação na geração e nas cargas presentes nos sistemas representa uma pressão adicional para melhorias na geração, transmissão e distribuição de energia elétrica. De acordo com Glover, Sarma e Overbye (2016), o crescimento da demanda por energia elétrica fomenta a necessidade de adaptação contínua e periódica dos sistemas, visando garantir a estabilidade e eficiência.

De fato, o desenvolvimento de dispositivos baseados em eletrônica de potência, intrinsicamente relacionados com os avanços em diversos setores como tecnologia, indústria, infraestrutura e energias renováveis, representa uma preocupação ambivalente para o fornecimento de energia elétrica. Esse tipo de equipamento possui capacidade de alterar as características de amplitude e frequência de sinais senoidais, no entanto, tal atributo tem como ponto negativo a introdução de distorções nas formas de onda tanto da tensão quanto da corrente. Em geral, esses equipamentos são tratados como cargas não lineares, pois não há uma relação de proporcionalidade linear entre os sinais de entrada e saída.

Segundo Kusko e Thompson (2007), a presença de cargas não lineares pode causar diversos problemas como aumento das perdas, sobreaquecimento de equipamentos e condutores e funcionamento inadequado dos dispositivos eletrônicos. Nesse contexto, um dos principais desafios atuais é a manutenção da Qualidade da Energia Elétrica (QEE).

A QEE está relacionada à manutenção dos padrões de qualidade no fornecimento de tensão e corrente, em relação ao sinal senoidal puro, tanto em termos de amplitude quanto em relação à frequência (Casazza; Delea, 2010; Grigsby, 2007). Conforme a Agência Nacional de Energia Elétrica (ANEEL), a QEE é definida como um conjunto de características que devem ser seguidas pela concessionária com o intuito de garantir que o produto e o serviço, ou seja, o fornecimento de energia elétrica, esteja adequado para o consumo. Por meio do Módulo 8 dos Procedimentos de Distribuição de Energia Elétrica Sistema Elétrico Nacional (PRODIST, 2024), a ANEEL estabelece parâmetros para regulamentação e fiscalização da QEE.

As diversas variações que podem ocorrer nos parâmetros de QEE são chamadas de distúrbios de QEE. Esses distúrbios podem estar relacionados com variações da amplitude da tensão e corrente, desbalanceamento entre as tensões e correntes de fase, interrupções do fornecimento da energia, presença de distorções harmônicas e presença de ruído elétrico (Diniz, 2018). A compreensão desses distúrbios, tanto para identificá-los quanto para diferenciá-los, é essencial para a manutenção da QEE.

Nesse contexto, propõe-se, neste Trabalho de Conclusão de Curso (TCC) o estudo sobre os principais distúrbios de QEE a fim de desenvolver um classificador capaz de diferenciá-los a partir de suas características mais relevantes utilizando como base o conceito de aprendizado de máquina.

1.1 OBJETIVOS

Esse trabalho tem como objetivo geral implementar um classificador de distúrbios de QEE, utilizando os conceitos de aprendizado de máquina adotando uma base de dados sintética fornecida pelo Instituto de Engenheiros Elétricos e Eletrônicos (IEEE). Como objetivos específicos destacam-se:

- Fazer um estudo sobre os distúrbios de QEE nos domínios do tempo e da frequência, buscando identificar possíveis atributos caracterizadores capazes de auxiliar em sua identificação e classificação dos distúrbios;
- Realizar um pré-processamento da base de dados;
- Selecionar e testar diferentes algoritmos de aprendizado de máquina no *software* MATLAB® (árvores de decisão, florestas aleatórias, *support vector machines*, redes neurais artificiais, entre outros);
- Identificar as técnicas de aprendizagem mais adequadas e atributos diferenciadores para o classificador de distúrbios por meio da análise de matrizes de confusão.

1.2 ORGANIZAÇÃO DO TEXTO

Este trabalho está organizado em cinco capítulos, incluindo esta introdução. O capítulo dois aborda os conceitos fundamentais relacionados à QEE e à aprendizado de máquina. No capítulo três descreve-se a metodologia adotada para o desenvolvimento do trabalho, incluindo a análise dos distúrbios de QEE, a seleção dos atributos classificadores, o pré-processamento dos dados e a aplicação das técnicas de aprendizado

de máquina. Ademais, são detalhados os procedimentos para a validação e teste dos modelos.

No capítulo quatro são apresentados os resultados obtidos e as respectivas discussões, incluindo a análise das matrizes de confusão e da acurácia dos modelos. Além disso, são discutidas as implicações dos resultados e a eficácia do classificador proposto. Por fim, o capítulo cinco traz as conclusões do trabalho, destacando as contribuições para o campo da QEE e sugerindo possíveis direções para pesquisas futuras.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os principais conceitos concernentes à QEE, a aprendizagem de máquina e às técnicas utilizadas para a obtenção dos parâmetros classificadores.

2.1 QUALIDADE DA ENERGIA ELÉTRICA

A QEE pode ser definida de várias formas a depender do contexto. Na perspectiva dos fabricantes de equipamentos, por exemplo, ela está ligada às características do fornecimento de energia, pois a adequação da tensão fornecida permite que os equipamentos operem de maneira correta. Por outro lado, sob a ótica do consumidor, refere-se ao desempenho e a durabilidade do equipamento, ao funcionamento adequado e ao custo da energia. Para as concessionárias, o conceito envolve a confiabilidade do sistema, ou seja, mantê-la significa garantir que o sistema atenda às demandas do consumidor sem interrupções (Dugan *et al.*, 2003).

De acordo com Dugan *et al.* (2003), a QEE pode ser definida como a capacidade de um sistema elétrico de fornecer a energia de forma estável e confiável, assegurando que os parâmetros de tensão, corrente, frequência e forma de onda estejam dentro dos padrões adequados. Em outras palavras, a QEE estabelece um parâmetro que avalia a eficácia com que a energia elétrica é disponibilizada para utilização pelo consumidor final. Dessa forma, a manutenção da QEE está ligada à necessidade de minimizar distúrbios (Deckmann; Pomilio, 2024).

Para compreender os diversos distúrbios de QEE é necessário compreender o que é um sinal senoidal. Um sinal senoidal puro tem amplitude e período constante, conforme a Equação (1):

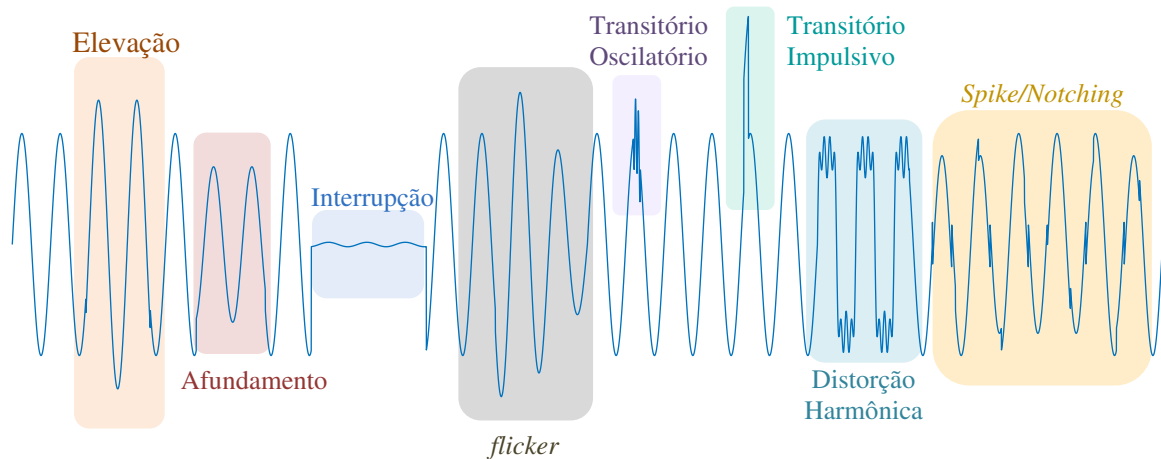
$$x(t) = A \sin(\omega t), \text{ ou } x(t) = A \cos(\omega t) \quad (1)$$

em que $x(t)$ é o sinal contínuo no tempo, t é o instante de tempo em segundos, ω é a frequência angular e A é a amplitude do sinal.

Segundo a norma do IEEE 1159-2019 (IEEE, 2019), os distúrbios de QEE podem ser classificados em sete categorias: transitórios, variações de tensão de curta duração, variações de tensão de longa duração, desequilíbrios, distorções na forma de onda, flutuações de tensão e variações de frequência. Os principais distúrbios estão

apresentados na Figura 1. Na Tabela 1, apresentam-se as principais características dos distúrbios definidos pela norma, bem como as subcategorias a eles associadas.

Figura 1: Distúrbios que alteram a QEE



Fonte: Autor (2025).

2.1.1 Transitórios

Os transitórios são perturbações de tensão ou de corrente em regime permanente com duração muito curta. Durante esse intervalo de tempo, o sistema elétrico pode sofrer variações bruscas dessas grandezas. Segundo Rocha (2016), eles podem ser classificados em:

- Impulsivo - mudança súbita na amplitude da tensão de regime permanente com natureza unidirecional na polaridade (sentido positivo ou negativo).
- Oscilatório - mudança súbita na amplitude de regime permanente da tensão, porém com mudanças na polaridade, provocando oscilações na forma de onda da tensão. Em geral, ocorre devido à energização de linhas, eliminação de faltas, chaveamento de banco de capacitores e transformadores.

Tabela 1: Características dos distúrbios de QEE

Categorias			Conteúdo Espectral	Duração	Amplitude
Transitórios	Impulsivos	Nanosegundo	-	< 50ns	
		Microsegundo	-	50 ns – 1 ms	

		Milisegundo	-	> 1 ms		
	Oscilatórios	Baixa Frequência	< 5 kHz	0,3 – 50 ms	0 – 4 pu	
		Média Frequência	5 – 500 kHz	20 μs	0 – 8 pu	
		Alta Frequência	0,5 – 5 MHz	5 μs	0 – 4 pu	
Variações de tensão de curta duração	Instantânea	Afundamento	-	0,5 – 30 ciclos	0,1 – 0,9 pu	
		Elevação	-	0,5 – 30 ciclos	1,1 – 1,8 pu	
	Momentânea	Interrupção	-	0,5 – 3 s	< 0,1 pu	
		Afundamento	-	30 ciclos – 3 s	0,1 – 0,9 pu	
		Elevação	-	30 ciclos – 3 s	1,1 – 1,4 pu	
	Temporária	Interrupção	-	3 s – 1 min	< 0,1 pu	
		Afundamento	-	3 s – 1 min	0,1 – 0,9 pu	
		Elevação	-	3 s – 1 min	1,1 – 1,2 pu	
	Variações de tensão de longa duração		Interrupção sustentada	-	> 1 min	0 pu
			Subtensão sustentada	-	> 1 min	0,8 – 0,9 pu
Sobretensão sustentada			-	> 1 min	1,1 – 1,2 pu	
Desequilíbrio de Tensão			-	Regime permanente	0,5 - 2%	
Distorções na forma de onda		Harmônicos	0 – 150º ordem H	Regime permanente	0 – 20%	
		Notching/Spike	-	Regime permanente		
Flutuação de tensão (<i>flicker</i>)			< 25 kHz	Intermitente	0,1 – 7%	
Variações de frequência			-	< 10 s	Não se aplica	

Fonte: IEEE 1159 (2019).

2.1.2 Variações de tensão de curta duração

As variações de tensão de curta duração são definidas por alterações na amplitude do sinal de tensão, que podem manifestar-se através de quedas, aumento ou ausência de tensão para níveis próximos de zero ou até mesmo zero (Macedo, 2017). Esses distúrbios podem ser classificados em:

- Afundamento - diminuição da tensão em torno de 10% a 90%, considerando a tensão base do sistema, com duração de 0,5 ciclo até 1 minuto. Esse fenômeno é frequentemente causado por eventos como faltas na rede, partidas de grandes motores ou a conexão de grandes cargas (Bollen *et al*, 2009).
- Elevação - aumento na amplitude da tensão de até 10% em relação a tensão base do sistema, com duração entre 0,5 ciclo até 1 minuto. Esse evento é menos comum que os afundamentos, e sua duração é crucial para determinar o tamanho do impacto nos equipamentos (Bollen *et al*, 2009). Ademais, elevações de tensão são frequentemente associadas a condições de faltas no sistema, embora também possam ocorrer devido a mudanças na carga ou na configuração do sistema (Pereira, 2017).
- Interrupção - distúrbio crítico cuja principal característica é a redução de tensão para valores inferiores a 10% da tensão base do sistema, com duração de até 1 minuto (Borges, 2019). É também influenciado pelo tempo de operação dos dispositivos de proteção, por exemplo os religadores automáticos, que podem limitar a duração da interrupção (Paulillo *et al.*, 2025).

2.1.3 Variações de tensão de longa duração

Análogas às variações de curta duração, as variações de longa duração são caracterizadas como sobretensões, subtensões ou até mesmo pela interrupção total da tensão, ocorrendo por um período superior a 1 minuto (IEEE, 2019).

A subtensão caracteriza-se pela redução de 10 a 90% da tensão com relação à tensão de base com período superior a 1 minuto. Ela ocorre, em alguns casos, devido à sobrecarga na rede de transmissão, problemas nos transformadores ou linhas de transmissão e falhas em subestações (GONÇALVES *et al.*, 2018). Já a sobretensão é

caracterizada pela elevação da tensão acima do valor dos 10% em relação a base do sistema (Borges, 2019).

A interrupção sustentada é outro distúrbio crítico nos sistemas elétricos, cuja principal característica é a redução do valor eficaz da tensão para valores inferiores a 10% da tensão de base. Esse tipo de ocorrência está geralmente relacionado a faltas na rede, falhas em equipamentos críticos e desligamentos programados para manutenção ou reparos (Bollen *et al.*, 2009).

2.1.4 Desequilíbrio de tensão

Os sinais gerados em sistemas trifásicos equilibrados são senoidais e possuem mesma amplitude com defasagem de 120° entre si. Caso contrário, o sistema trifásico é considerado desequilibrado, afetando a QEE e os equipamentos conectados a essa rede.

A principal causa de desequilíbrio nas correntes e nas tensões é a distribuição inadequada de cargas monofásicas, que pode gerar correntes desequilibradas ao longo das fases do sistema (Dugan *et al.*, 2003). Outro fator que contribui para o desequilíbrio são as impedâncias e admitâncias das linhas aéreas, que, quando tem condutores com espaçamentos não uniformes, podem ocasionar quedas de tensão distintas para cada fase. Para atenuar esse problema, utilizam-se as transposições das linhas (Borges, 2019).

O controle e a regulamentação dos níveis de desequilíbrio de tensão nas redes de distribuição internacionais são definidos pelas normas IEC 61000 e EN-50160 (CENELEC, 2022; IEC. 2023.). No Brasil, a regulamentação é dada pelo módulo 8 PRODIST, estabelecido pela ANEEL. Para os barramentos de distribuição, o desequilíbrio deve ser de menor ou igual a 2%, com exceção para os de baixa tensão, que pode ser até 3%.

2.1.5 Distorções harmônicas

Os harmônicos são componentes de tensão ou corrente que consistem em múltiplos inteiros da frequência fundamental da rede e, em conjunto com a componente fundamental, resultam em uma forma de onda distorcida em comparação com a senoide pura (Borges, 2019).

A forma de onda distorcida pode ser decomposta matematicamente através da Transformada de Fourier (TF) em uma série de ondas senoidais puras com frequências múltiplas da frequência fundamental (Arrillaga; Watson, 2003).

De acordo com Franco (2015), a principal fonte de harmônicos são os equipamentos eletrônicos, em particular, reatores eletrônicos, inversores de frequência, retificadores e cargas não lineares. Os componentes harmônicos presentes no sistema podem provocar o sobreaquecimento de cabos e equipamentos, a diminuição do desempenho de motores e a operação indevida de disjuntores, relés e fusíveis (Fortes, 2016).

2.1.6 Flutuação de tensão

O fenômeno da flutuação de tensão (*flicker*) caracteriza-se pela variação do valor eficaz ou do pico de tensão instantânea em um sistema elétrico. Assim sendo, essa variação pode ocorrer de forma aleatória, repetitiva ou esporádica, o que traz como resultado uma oscilação que ocorre na forma variável de onda senoidal da tensão se comparada com a onda fundamental (Arrilla; Watson, 2003).

De acordo com Paulillo e Teixeira (2013), flutuações na tensão são variações sistemáticas dos valores eficazes de tensão, ou uma série de mudanças aleatórias cujas magnitudes normalmente não excedem faixas de valores preestabelecidos (faixa compreendida entre 95% e 105%) da base do sistema.

Essas variações são, em geral, provocadas pelas alterações rápidas nas potências ativas e reativas das cargas elétricas, como em fornos a arco. Outras causas incluem sobrecarga na rede elétrica, além de falhas nos equipamentos de transmissão e distribuição, também condições climáticas adversas, alterações bruscas de carga e operação inadequada de geradores de energia (Dugan *et al.*, 2003).

No Brasil, esse parâmetro é normatizado pelo PRODIST módulo 8, o qual orienta a forma de medição da flutuação de tensão, bem como os parâmetros que devem ser medidos para a realização do cálculo do indicador associado ao distúrbio.

2.1.7 Notching e Spike

Os distúrbios *notching* e *spike*, que já foram considerados distintos, são tratados atualmente, na literatura, como um distúrbio único com características variáveis. O *notching* trata-se de um distúrbio periódico na forma de onda da tensão elétrica, caracterizado por perturbações instantâneas e repetitivas que ocorrem durante a comutação de correntes em dispositivos eletrônicos de potência. Ele é caracterizado por cortes periódicos na forma de onda da tensão, com duração inferior a 0,5 ciclos. Essa duração está relacionada com a operação de equipamentos como conversores estáticos,

retificadores e inversores. Esse distúrbio gera distorções de alta frequência (acima de 1 kHz) que se propagam pela rede elétrica (IEEE, 1992; Starosta, 2011).

O *spike* é considerado com um caso especial do *notching*, logo trata-se de um distúrbio que se caracteriza por um aumento abrupto na tensão elétrica, com magnitude que pode atingir níveis da ordem de quilovolts (kV), mesmo em sistemas de baixa tensão. Sua duração varia de microssegundos a milissegundos, conforme definido na literatura técnica e em normas como a IEEE 1159-2019. Ele está associado a eventos de alta energia, como, por exemplo, descargas atmosféricas, chaveamento de cargas indutivas, e faltas bifásicas/monofásicas (Almeida; Moreira; Delgado, 2003; Fortunato, 2016). Salienta-se que o *spike*, muitas vezes, é confundido com o transitório impulsivo e é tratado como tal.

2.2 TRANSFORMADA DE FOURIER

A TF é um operador matemático essencial para a análise de funções e sinais, pois possibilita a decomposição de formas de onda complexas em componentes de frequência. Assim, ao lidar com um sinal não senoidal, a TF possibilita sua análise no domínio da frequência, representando-o por meio de componentes senoidais, cujas amplitudes variam em função da frequência.

A TF foi desenvolvida para o estudo da equação do calor por Joseph Fourier e sua formulação moderna é amplamente aplicada em áreas como engenharia, física e processamento de sinais (Bracewell, 2000; Pereira, 2017). Matematicamente, ela pode ser definida de acordo com a Equação (2):

$$\mathcal{F}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt, \quad (2)$$

em que $\mathcal{F}(\omega)$ é a TF de uma função no tempo $f(t)$, e ω a frequência angular

Para reconstrução do sinal utiliza-se a transformada inversa de Fourier que permite, a partir das componentes no domínio da frequência, obter o sinal no domínio do tempo. Em termos matemáticos ela pode ser definida conforme a Equação (3):

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega)e^{j\omega t} d\omega. \quad (3)$$

Na prática, sinais digitais devem ser analisados de forma discreta, portanto utiliza-se a TF discreta, que pode ser obtida por meio da Equação (4):

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\omega_0 n}, \quad (4)$$

sendo $X[k]$ os coeficientes da TF no domínio da frequência, $x[n]$ amostras do sinal no domínio do tempo, N o número de amostras do sinal; e ω_0 a frequência fundamental do sinal, com $n = 0, 1, 2, \dots, N - 1$; e $k = 0, 1, 2, \dots, N - 1$.

Uma forma de implementar a TF visando um melhor desempenho computacional é utilizando um conjunto de algoritmos conhecido como *Fast Fourier Transform* (FFT), que permite a decomposição de sinais em suas componentes de frequência de maneira eficiente (Fortunato, 2016). Além disso, a FFT pode ser utilizada em diversas áreas, como, por exemplo, o processamento de sinais, no qual possibilita realizar uma análise no domínio da frequência, o que facilita a identificação de padrões.

2.3 APRENDIZADO DE MÁQUINA

A aprendizagem de máquina é um campo dentro da inteligência artificial (IA) utilizado para desenvolver algoritmos que aprendam a partir de um conjunto de dados. O objetivo é que os algoritmos realizem, prevejam ou tomem decisões baseadas nos dados e aumentem seu desempenho de acordo com novos dados inseridos (Bishop, 2006).

Essa técnica pode ser utilizada para diversas finalidades, dentre elas se destacam: a regressão, que visa realizar previsões de variáveis contínuas e dependentes a partir de variáveis independentes; a classificação, que consiste em atribuir categorias a diversos itens; e a *clusterização* que busca dividir dados em subconjuntos semelhantes (Lima, 2021).

O aprendizado de máquina pode ser classificado em aprendizagem supervisionada, não supervisionada, semi-supervisionada e por reforço. A aprendizagem supervisionada possui um conjunto de dados de treinamento com rótulos, por meio do qual o modelo aprende a associar as entradas às saídas esperadas. Esse tipo de aprendizado pode ser empregado por meio de diversas técnicas, como a regressão, máquinas de vetores de suporte (SVM, do inglês *support vector machines*), árvores de decisão, florestas aleatórias, redes neurais artificiais (RNA), dentre outros.

A regressão é um método estático utilizado para presumir a possibilidade de ocorrer um evento com base em um ou alguns fatores de dados (Montgomery; Peck; Vining, 2012). Já as SVM são algoritmos que se destacam por sua eficiência em

problemas de classificação e regressão, operando através da identificação do hiperplano¹, ou seja, as SVM consideram um limite ideal, que maximiza a margem entre classes em espaços multidimensionais (Vapnik, 2000).

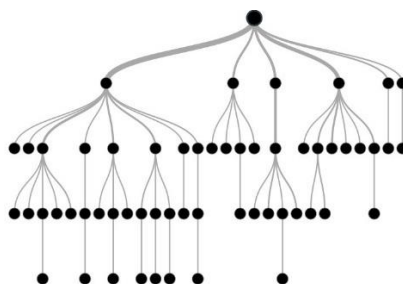
As árvores de decisão são modelos de aprendizado supervisionado que utilizam uma estrutura hierárquica para realizar a divisão de dados em categorias discretas, baseando-se em condições e atributos para tomar as decisões. Elas se destacam pela interpretabilidade e pela sua capacidade de tratar as relações que não são lineares entre variáveis (Breiman *et al.*, 1984). Nas florestas aleatórias, por sua vez, é realizada a combinação de múltiplas árvores de decisão, visando melhorar a precisão e a robustez das previsões. Esse tipo de algoritmo tem se destacado por sua capacidade de reduzir o *overfitting*, que ocorre quando um modelo se adapta bem ao treinamento, mas não consegue generalizar padrões de dados (Breiman, 2001).

Os quatro modelos de árvores de decisão que são utilizadas nesse trabalho foram:

- árvore fina (*fine tree*) – caracterizada por ter a presença de mais nós, ou seja, cria mais classes e possui um detalhamento profundo com relação aos dados (Hastie; Tibshirani; Friedman, 2009);
- árvore média (*medium tree*) – possui um número de nós inferior ao da árvore fina, implicando em uma menor precisão com relação aos dados (Hastie; Tibshirani; Friedman, 2009);
- árvore grossa (*coarse tree*) – possui um número de nós inferior se comparada com suas antecessoras e, dessa forma, é um modelo mais simples, que realiza uma análise mais superficial dos dados com uma menor precisão. Por outro lado, ela necessita de menos processamento computacional (Bishop, 2006);
- árvore *boosted trees* - realiza a combinação de múltiplos modelos simples (como árvores de decisão pequenas) em um modelo composto mais completo (Hastie; Tibshirani; Friedman, 2009).

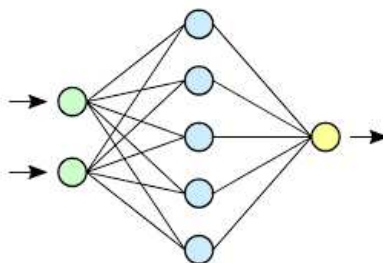
Cada novo modelo é cuidadosamente ajustado para identificar e corrigir os erros dos modelos anteriores, resultando em uma melhoria iterativa do desempenho tanto no desempenho preditivo quanto na capacidade de generalização (Hastie; Tibshirani; Friedman, 2009). Apresenta-se, na Figura 2, um modelo simples geral para uma árvore de decisão.

¹ Em termos de aprendizado de máquinas, o hiperplano é uma superfície de decisão que separa classes de dados em um espaço multidimensional.

Figura 2: Exemplo de árvore de decisão

Fonte: Rocha (2017).

As RNA, por sua vez, são sistemas computacionais inspirados na estrutura e no funcionamento do cérebro humano. Elas são projetadas para reconhecerem padrões e solucionarem problemas complexos. Essa técnica é bastante utilizada em tarefas de classificação, regressão e em processamento de dados não lineares, destacando-se por sua capacidade de generalização e de adaptação (Haykin, 2001). Um exemplo simples de RNA pode ser observado na Figura 3.

Figura 3: Exemplo de RNA

Fonte: Costa (2023).

Na aprendizagem não supervisionada, o algoritmo não recebe rótulos e deve encontrar padrões e estruturas nos dados por conta própria. Consiste, então, em treinar os algoritmos para identificar se existem padrões ocultos ou agrupamentos naturais em conjuntos não categorizados. Dessa forma, como não há respostas predefinidas, o modelo precisa descobrir autonomamente as relações existentes entre os dados (IBM, 2024).

A aprendizagem semi-supervisionada é treinada por uma combinação de uma baixa quantidade de dados rotulados com uma alta quantidade de dados não rotulados. A combinação de dados rotulados e não rotulados pode melhorar a classificação de novas

instâncias. Um exemplo desse tipo de aprendizagem são as máquinas de *Boltzmann* (Larochelle, 2012).

Por fim a aprendizagem por reforço pode ser treinada por meio da interação com um ambiente em tempo real, de tal maneira que uma solução é encontrada empregando tentativa e erro de um determinado problema. Dessa forma, o sistema de aprendizado recebe o nome de agente, sendo capaz de observar o ambiente, selecionar e executar ações, além de receber recompensas ou penalidades em resposta do seu desempenho. Esse tipo de aprendizagem é utilizado, por exemplo, na robótica (IBM, 2024).

2.3.1 Técnicas para validação e testes

Em geral, os dados utilizados nos modelos de aprendizado de máquina são divididos em dados de treinamento e de testes. Sendo comum utilizar 70% dos dados pois é uma quantidade suficiente para cobrir a diversidade dos dados (variabilidade, distribuição de classes e possíveis ruídos). Os 30% restantes são utilizados para testes, sendo uma amostra grande o suficiente para generalizar os dados não vistos no treinamento. Os dados de treinamento são o conjunto utilizado para ensinar um modelo de aprendizado de máquina, sendo essenciais para que o algoritmo aprenda padrões e faça previsões. Durante o treinamento, o modelo ajusta seus parâmetros para minimizar o erro entre suas previsões e os rótulos verdadeiros no conjunto de treinamento (Hastie; Tibshirani; Friedman, 2009).

Já os dados de teste são utilizados para avaliar o desempenho de um modelo de aprendizagem de máquina após o treinamento. São essenciais para verificar a capacidade do modelo de analisar para novos dados, ou seja, prever corretamente resultados em situações que não foram vistas durante o treinamento (Russell; Norvig, 2020).

Para realizar a divisão dos dados de forma imparcial, sem grandes influências do projetista, utiliza-se a técnica da validação cruzada (*cross validation*), por meio da qual se divide os dados em várias partes, treina-se o modelo em uma das partes e o testa nas demais. Esse processo é repetido várias vezes, com a finalidade de garantir sua eficiência (Hastie; Tibshirani; Friedman, 2009).

Uma das formas mais intuitivas de se avaliar a pertinência de um modelo de classificação é por meio das matrizes de confusão. A matriz de confusão é uma ferramenta visual que organiza as previsões em quatro categorias: verdadeiros positivos (VP), que ocorrem quando o modelo prevê que a classe é positiva e, de fato, ela é; os falsos positivos (FP), que ocorrem quando o modelo prevê uma classe positiva, mas ela é negativa;

verdadeiros negativos (VN), quando o modelo prevê corretamente uma classe negativa; e os falsos negativos (FN) que ocorrem quando o modelo prevê uma classe negativa, mas ela é, na verdade, positiva. Essa matriz é apresentada na Figura 4.

A matriz de confusão também pode ser composta pelo número de acertos e erros do modelo. No que diz respeito aos termos de classificação, as linhas da matriz correspondem as classes reais dos sinais testados e as colunas correspondem as classes previstas pelo modelo. Portanto, os elementos da diagonal principal indicam o número de acertos, enquanto os elementos fora da diagonal indicam o número de erros.

Figura 4: Matriz confusão.

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Fonte: Nogere, (2020).

Além da matriz de confusão, um parâmetro numérico muito utilizado para avaliar os modelos de aprendizado de máquina é a acurácia, uma medida estatística que indica a porcentagem de previsões corretas em relação ao total de previsões realizadas (Hastie; Tibshirani; Friedman, 2009). Matematicamente, a acurácia pode ser compreendida de acordo com a Equação (5):

$$acurácia = \frac{VP + VN}{VP + FP + VN + FN} = \frac{\text{previsões corretas}}{\text{total de previsões}} * 100. \quad (5)$$

2.3.2 Escolha dos atributos caracterizadores

Os dados de entrada de um modelo de aprendizado de máquina são, na verdade, atributos que caracterizam os sinais. A escolha desses atributos, ou parâmetros, deve propiciar a classificação correta, sendo, portanto, a base para o treinamento, teste e eventual uso dos modelos.

Os atributos de entrada dependem, principalmente, da visão do especialista que, com o devido conhecimento do assunto é capaz de identificar as principais características que definem as classes a serem estudadas. No entanto, é possível aprimorar os resultados reduzindo o conjunto de parâmetros escolhidos. Para isso, alguns algoritmos podem ser

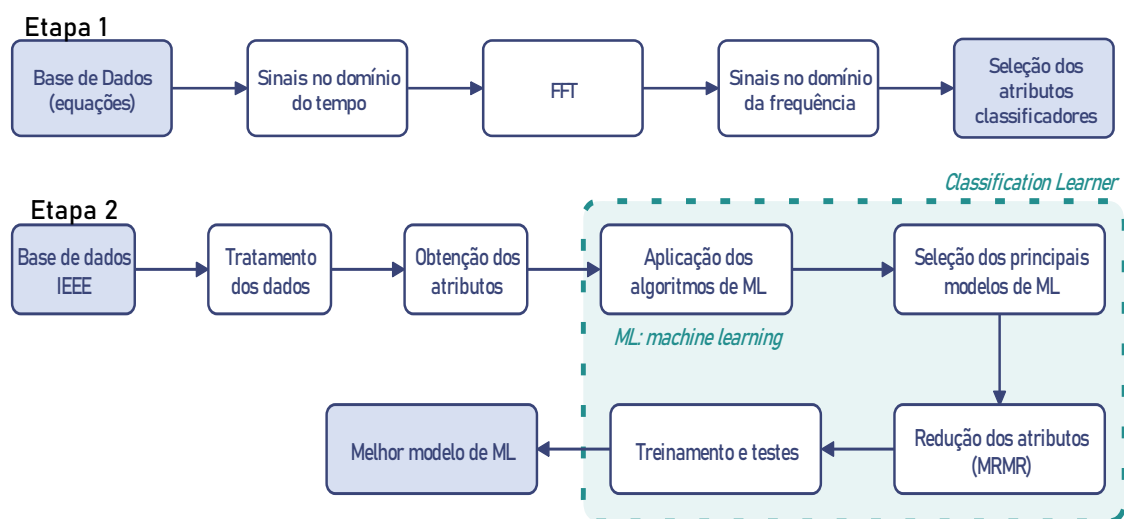
utilizados, dentre os quais se destaca o algoritmo de máxima relevância e mínima redundância (MRMR).

O MRMR foi inicialmente proposto por Ding e Peng (2005) e posteriormente aprimorado por Zhao, Anand e Wang (2019). Em síntese, os autores estabelecem duas condições principais: uma de máxima relevância e outra de mínima redundância, que são aplicadas às variáveis. Essas condições são utilizadas em um algoritmo responsável por atribuir uma pontuação a cada uma das variáveis e definir quais são mais importantes (Froes, 2022). Destaca-se que a análise é realizada entre as variáveis preditivas, ou atributos, e a variável alvo, sendo utilizado em diversas áreas dentro da IA. Para a aplicação em aprendizado de máquina, o algoritmo pode reduzir o número de atributos e diminuir o esforço computacional necessário para o uso do modelo escolhido.

3 METODOLOGIA

O presente trabalho tem característica quantitativa e exploratória, visando analisar dados sintéticos em busca de padrões e possíveis classificações para distúrbios de QEE. Para isso, desenvolveu-se um classificador de avaliando os principais atributos de sinais de tensão e testando diferentes métodos de aprendizado de máquina. Apresenta-se, no Fluxograma da figura 5, os procedimentos realizados para a escolha dos atributos classificadores (Etapa 1) e seleção do modelo de aprendizado de máquina (etapa 2) adequado ao problema supracitado.

Figura 5 – Fluxograma referente as etapas do trabalho.



Fonte: Autor, (2025).

Inicialmente, adotou-se uma abordagem teórica para a compreensão sobre os conceitos básicos relacionados à pesquisa, como aprendizado de máquina, QEE e análises de sinais no domínio do tempo e da frequência. Adicionalmente foi realizada uma revisão bibliográfica a fim de identificar possíveis pontos de contribuição. Para essas etapas foram utilizados livros, artigos científicos e outras fontes relevantes.

Em seguida, no *software* MATLAB[®], foram analisados sinais gerados por meio de equações matemáticas, visando simular os distúrbios e identificar os possíveis atributos caracterizadores para classificação. O objetivo desta etapa foi encontrar quais os principais atributos que podem ser utilizados para desenvolver o classificador da QEE. Para tal, foram adotadas equações descritivas para cada um dos distúrbios como apresentado na Tabela 2, na qual a função $u(t)$ refere-se à função degrau.

Todos os distúrbios descritos na Tabela 2 foram estudados tanto no domínio do tempo quanto no domínio da frequência e, por meio das análises, diversos parâmetros foram separados a fim de serem avaliados como atributos para a classificação dos distúrbios. Destaca-se que, para este estudo, foi utilizada a frequência de amostragem de 15360 Hz e a frequência fundamental de 60 Hz.

Em seguida, adotou-se uma base de dados sintética fornecida pelo IEEE, na qual constam senoides puras e diversos distúrbios, tanto isolados quanto combinados (IEEE Dataport, 2024). A base utilizada é fundamentada em critérios técnicos, científicos e práticos que conferem robustez, credibilidade e reprodutibilidade. Além disso, o IEEE *Dataport* fornece arquivos para utilização no Matlab® ou Microsoft Excel®, contendo os dados e os rótulos associados. As especificações dos sinais estão apresentadas na Tabela 3.

Tabela 2: Modelo matemático dos distúrbios.

Distúrbio	Modelo	Parâmetro IEEE
Senoide Pura	$v(t) = A * \sin(\omega t)$	$\omega = 2\pi f$
Afundamento de tensão	$v(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) * \sin(\omega t)$	$0,1 \leq \alpha \leq 0,9$ $T \leq t_2 - t_1 \leq 9 T$
Elevação de tensão	$v(t) = (1 + \alpha(u(t - t_1) - u(t - t_2))) * \sin(\omega t)$	$0,1 \leq \alpha \leq 0,8 T$ $\leq t_2 - t_1 \leq 9 T$
Interrupção	$v(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) * \sin(\omega t)$	$0,9 \leq \alpha \leq 1,0$ $T \leq t_2 - t_1 \leq 9 T$
<i>Flicker</i>	$V(t) = (1 - \alpha_f \sin(\beta \omega t)) * \sin(\omega t)$	$0,1 \leq \alpha_1 \leq 0,2$ $5 \leq \beta \leq 20 \text{ Hz}$
Transitório Oscilatório	$v(t) = \left(\sin(\omega t) + \left(\alpha e^{\left(\frac{(t-t_1)}{\tau}\right)} \right) * (\omega_n(t - t_2)) \right) * (u(t_2 - t_1))$	$0,1 \leq \alpha \leq 0,8$ $0,05 T \leq t_2 - t_1 \leq 3 T$ $8 \text{ ms} \leq \tau \leq 40 \text{ ms}$ $300 \leq \tau f_n \leq 900 \text{ Hz}$
Transitório Impulsivo	$v(t) = \sin(\omega t) + \left(\alpha e^{\left(\frac{(t-t_1)}{\tau}\right)} \right) * (u(t_2 - t_1))$	$1 \leq \alpha \leq 10$ $0,05 T \leq t_2 - t_1 \leq 3 T$ $8 \text{ ms} \leq \tau \leq 40 \text{ ms}$
Harmônicos	$v(t) = \alpha_1 \sin(\omega t) + \alpha_2 \sin(\omega t) + \alpha_3 \sin(\omega t) + \alpha_4 \sin(\omega t) + \alpha_5 \sin(\omega t) + \alpha_6 \sin(\omega t) + \alpha_7 \sin(\omega t)$	$0,1 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0,15$
<i>Notching</i>	$v(t) = \sin(\omega t) - \text{sign}(\sin(\omega t)) * \left[\sum_{n=0}^9 K * \{u(t - (t_1 + 0,02n)) - u(t - (t_2 + 0,02n))\} \right]$	$0,1 \leq K \leq 0,4$ $0 \leq t_1, t_2 \leq 0,5 T$ $0,01 T \leq t_2 - t_1 \leq 0,05 T$

Fonte: Fortunato (2016).

Tabela 3: Especificação dos sinais presentes na base de dados sintética.

Amplitude	Normalizada em 1 pu
Frequência fundamental	50 Hz
Frequência de amostragem	5 kHz
Tamanho dos sinais	10 ciclos da fundamental
Amostras por ciclo	100 amostras
Número de amostras de cada sinal	1000 amostras

Fonte: IEEE Dataport (2024).

A base de dados sintética contém 7500 sinais divididos em 15 classes pré-definidas. Dessa forma, cada classe contém 500 sinais para cada tipo de distúrbio ou combinação de distúrbios. A base de dados foi analisada no *software* Matlab® e adaptada visando diversificar os sinais correspondentes à senoide pura, visto que todos os 500 sinais para esta classe eram iguais. Nesse caso, adotou-se uma fase aleatória em radianos para cada sinal senoidal puro. Salienta-se que os dados são rotulados pela classe, sendo assim a aprendizagem supervisionada foi utilizado. As classes e rótulos utilizados são apresentados na Tabela 4.

Tabela 4: Classificação dos distúrbios na base de dados sintética.

Rótulo/Classe	Distúrbio
1	Senoide pura
2	Afundamento
3	Elevação
4	Interrupção
5	Transitório impulsivo e <i>Spike</i>
6	Transitório oscilatório
7	Harmônicos
8	<i>Flicker</i>
9	<i>Notching</i>
10	Harmônicos com afundamento
11	Harmônicos com elevação
12	Afundamento com transitório oscilatório
13	Elevação com transitório oscilatório
14	Harmônicos com afundamento e transitório oscilatório
15	Harmônicos com elevação e transitório oscilatório

Fonte: IEEE Dataport (2024).

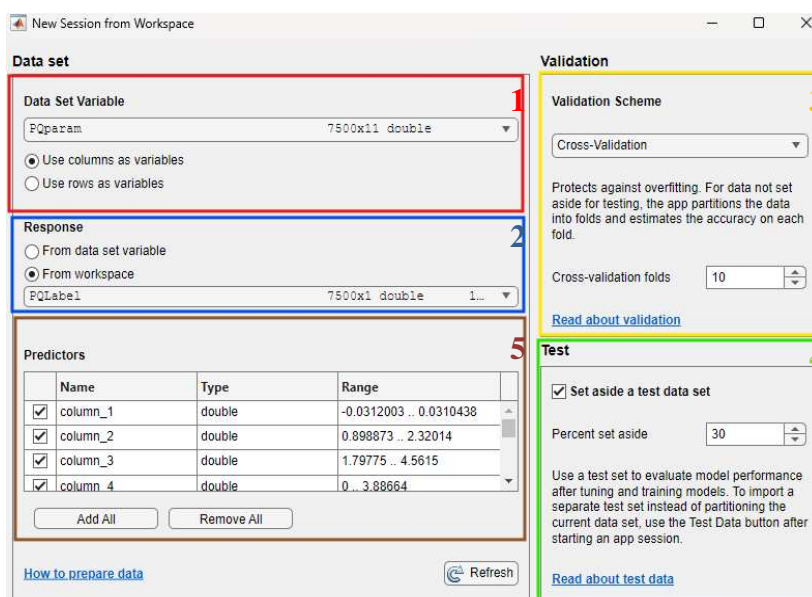
Por fim, os dados foram divididos em treinamento e testes, e os atributos selecionados foram utilizados como entrada para diferentes métodos de aprendizagem visando selecionar o mais adequado para a aplicação em questão.

Nesta etapa, realiza-se um pré-processamento dos dados com a finalidade de obter os parâmetros necessários para realizar o teste e o treinamento. Os parâmetros foram

obtidos utilizando a matriz PQData, fornecida pelo IEEE, com 7500 linhas e 1000 colunas e o vetor PQLabel, também fornecido pelo IEEE, e que contém as classes de cada um dos distúrbios. Dessa forma, uma matriz PQparam foi criada, contendo 7500 linhas e 11 colunas, a qual detém os parâmetros extraídos dos sinais da base de dados.

Em seguida, utilizou-se o aplicativo *Classification Learner* do software Matlab®, cuja janela para construção de uma nova sessão pode ser observada na Figura 6. Esta janela é utilizada para configuração inicial do método, e está dividida em alguns itens numerados, para fins explicativos. Os itens de número 1 e 2 representam os dados de entrada e a resposta, ou seja, PQparam e PQLabel respectivamente. No item 3, escolhe-se qual técnica de validação cruzada utilizar para realizar os procedimentos de treino e teste, e no item 4 reserva-se uma porcentagem dos dados para teste, sendo essa porcentagem excluída dos dados de treinamento. No item 5 estão presentes todos os atributos usados como entrada, os quais podem ser selecionados ou não, a depender da decisão do projetista.

Figura 5: Configuração do *Classification Learner*.



Fonte: Autor (2025).

4 RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os principais resultados obtidos ao longo da pesquisa. Inicialmente, destaca-se a seleção dos atributos classificadores, obtidos por meio da análise de sinais teóricos para os diversos distúrbios associados à QEE. Em seguida, é apresentado o desempenho de cada um dos modelos de aprendizado de máquina utilizados para classificação dos distúrbios.

4.1 SELEÇÃO DOS ATRIBUTOS CLASSIFICADORES

Inicialmente, para a escolha dos principais atributos, foram analisados sinais obtidos por meio das equações matemáticas de cada distúrbio fornecidas pelo IEEE e destacadas na Tabela 2. Nesta etapa o MATLAB® foi utilizado tanto para implementação das equações temporais, quanto para análise no domínio da frequência utilizando a FFT. Dessa forma, esse conjunto de dados permitiu uma inspeção visual dos distúrbios de QEE em ambos os domínios, e, em conjunto com o estudo previamente realizado sobre o assunto, foi possível identificar as principais características classificadoras. Apresenta-se, na Tabela 5, os parâmetros utilizados para simulação de cada um dos distúrbios.

Tabela 5: Parâmetros para simulação dos distúrbios por equações.

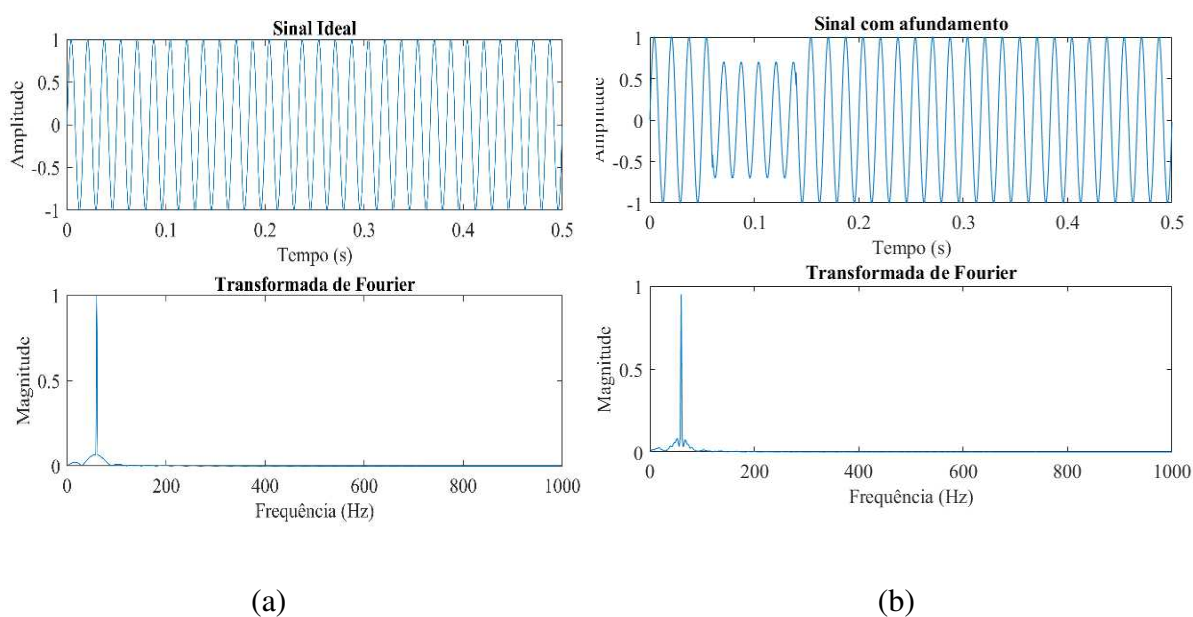
Distúrbio	Parâmetros para Simulação
Senoide Pura	$A = 1, f = 60Hz$
Afundamento de tensão	$\alpha = 0.3, t_1 = 0.06, t_2 = 0.14$
Elevação de tensão	$\alpha = 0.3, t_1 = 0.06, t_2 = 0.14$
Interrupção	$\alpha = 0.95, t_1 = 0.06, t_2 = 0.14$
Flutuação de tensão	$\alpha_f = 0.15, \beta = 15$
Transitório Oscilatório	$\alpha = 0.8, t_1 = 0.08, t_2 = 0.1, \tau = 0.02$
Transitório Impulsivo	$\alpha = 2, t_1 = 0.085, t_2 = 0.088, \tau = 0.02$
Harmônicos	$\alpha_3 = 0.05, \alpha_5 = 0.10, \alpha_7 = 0.15$
Notching	$K = 0.4, t_1 = 0.006, t_2 = 0.0065$

Fonte: Autor (2025).

Como esperado, no sinal senoidal puro, apresentado na Figura 7(a), observou-se a periodicidade, a amplitude máxima inalterada, a semelhança entre os semiciclos positivo e negativo e a tensão média nula. Além disso, no domínio da frequência apenas

a componente fundamental está presente. Já no sinal com afundamento, apresentado na Figura 7(b), observa-se a redução na amplitude por alguns ciclos. Dessa forma, características como tensão média e tensão de pico a pico apresenta valores reduzidos quando comparadas à senoide pura. Nota-se ainda que, neste caso, as tensões de pico a pico variam em diferentes em instantes de tempo diferentes. No domínio da frequência, no entanto, observa-se apenas a alteração na amplitude da componente fundamental.

Figura 6: (a) Sinal ideal, (b) Sinal com afundamento.



Fonte: Autor (2025).

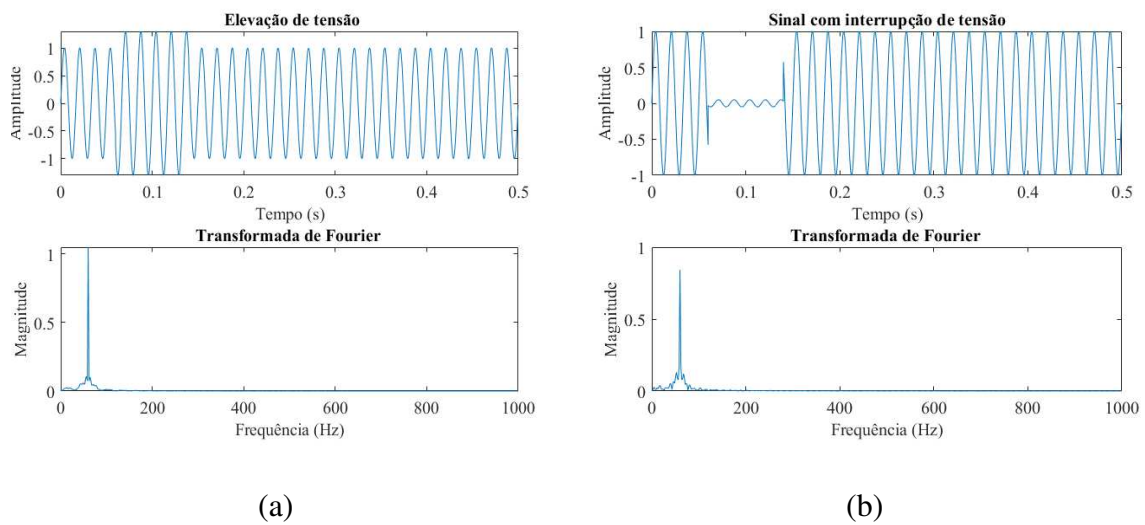
Analisando a elevação de tensão, apresentada na Figura 8(a), nota-se o aumento esperado na amplitude durante o distúrbio, o que deve alterar características como tensão média e tensão de pico a pico. No domínio da frequência, a componente fundamental também se altera, assumindo valores maiores. O sinal com interrupção de tensão, ilustrado na Figura 8(b), apresenta características semelhantes ao sinal com afundamento, porém com reduções maiores de amplitude e, conseqüentemente, no valor médio e no valor de pico a pico durante o evento. No domínio da frequência, além da redução na amplitude da componente fundamental, observam-se componentes de frequência próximas à fundamental, que podem ser provocadas pelas discontinuidades presentes no sinal no domínio do tempo.

O transitório impulsivo, apresentado na Figura 9(a), tem uma tensão máxima significativamente maior que os demais distúrbios citados, o que eleva também a tensão de pico a pico e a tensão média. Além disso, o sinal pode ser considerado assimétrico entre os semiciclos no instante do distúrbio. No domínio da frequência observam-se

componentes de baixa frequência próximas à fundamental, assim como um leve aumento dessa componente.

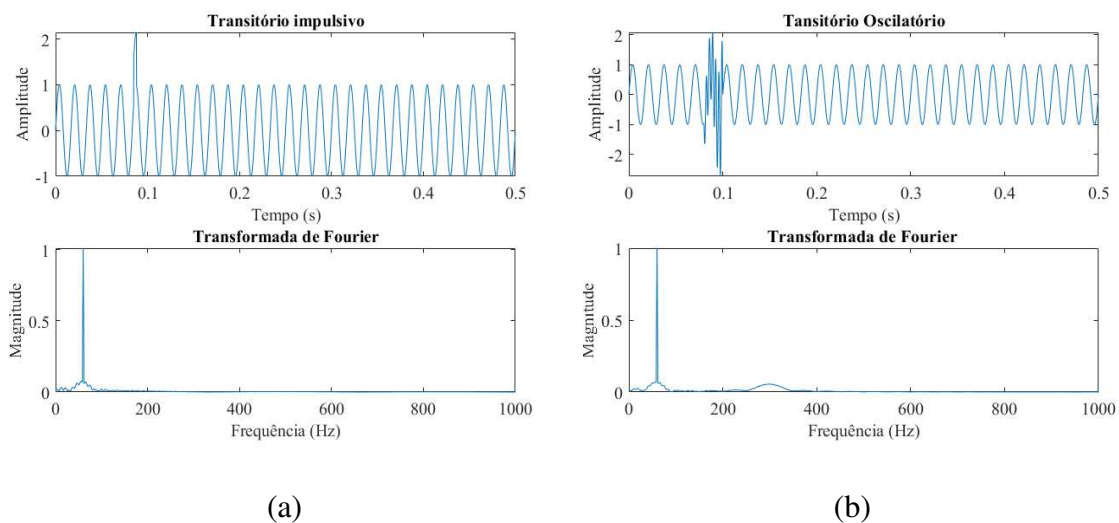
No sinal referente ao transitório oscilatório, observado na Figura 9(b), também há alterações na tensão média e tensão de pico a pico. No entanto, observa-se que o número de picos (máximos locais), durante os dez ciclos observados, é maior do que nos sinais anteriores. Destaca-se ainda que, no domínio da frequência, além das componentes próximas à componente fundamental, observa-se a presença significativa do 5º harmônico.

Figura 7: (a) Elevação de tensão, (b) interrupção de tensão.



Fonte: Autor (2025).

Figura 8: (a) Transitório impulsivo, (b) Transitório oscilatório.



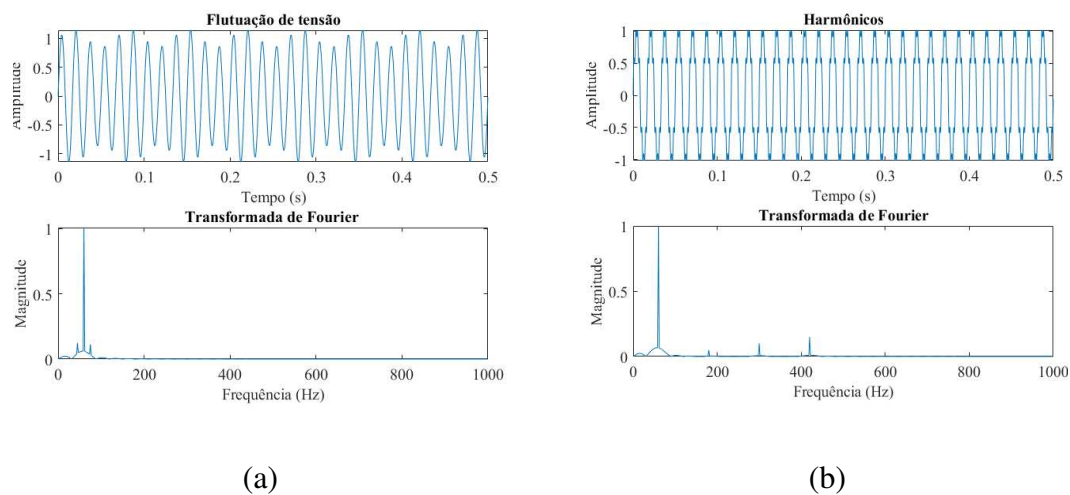
Fonte: Autor (2025).

Para a flutuação de tensão (Figura 10(a)), nota-se que, embora o número de picos não se altere, a amplitude máxima varia durante todo o período analisado, causando

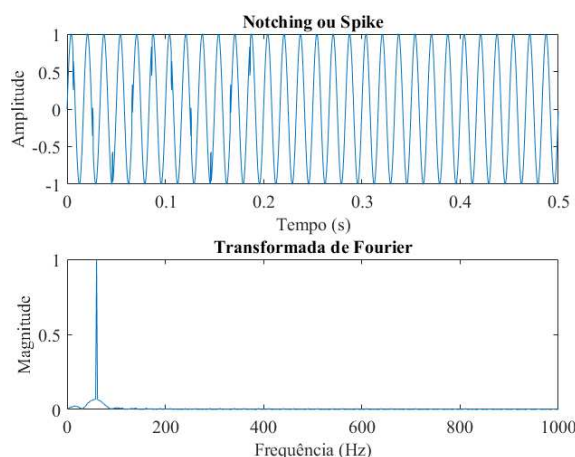
alterações na tensão média e tensão de pico a pico. No domínio da frequência, surgem componentes de baixa frequência e baixa amplitude próximas à frequência fundamental, sem alteração na amplitude da componente de 60 Hz. Para a análise dos harmônicos adotou-se a inserção do 3º, 5º e 7º harmônicos, por serem os mais proeminentes no estudo de QEE aplicado a sinais de tensão. Essas componentes podem ser observadas na Figura 10(b). Além disso, no domínio do tempo, observa-se um estreitamento dos picos, podendo causar alterações no valor máximo, na tensão eficaz e até mesmo no número de vezes que a função atinge o máximo local durante o período analisado.

Por fim, os distúrbios *notching* e *spike* foram analisados em conjunto. Na Figura 11 apresenta-se um caso de *notching*. No domínio do tempo, observam-se cortes no sinal, que podem reduzir a tensão média, a tensão de pico e a tensão eficaz. Além disso, o número de picos também é alterado. Para o *spike*, a diferença principal é que os cortes acontecem no sentido oposto, resultando em um aumento nos parâmetros relacionados com a amplitude da tensão. Não foram observados, no entanto, grandes divergências no domínio da frequência quando comparados à senoide pura.

Figura 9:(a) Flutuação de tensão; (b) Harmônicos



Fonte: Autor (2025).

Figura 10: Notching ou Spike.

Fonte: Autor (2025).

Após esta análise, foi possível destacar algumas características principais. Atributos como tensão máxima (V_{max}), tensão média (V_{med}) e tensão eficaz (V_{rms}) podem ser utilizados para identificar distúrbios como elevação, afundamento, interrupção e até mesmo auxiliar na classificação de transitórios ou distorções harmônicas. Adotou-se também a máxima diferença de amplitude (V_{amax}), a diferença entre o maior pico e o menor pico (V_{ppdif}), o número de picos (N_{picos}), que identificam os máximos locais e podem contribuir na classificação de transitórios oscilatórios, flutuação de tensão e até mesmo harmônicos. Este último é reforçado pela análise de parâmetros como a componente fundamental ($fund$), o segundo harmônico (h_2), capaz de identificar assimetrias entre os semiciclos do sinal, o terceiro (h_3), o quinto (h_5) e o sétimo harmônico (h_7). Os atributos escolhidos estão apresentados na Tabela 6, na qual se observa que cada coluna da matriz PQparam, utilizada como entrada dos modelos de aprendizagem, corresponde a um parâmetro específico. Para facilitar identificações futuras, esse grupo de parâmetros será chamado de grupo A11.

Tabela 6: Atributos escolhidos e sua posição na matriz PQparam (Grupo A11)

Parâmetros	Distúrbio	Coluna
V_{med}	Valor médio do sinal	Coluna_1
V_{max}	Valor máximo do sinal	Coluna_2
V_{amax}	Amplitude máxima – amplitude mínima	Coluna_3
V_{ppdif}	Diferença entre o pico máximo e o pico mínimo	Coluna_4
V_{rms}	Valor RMS	Coluna_5
h_2	2° Harmônicos	Coluna_6
h_3	3° Harmônicos	Coluna_7

h_5	5° Harmônicos	Coluna_8
h_7	7° Harmônicos	Coluna_9
F_{und}	Frequência fundamental	Coluna_10
N_{picos}	Número de picos	Coluna_11

Fonte: Autor, 2025.

4.2 CLASSIFICAÇÃO POR APRENDIZADO DE MÁQUINA

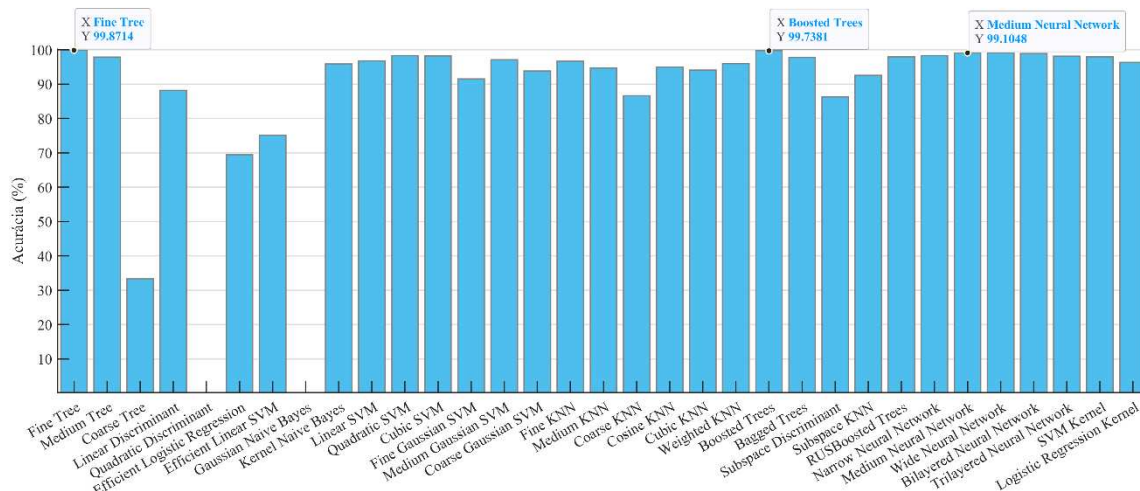
Para esta etapa, foi utilizado um computador com processador Intel® Core-i5 da 8ª geração com 8 GB de memória RAM (DDR4). A configuração afeta diretamente o tempo de processamento computacional necessário para o treinamento das técnicas de aprendizagem.

A matriz PQparam, que contém os atributos escolhidos para os 7500 sinais da base sintética, foi utilizada como entrada de dados no aplicativo *Classification Learner* do Matlab®. Os rótulos correspondentes aos resultados foram inseridos empregando o vetor PQLabel, a validação cruzada em dez subconjuntos (*10-fold cross-validation*) foi realizada nessa etapa para garantir a robustez e evitar o *overfitting*. Adotou-se, seguindo a literatura acerca do tema (Murphy, 2012; Bishop, 2006), 70% dos dados para treino e 30% para testes.

Inicialmente foram adotados todos os modelos de aprendizagem disponíveis no *Classification Learning* para treinamento e teste dos dados. Dessa forma, para o grupo A11, foram testadas as seguintes técnicas e suas subdivisões: árvore de decisão, regressão, SVM e redes neurais.

Na Figura 12 apresentam-se os resultados para todas as técnicas analisadas, com destaque para os três melhores modelos que serão utilizados como referência para a análise dos parâmetros: a árvore fina (*fine tree*) com 99,87% de acurácia, a árvore do tipo *boosted* com 99,73% e a RNA média com 99,10%. Destaca-se que a árvore do tipo floresta aleatória (*boosted*) é um combinado (*ensemble*) de árvores simples que busca corrigir possíveis erros em árvores finas.

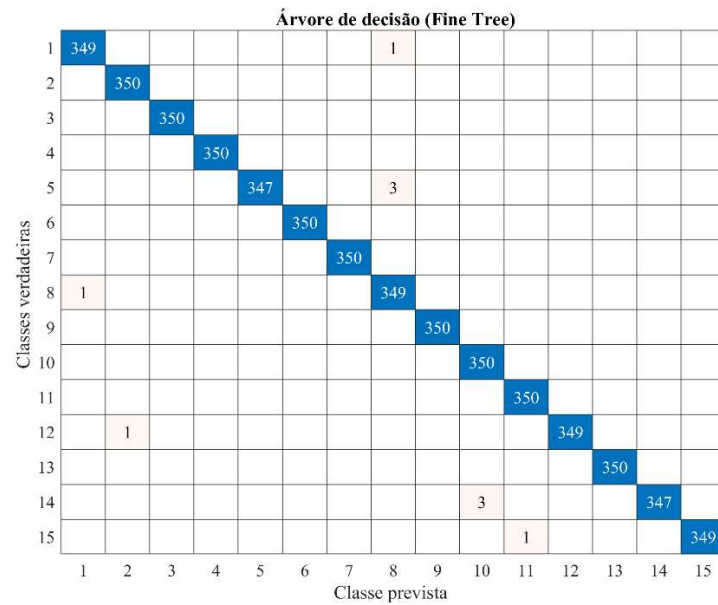
Figura 11: Resultados para todas as técnicas analisadas com os 11 atributos.



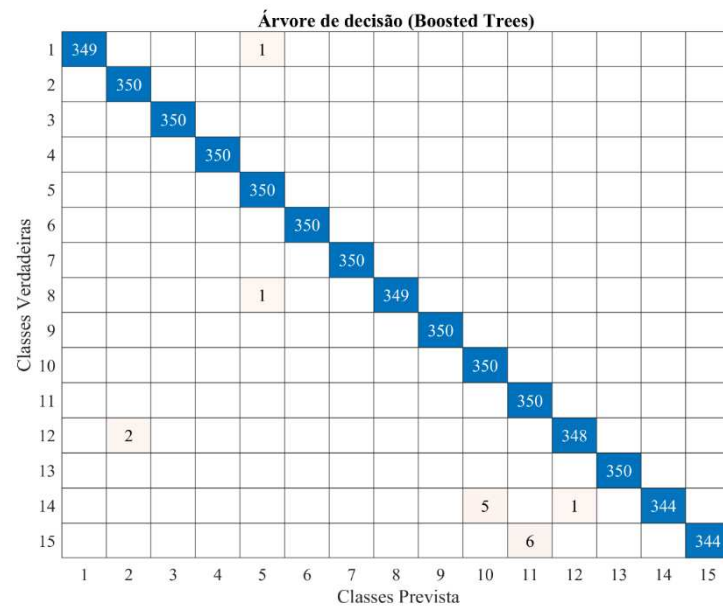
Fonte: Autor (2025).

A matriz de confusão para a *fine tree* está apresentada na Figura 13 e, analisando a diagonal principal, observa-se uma taxa de acertos elevada. Por outro lado, os números fora da diagonal evidenciam casos em que o modelo previu incorretamente, com destaque para o transitório impulsivo/*Spike* (classe real) com *flicker* (classe prevista), resultando em três erros. Da mesma forma, os harmônicos com afundamento e transitório oscilatório (classe real) foram confundidos com harmônicos com afundamento (classe prevista), também com três erros cada.

A segunda técnica que se destacou durante o treinamento foi a *Boosted Tree*, cuja matriz de confusão é apresentada na Figura 14. Nota-se que o maior número de erros (seis) apresentado foi na classe harmônicos com elevação e transitório oscilatório (classe real) e harmônicos com elevação (classe prevista). Em segundo lugar harmônicos com afundamento e transitório oscilatório (classe real) com harmônicos com afundamento (classe prevista) com cinco erros.

Figura 12: Matriz de confusão para árvore de decisão (*fine tree*).

Fonte: Autor (2025).

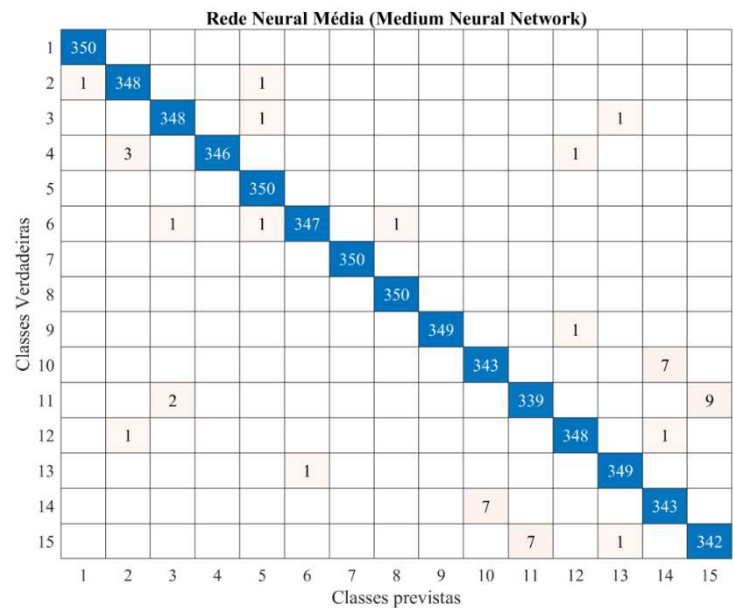
Figura 13: Matriz de confusão para árvore de decisão (*Boosted Trees*).

Fonte: Autor (2025).

A terceira melhor técnica foi a RNA média, cuja matriz de confusão para o teste com todos os parâmetros está apresentada na Figura 15. Observando a diagonal principal, nota-se o seu bom desempenho. O maior número de erros ocorreu para harmônicos com elevação (classe real) e a harmônicos com elevação e transitório oscilatório (classe prevista) totalizando nove previsões erradas. Ademais, algumas outras classes

apresentaram sete erros cada, sendo elas: harmônicos com afundamento, elevação com transitório oscilatório, harmônicos com elevação e transitório oscilatório (todas classes reais), os quais podem ser selecionados ou não harmônicos com afundamento e transitório oscilatório, harmônicos com afundamento e harmônicos com elevação, respectivamente (classes previstas).

Figura 14: Matriz de confusão para RNA média (*Medium Neural Network*).



Fonte: Autor (2025).

Em suma, na Tabela 7, destacam-se as principais características de cada modelo utilizado durante o treinamento. Nota-se que o modelo da árvore de decisão fina obteve os melhores resultados tanto no critério de tempo quanto de resultado, foi mais rápido e preciso.

Tabela 7: Resultados para o treinamento com todos os parâmetros.

Modelo	Acurácia (%)	Tempo de simulação (s)
<i>Fine Tree</i>	99,8	46,84
<i>Boosted Trees</i>	99,7	272,82
<i>Medium Neural Network</i>	99,1	356,29

Fonte: Autor (2025).

A etapa seguinte consistiu em realizar os testes com os mesmos modelos escolhidos durante o treinamento, tendo como objetivo principal verificar a consistência dos resultados obtidos na etapa de treino e, utilizando os 30% dados restantes, a fim de verificar se não ocorreu o *overfitting*.

Os resultados para os testes estão sintetizados na Tabela 8, na qual observa-se que a árvore de decisão fina teve um desempenho ainda melhor, com acurácia de 99,9%. O

principal erro nesse caso foi entre o transitório impulsivo e *spike* (classe real) e a classe prevista *flicker*, com dois erros.

O modelo *boosted trees* também teve um desempenho superior com acurácia de 99,8%. O maior número de erros durante o teste foi entre os distúrbios transitório impulsivo e *Spike* (classe real) e a classe prevista *Flicker*, com dois erros.

A RNA média também elevou a acurácia para 99,3 % e o maior número de previsões erradas durante o teste foi os harmônicos com elevação e transitório oscilatório (classe prevista) e harmônicos com elevação (classe real), com 3 erros.

Tabela 8: Resultados para os testes com todos os parâmetros.

Modelo	Acurácia (%)
<i>Fine Tree</i>	99,9
<i>Boosted Trees</i>	99,8
<i>Medium Neural Network</i>	99,1

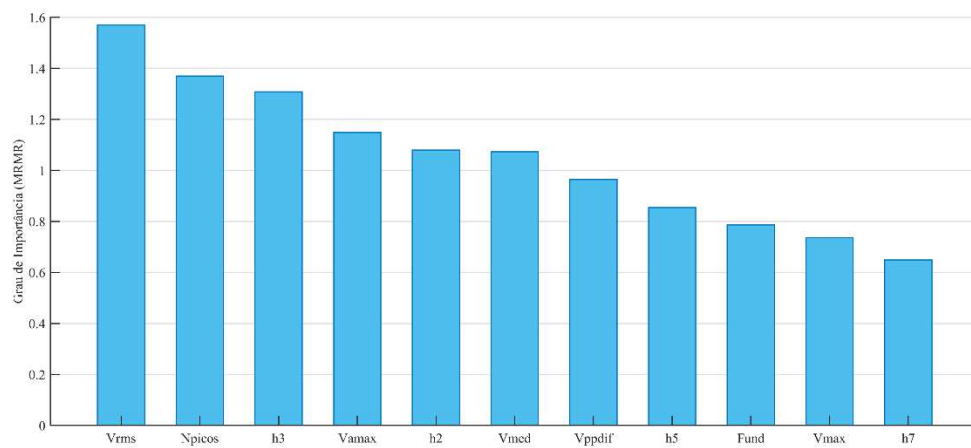
Fonte: Autor (2025).

4.2.1 Redução dos parâmetros classificadores

O objetivo principal da redução de parâmetros é buscar uma forma de diminuir o tempo de processamento sem comprometimento demasiado na eficácia do modelo. Além disso, previne-se o *overfitting* pois a presença de muitos parâmetros pode fazer com que o modelo se adapte excessivamente aos dados.

Para selecionar quais parâmetros seriam removidos, observou-se quais classes apresentavam mais previsão incorretas, que, neste caso, foram as classes que combinam distúrbios.

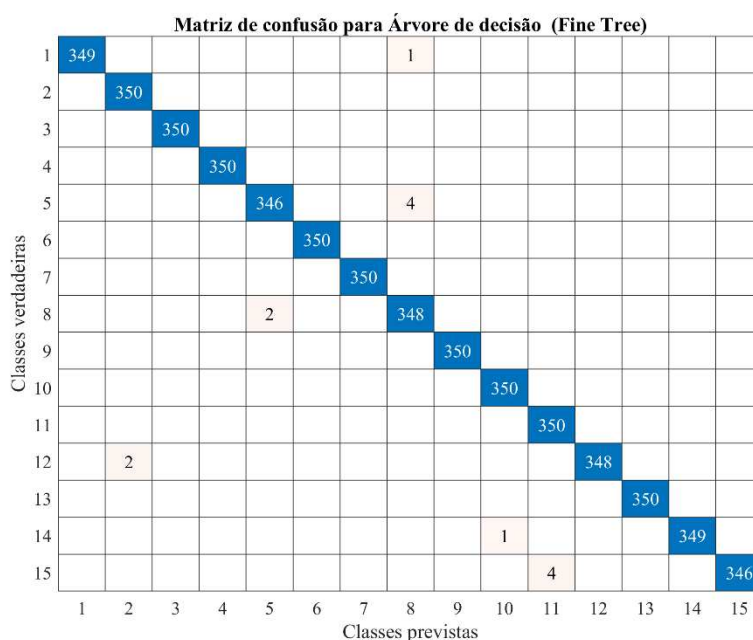
Sendo assim, os parâmetros foram escolhidos com base no conhecimento teórico e no algoritmo disponível dentro do próprio aplicativo (*Classification Learner*): o MRMR. Com esse algoritmo é possível identificar quais parâmetros mais influenciam os modelos, reduzindo a redundância entre eles. O resultado para essa análise está apresentado na Figura 16, por meio da qual constatou-se a possibilidade de remover três parâmetros: *fund*, *Vmax* e *h7*. Dessa forma, foi possível treinar os modelos pré-selecionados com os oito parâmetros restantes (Grupo A8), e avaliar a acurácia e o tempo necessário para o treinamento. Isso foi realizado removendo também os parâmetros *h5*, *Vppdif* e *Vmed*, formando o Grupo A5, com os cinco parâmetros restantes: *Vrms*, *Npicos*, *h3*, *Vamax* e *h2*.

Figura 15: Resultados do algoritmo MRMR.

Fonte: Autor (2025).

4.2.1.1 Treinamento e testes para o Grupo A8

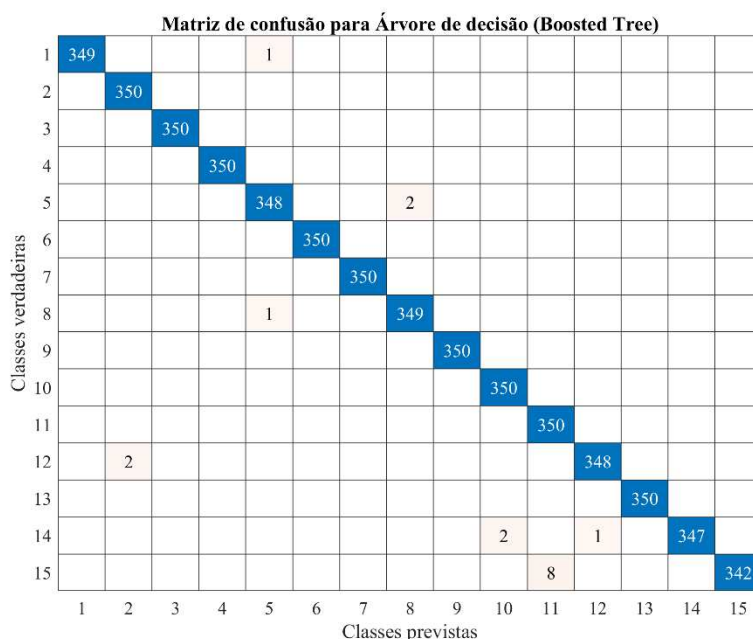
Para o treino da árvore de decisão fina, foi obtida a matriz de confusão apresentada na Figura 17, na qual nota-se um aumento das previsões erradas, com destaque para as classes que já apresentaram previsões incorretas anteriormente. Os principais foram erros foram entre o transitório impulsivo/*spike* (classe real) e *flicker* (classe prevista), e entre harmônicos com elevação e transitório oscilatório (classe real) e harmônicos com elevação (classe prevista), com quatro erros cada uma. O *flicker* também foi previsto como transitório impulsivo/*spike* e o transitório oscilatório com afundamentos também foi previsto com afundamento, ambos os casos com dois erros. Apesar dessas previsões incorretas, vale salientar que a acurácia foi reduzida em apenas 0,1%, finalizando o treino com 99,7% em apenas 2,38 segundos, uma redução de 44,46 segundos em relação ao treinamento com 11 atributos.

Figura 16: Matriz de confusão para árvore de decisão (Grupo A8).

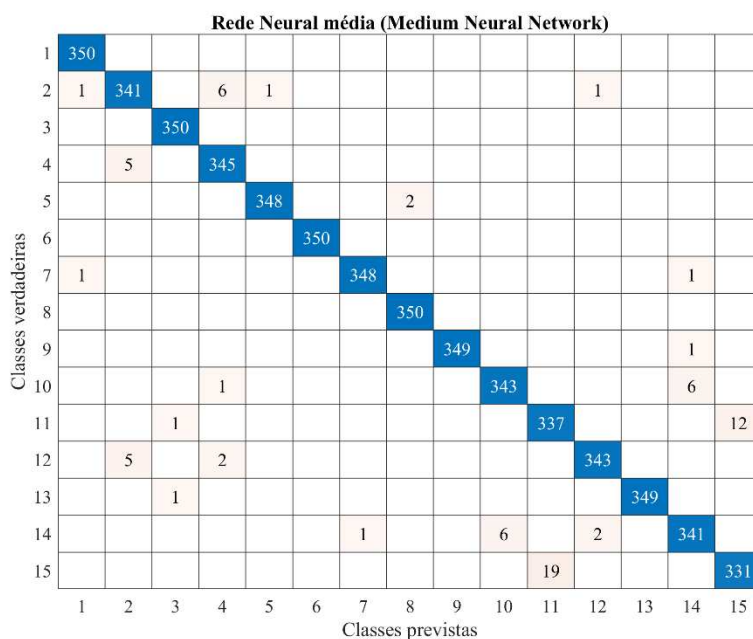
Fonte: Autor (2025).

Analisando a técnica *boosted trees*, observa-se que a redução dos três parâmetros elevou os números de previsões incorretas com destaque para as classes harmônicos com elevação e transitório oscilatório (classe real) e harmônicos com elevação (classe prevista), com oito erros. Além disso, houve dois erros na classificação da classe harmônicos com afundamento e transitório oscilatório (classe real) como harmônicos com afundamento (classe prevista), com dois erros, como apresentado na Figura 18. Para essa técnica, a acurácia se manteve em 99,7%, enquanto o tempo de treinamento reduziu para de 12,76 segundos, o que representa um ganho superior a 4 minutos (260,06 segundos) com relação ao treinamento com 11 atributos.

Para a RNA média também houve um aumento nas previsões erradas, principalmente entre as classes harmônicos com elevação e transitório oscilatório (classe real) e a harmônicos com elevação (classe prevista) com 19 erros, e no sentido inverso com 12 erros, conforme a matriz confusão de apresentada na Figura 19. Essa técnica apresentou uma redução de 0,5% na acurácia, finalizando o treinamento em 67,06 segundos, uma redução de quase 5 minutos para o treino.

Figura 17: Matriz de confusão para *boosted trees* (Grupo A8).

Fonte: Autor (2025).

Figura 18: Matriz de confusão para RNA média (Grupo A8).

Fonte: Autor (2025).

Apresentam-se na Tabela 9 os resultados comparativos entre as três técnicas ao grupo A8. Notou-se que diminuição do número de parâmetros implicou em uma redução inferior a 1% na acurácia. No entanto, houve uma significativa redução no tempo de processamento em todos os modelos. Tal resultado melhora as condições de

implementação das técnicas, tanto no que diz respeito à velocidade de classificação quanto ao aspecto econômico, possibilitando uma redução do esforço computacional.

Tabela 9: Resultados para o treinamento do Grupo A8.

Modelo	Acurácia (%)	Tempo de simulação (s)
<i>Fine Tree</i>	99,7	2,38
<i>Boosted Trees</i>	99,7	12,76
<i>Medium Neural Network</i>	98,6	81,64

Fonte: Autor (2025).

Além do treinamento, também foram realizados os testes do grupo A8, cujos resultados estão apresentados na Tabela 10. Nela, observa-se a acurácia de 99,7% para as árvores finas, que apresentaram no máximo 5 erros na previsão das classes. Para as *boosted trees*, a acurácia reduziu para 99,6%, em comparação ao teste com 11 atributos, enquanto a RNA média manteve a acurácia em 99,1%. Dessa forma, mesmo com dados diferentes do treinamento, o classificador para o grupo A8 manteve uma acurácia elevada.

Tabela 10: Resultados para os testes do Grupo A8.

Modelo	Acurácia (%)
<i>Fine Tree</i>	99,7
<i>Boosted Trees</i>	99,6
<i>Medium Neural Network</i>	99,1

Fonte: Autor (2025).

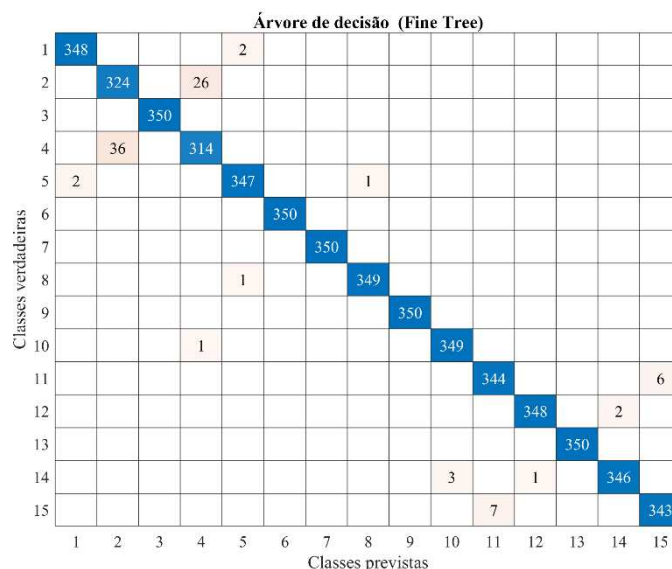
4.2.1.2 Treinamento e testes para o Grupo A5

O mesmo processo foi repetido para o grupo A5, composto por cinco parâmetros, visando analisar a redução do tempo com relação aos grupos anteriores. Ademais, essa avaliação teve como objetivo verificar se, em termos de aprendizagem, as técnicas são capazes de manter uma boa margem de acerto e reduzir o viés financeiro do projeto, caso venha a ser implementado.

Para a árvore de decisão, foi obtida uma acurácia de 98,3% com um tempo de processamento de 6,58 segundos. Nesse caso, houve uma redução de 1,4% na acurácia e um ganho de mais de 4 segundos no tempo de processamento, evidenciando que a redução demasiada dos parâmetros pode dificultar a classificação dos distúrbios, além de aumentar o tempo de treinamento. Ademais, como pode ser observado na Figura 20, a matriz de confusão enfatiza o aumento de previsões incorretas, destacando-se a classe afundamento (classe prevista) confundida com a classe interrupção (classe real), com 36

erros, e no sentido oposto, com 26 erros. Observa-se, então, um aumento significativo no número de erros mesmo em sinais com distúrbios não combinados.

Figura 19: Matriz de confusão para o treino da *fine tree* (Grupo A5).

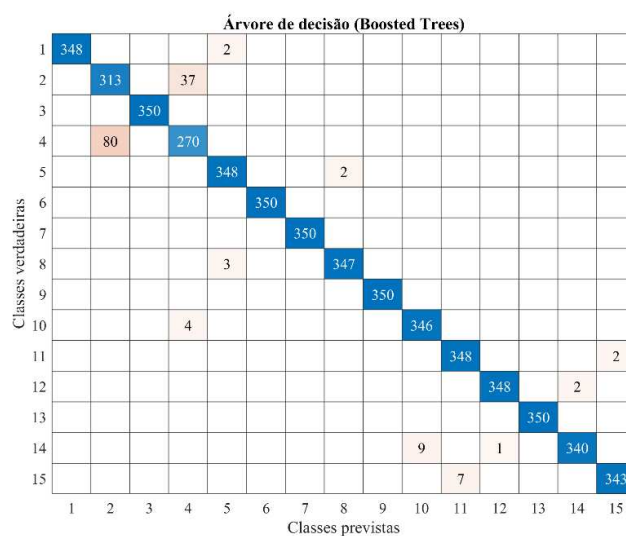


Fonte: Autor (2025).

O modelo *boosted trees*, por sua vez, apresentou uma acurácia de 97,2%, com o tempo de processamento de 14,51 segundos. Com relação às previsões incorretas, mais uma vez destacou-se a classe interrupção (classe real) com a classe afundamento (classe prevista), somando 80 erros, conforme mostra a matriz de confusão da Figura 21. Esse resultado evidencia que a redução dos parâmetros não foi benéfica para o treinamento do classificador.

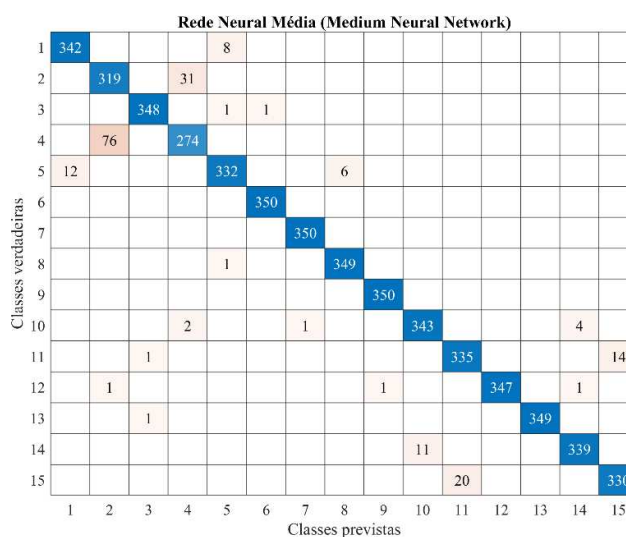
Por fim, a RNA média obteve uma acurácia de 96,6%, uma redução de 2% quando comparada ao grupo A8, e um tempo de processamento de 73,87 segundos, que representa um ganho de aproximadamente 10 segundos quando comparada ao grupo anterior. No entanto, essa técnica apresentou diversas previsões erradas, somando 12 classes previstas indevidamente, com destaque mais uma vez para as classes afundamento (prevista) e interrupção (real), como pode ser observado na matriz de confusão da Figura 22.

Figura 20: Matriz de confusão para o treino da *Boosted trees* (Grupo A5).



Fonte: Autor (2025).

Figura 21: Matriz de confusão para o treino da RNA (Grupo A5).



Fonte: Autor (2025).

Na Tabela 11 estão resumidas as informações referentes ao treinamento e aos testes do grupo A5. Nela observa-se, além dos dados já comentados, a redução da acurácia no grupo de testes quando comparada com o grupo A5.

Tabela 11: Resultados para o treinamento e testes do Grupo A5.

Modelo	Treinamento		Testes
	Acurácia (%)	Tempo de simulação (s)	Acurácia (%)
<i>Fine Tree</i>	98,3	6,58	98,3
<i>Boosted Trees</i>	97,2	14,51	97,2
<i>Medium RNA</i>	96,6	70,64	96,6

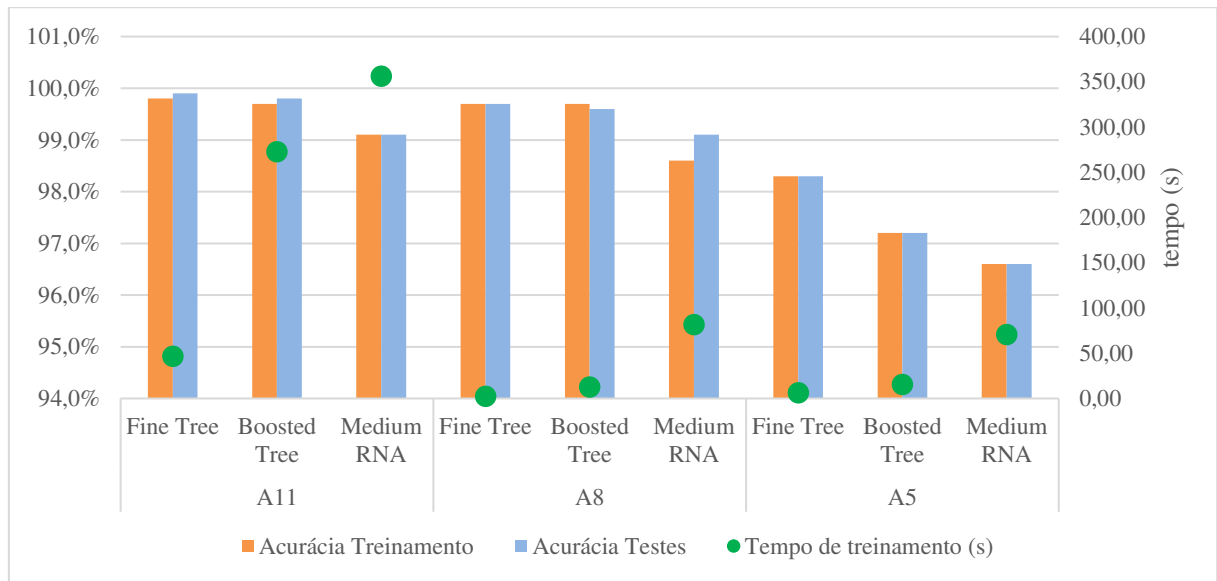
Fonte: Autor, 2025.

4.2.2 Escolha do modelo de classificador

A fim de sintetizar todas as informações dos testes realizados, apresenta-se o gráfico da Figura 23, na qual observam-se os dados de acurácia para treinamento e testes nos três grupos e o tempo de treinamento necessário para cada um deles.

Nota-se que, para os três grupos, a técnica da árvore de decisão fina (*fine tree*) apresentou os melhores resultados, com acurácia superior as demais tanto no treinamento quanto nos testes. Ademais, nota-se que a redução de 11 para 8 parâmetros foi, de fato, benéfica para o modelo, resultando em uma redução significativa no tempo de treinamento em todos os casos. A redução para 5 parâmetros, no entanto, não apresentou vantagens para o classificador, diminuindo a acurácia e, por vezes, aumentando o tempo de processamento.

Portanto, com base em todos os dados, conclui-se que o modelo *fine tree* com oito parâmetros classificadores apresentou os melhores resultados, com acurácia de 99,7% , tanto no treinamento quanto nos testes, e um tempo de processamento de apenas 2,38 segundos, evidenciando sua adequação à aplicação proposta.

Figura 22: Comparativo de resultados para os três grupos de teste.

Fonte: Autor (2025).

5 CONCLUSÕES

O presente trabalho teve como objetivo principal desenvolver um classificador de distúrbios de QEE utilizando técnicas de aprendizado de máquina, baseada nos dados sintéticos fornecidos pelo IEEE *Dataport*.

Ao longo do estudo, foram analisados diversos distúrbios de QEE, como transitórios oscilatórios e impulsivos, variações de tensão de curta duração, distorções harmônicas, flutuações de tensão, *spike* e *notching*, tanto no domínio do tempo quanto no domínio da frequência. A partir dessas análises, foram selecionados atributos característicos de cada distúrbio, como valor médio, valor máximo, amplitude, número de picos, harmônicos específicos, frequência fundamental, dentre outros. Esses atributos serviram como entrada para os modelos de aprendizagem de máquina.

Diferentes algoritmos de classificação foram testados, incluindo árvores de decisão, florestas aleatórias, SVM e RNA. Dentre esses, a árvore de decisão fina (*fine tree*) destacou-se como o modelo mais eficiente e eficaz, alcançando uma acurácia de 99,8% no treinamento e 99,9% nos testes, com um tempo de processamento significativamente menor em comparação aos demais modelos. O modelo combinado de árvores chamado *boosted trees* e a RNA média também apresentaram acurácia elevada, porém com tempos de processamento mais elevados.

A redução do número de parâmetros foi explorada com o intuito de otimizar o tempo de processamento e reduzir custos computacionais. A análise mostrou que a redução para oito parâmetros manteve uma alta acurácia (99,6% para a árvore fina) enquanto reduziu consideravelmente o tempo de processamento. No entanto, a redução para cinco parâmetros resultou em uma queda significativa na acurácia, inviabilizando sua utilização. Em suma, os resultados demonstraram que a utilização de técnicas de aprendizagem de máquina, em particular a árvore de decisão fina, é viável e eficaz para a classificação de distúrbios de QEE.

A escolha adequada dos parâmetros e a otimização do modelo permitiram alcançar resultados satisfatórios, tornando-o uma solução promissora para a detecção e classificação de distúrbios em sistemas elétricos. Ademais, essa linha de pesquisa é promissora e apta para evolução. Nesse sentido, como sugestão de trabalhos futuros, pode-se explorar a aplicação desses modelos em dados reais, bem como a integração com sistemas de monitoramento em tempo real, visando a melhoria contínua da qualidade da energia elétrica.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, A.; MOREIRA, L.; DELGADO, J. **Power Quality problems and new solutions” International conference on renewable energies and power quality. 03,** Vigo, Spain Vol.1, 2003.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST: Módulo 1 - Glossário de termos técnicos do PRODIST.** Brasília: ANEEL, 2021a. Revisão 11.

ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST: Módulo 8 – Qualidade do fornecimento de energia elétrica.** Brasília: ANEEL, 2021. Revisão 13.

ARRILLAGA, J.; WATSON, N. R. **Power System Harmonics.** 2. ed. Chichester: John Wiley & Sons, 2003.

BISHOP, Christopher M. **Pattern Recognition and Machine Learning.** New York: Springer, 2006.

BOLLEN, M. H. J.; RIBEIRO, P. F.; GU, I. Y. H.; DUQUE, C. A. **“Trends, challenges and opportunities in power quality research”**, European Transactions on Electrical Power, v. 20, pp. 3–18, 2009.

BORGES, Caio Franco. **Análise da qualidade de energia elétrica em um sistema fotovoltaico conectado a rede.** Trabalho de conclusão de curso (Bacharelado em Engenharia Elétrica), Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Itumbiara, 2019.

BRACEWELL, R. N. **The Fourier Transform and Its Applications.** 3. ed. New York: McGraw-Hill, 2000.

BREIMAN, Leo; friedman, Jerome; olshen, Richard; stone, Charles. **Classification and Regression Trees**. Belmont: Wadsworth International Group, 1984.

BREIMAN, Leo. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5-32, 2001.

CASAZZA, John; delea, Frank. **Understanding Electric Power Systems: An Overview of the Technology and the Marketplace**. 2. ed. Hoboken: John Wiley & Sons, 2010.

COSTA, Luisa. **SISTEMA DE IDENTIFICAÇÃO DE DIREÇÃO PERIGOSA UTILIZANDO ACELERAÇÃO E VELOCIDADE ANGULAR COLETADOS PELO MPU-6050**. 2023. Trabalho de Conclusão de Curso. Graduação em CURSO DE BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO — Universidade Federal de Pernambuco, Recife, 2023.

CENELEC. *EN 50160:2022 - Voltage characteristics of electricity supplied by public electricity networks*. Bruxelas: CENELEC, 2022.

DECKMANN, S. M.; POMILIO, J. A. **Qualidade da Energia Elétrica: Conceituação e Processamento Digital**. São Paulo: Edgard Blücher Ltda., 2024.

DING, C.; PENG, H. **Minimum redundancy feature selection from microarray gene expression data**. *Journal of Bioinformatics and Computational Biology*, v. 3, n. 2, 2005.

DUGAN, Roger C.; MCGRANAGHAN, Mark F.; BEATY, Wayne H.; SANTOSO, Surya. **Electrical power systems quality**. 2.ed. New York, NY: McGraw-Hill, 2003.

FORTES, Rárisson Roberto Acácio. **Distorções harmônicas produzidas por inversores de geração distribuída fotovoltaica em condição de ressonância paralela**. Dissertação (Mestrado em Automação) – Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2016.

FORTUNATO, Laís M. A. Análise nos domínios do tempo e da frequência dos distúrbios de Qualidade da Energia Elétrica. TCC. Graduação em Engenharia Elétrica. Universidade Federal de Campina Grande. 2016.

FROES, Filipe. **Comparação de modelos de desenvolvimento de sensores virtuais baseados em inteligência artificial**. 2022. Trabalho de Conclusão de Curso (Graduação em Engenharia) – Centro de Energias Alternativas e Renováveis, Universidade Federal da Paraíba, João Pessoa, 2022.

GLOVER, J. Duncan; SARMA, Mulukutla S.; OVERBYE, Thomas. *Power System Analysis and Design*. 5. ed. Stamford: Cengage Learning, 2016.

GONÇALVES, F. A. S.; MARTINS, H. J. A. **Qualidade da Energia Elétrica: Conceitos e Tópicos**. 2. ed. São Paulo: Editora Manole, 2018.

GRIGSBY, Leonard. **The Electric Power Engineering Handbook**. 2. ed. Boca Raton: CRC Press, 2007.

HAYKIN, Simon. **Redes Neurais: Princípios e Práticas**. 2. ed. Porto Alegre: Bookman, 2001.

HASTIE, Trevor.; Tibshirani, Robert; Friedman, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. Nova York: Springer, 2009.

IBM. **O que é aprendizado não supervisionado**. Disponível em: <https://www.ibm.com/br-pt/topics/unsupervised-learning>. Acesso em: 15 de março 2025

IEEE, 1159. Instituto de Engenheiros Eletricistas e Eletrônicos, **Recommended practice for monitoring electric power quality**. (IEEE Standard 1159). New York: [s.n.], 2019.

IEC. *Compatibilidade Eletromagnética (EMC) - Parte 3*. Genebra: **International Electrotechnical Commission**, 2023.

KUSKO, A.; THOMPSON, M. T. *Power Quality in Electrical Systems*. New York: McGraw-Hill, 2007.

LAROCHELLE, Hugo. **Learning Algorithms for the Classification Restricted Boltzmann Machine**. *Journal of Machine Learning Research*, v. 13, p. 643–669, 2012.

LIMA, Érica Manguiera. **Detecção e localização de faltas de alta impedância baseadas na transformada de Stockwell**. 2021. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, Campina Grande, 2021.

MARIANO, Juliana D'Angela; URBANETZ JUNIOR, Jair. **Energia Solar Fotovoltaica: Princípios fundamentais**. Ponta Grossa - PR: Atena, 2022.

MACEDO, Júnior. **Variações de tensão de curta duração – Parte 1**. Faraday Engenharia. Disponível em: <http://www.faraday.com.br/tutorial-variacoes-de-tensao-de-curta-duracao-parte-1/>. 2017. Acesso em: 13 Março 2025.

MONTGOMERY, Douglas; Peck, Elizabeth; Vining, Geoffrey. **Introduction to Linear Regression Analysis**. 4. ed. Hoboken: John Wiley & Sons, 2012.

MURPHY, Kevin P. **Machine learning: a probabilistic perspective**. Cambridge, MA: MIT Press, 2012.

NOGARE, Diego. **Performance de Machine Learning: Matriz de Confusão**. 2020. Disponível em: <https://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>. Acesso em: 01 de Março 2025.

PAULILLO, Gilson. TEIXEIRA, Mateus. BACCA, Ivandro. FILHO, José. O Setor Elétrico **Capítulo VI Variações de tensão de curta duração - Parte II**. Disponível em: https://www.osetoreletrico.com.br/wp-content/uploads/2013/07/ed-89_Fasciculo_Cap-VI-Qualidade-de-energia.pdf.

PAULILLO, Gilson. TEIXEIRA, Mateus. **Capítulo VII Variações de tensão de curta duração - Parte II.** Disponível em: https://www.osetoreletrico.com.br/wp-content/uploads/2013/09/ed-90_Fasciculo_Cap-VII-Qualidade-de-energia.pdf

PEREIRA, Henrique. **Aula 07 – Elevação e Desequilíbrio de Tensão. Departamento de Engenharia Elétrica** – Universidade Federal Viçosa, 2017.

ROCHA, J.E. **Qualidade da Energia Elétrica. Departamento de Eletrotécnica.** Universidade Tecnológica Federal do Paraná. Curitiba, 2016.

ROCHA, Júlio. **Uma introdução à Decision Tree.** *LinkedIn*, [S. l.], 19 ago. 2021. Disponível em: <https://www.linkedin.com/pulse/uma-introdu%C3%A7%C3%A3o-decision-tree-julio-cesar-pereira-rocha/>. Acesso em: 20 mar. 2025.

RUSSELL, Stuart; norvig, Peter. **Artificial Intelligence: A Modern Approach.** 4. ed. São Paulo: Pearson, 2020.

STAROSTA, R. **Power Quality in Electrical Systems.** New York: McGraw-Hill, 2011

VAPNIK, Vladimir. **The Nature of Statistical Learning Theory.** 2. ed. Berlim: Springer, 2000.

ZHAO, Z.; ANAND, R.; WANG, M. **Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform.** In: 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2019.