

UNIVERSIDADE FEDERAL DO PARANÁ

LEANDRO AUGUSTO ENSINA

FRAMEWORK PARA ANÁLISE DE FALTAS EM LINHAS DE TRANSMISSÃO:  
UM MÉTODO BASEADO EM MODELOS MÚLTIPLOS PREDITIVOS

CURITIBA PR

2024

LEANDRO AUGUSTO ENSINA

FRAMEWORK PARA ANÁLISE DE FALTAS EM LINHAS DE TRANSMISSÃO:  
UM MÉTODO BASEADO EM MODELOS MÚLTIPLOS PREDITIVOS

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciência da Computação no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Prof. Dr. Luiz Eduardo Soares de Oliveira.

CURITIBA PR

2024

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
UNIVERSIDADE FEDERAL DO PARANÁ  
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Ensina, Leandro Augusto

Framework para análise de faltas em linhas de transmissão: um método baseado em modelos múltiplos preditivos / Leandro Augusto Ensina. – Curitiba, 2024.

1 recurso on-line: PDF.

Tese (doutorado) – Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Informática. Defesa: Curitiba, 19/02/2024.

Orientador: Prof. Dr. Luiz Eduardo Soares de Oliveira

1. Inteligência artificial. 2. Energia elétrica – Qualidade. I. Oliveira, Luiz Eduardo Soares de. II. Título

CDD 006.3



MINISTÉRIO DA EDUCAÇÃO  
SETOR DE CIÊNCIAS EXATAS  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA -  
40001016034P5

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **LEANDRO AUGUSTO ENSINA** intitulada: **FRAMEWORK PARA ANÁLISE DE FALTAS EM LINHAS DE TRANSMISSÃO: UM MÉTODO BASEADO EM MODELOS MÚLTIPLOS PREDITIVOS**, sob orientação do Prof. Dr. LUIZ EDUARDO SOARES DE OLIVEIRA, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 19 de Fevereiro de 2024.

Assinatura Eletrônica

26/02/2024 06:46:57.0

LUIZ EDUARDO SOARES DE OLIVEIRA  
Presidente da Banca Examinadora

Assinatura Eletrônica

27/02/2024 10:19:05.0

HUEI DIANA LEE  
Avaliador Externo (UNIVERSIDADE ESTADUAL DO OESTE DO  
PARANÁ)

Assinatura Eletrônica

23/02/2024 14:11:42.0

EDUARDO CUNHA DE ALMEIDA  
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

28/02/2024 11:41:52.0

RAFAEL MENELAU OLIVEIRA E CRUZ  
Avaliador Externo (ÉCOLE DE TECHNOLOGIE SUPÉRIEURE)

*À minha família...*

## AGRADECIMENTOS

Agradeço primeiramente a Deus e aos meus pais, Marlene e Gerson, por todo incentivo, confiança e apoio incondicional ao longo destes anos. Agradeço também ao meu irmão Gustavo e à minha namorada Kelly pelo apoio, paciência (principalmente) e auxílio para que eu pudesse me dedicar ao projeto. Vocês foram fundamentais, muito obrigado por acreditarem em mim!

Agradeço ao professor Luiz Oliveira pelas inúmeras oportunidades, amizade, confiança e ensinamentos ao longo do doutorado. Fico honrado e feliz pelo privilégio de ser seu aluno. Sou imensamente grato pela generosidade e pelo tempo que o senhor dedicou à minha orientação, zelando pela ética e pela excelência técnica, acadêmica e científica.

Desde o trabalho de conclusão do curso de graduação faço questão de sempre agradecer aos professores Fabio Spanhol, Wesley Klewerton Assunção, Alexandre Huff e Vilson Dalle Mole que foram essenciais para que eu pudesse chegar até aqui. Antes de ingressar na graduação, ainda no final do curso técnico na UTFPR/Toledo, os professores foram pacientes e atenciosos ao discutir minhas dúvidas sobre os cursos de graduação que eu pensava seguir, propiciando para que eu pudesse fazer a melhor escolha para aquilo que eu almejava. E graças a eles, acredito que fiz, sim, a melhor escolha.

Agradeço aos professores e amigos Huei Diana Lee e Wu Feng Chung por toda orientação, amizade, ensinamentos e oportunidades que me proporcionaram desde o início desta trajetória, ainda na graduação. Inspirado em vocês, decidi seguir o caminho da pesquisa e do ensino, muito obrigado!

Aos amigos dos laboratórios LABI/UNIOESTE e HiPES/UFPR pelas experiências e conhecimentos compartilhados, em especial aos professores e amigos Newton Spolaôr, Weber Takaki, Marco Zanata e Eduardo Almeida pelas trocas de ideias, suporte e amizade, bem como a confiança depositada em mim para a colaboração em diversos trabalhos.

Ao meu grande amigo César Augusto pela amizade, incentivo e pelos momentos de descontração ao longo das conversas que certamente me ajudaram bastante.

À Ivonete Borne, ao Oswaldo Hideo Ando Jr. e ao Pietro por todo apoio, incentivo, compreensão e amizade, obrigado por tornarem-se esse período mais leve.

À minha avó Hedy e aos meus demais familiares que me apoiaram neste período.

Aos novos e “velhos” amigos da COGETI e da COENC da UTFPR/Toledo, que tive o prazer de conhecer e rever no período que estive como professor substituto na instituição. Em particular aos professores e amigos Maurício Zardo, Elder Schemberger, Daniel Cavalcanti e Lígia Mendes pelo apoio incondicional pela minha conciliação das aulas com o doutorado.

Agradeço aos professores Luiz Oliveira, George Cavalcanti e Rafael Cruz pela oportunidade na realização do período sanduíche na ÉTS/Montreal. Ao professor Rafael, agradeço ainda por me receber tão bem e dedicar seu tempo para me ensinar e orientar. Agradeço ainda aos novos amigos Juscimara Avelino, Cristiano Garcia e Heitor Rapela que o período sanduíche proporcionou, obrigado pela amizade, pelos passeios e pelo apoio pessoal!

Ao Vincent Mercier pela amizade e pela excelente hospedagem em sua casa em Montreal durante minha estadia no Canadá. Vincent, mercie!

Agradecimentos ainda a COPEL (projeto P&D 06491-0407/2015) e ao CNPq (projeto 401992/2022-5) pelas bolsas concedidas durante este projeto.

## RESUMO

As linhas de transmissão são elementos cruciais do sistema elétrico de potência, transportando a energia gerada nas usinas até subestações próximas aos centros consumidores. Para isso, essas linhas percorrem longas distâncias e, conseqüentemente, estão suscetíveis a diversos tipos de perturbações que afetam o seu estado normal de operação, como a incidência de descargas atmosféricas e queimadas sob as linhas. Tais perturbações são conhecidas como faltas, as quais ocorrem de modo imprevisível em qualquer ponto da linha de transmissão, resultando em fluxos anormais nas grandezas elétricas nos componentes do sistema de potência. Os efeitos de uma falha são preocupantes, uma vez que podem acarretar na interrupção do fornecimento de energia elétrica e em danos aos equipamentos e componentes envolvidos na transmissão. Frente a isso, a análise de faltas torna-se indispensável, com destaque para duas tarefas: a classificação do tipo da falha e a sua localização. Portanto, este trabalho tem como objetivo propor e avaliar um método para classificar e localizar faltas em linhas aéreas de transmissão a partir de algoritmos de aprendizado de máquina, buscando alcançar duas propriedades: flexibilidade de taxa de amostragem e generalização para linhas com diferentes configurações. Sobretudo, exploramos o uso de modelos múltiplos preditivos (*ensemble*) para a realização de ambas as tarefas devido a lacuna identificada sobre o tema. Especialmente para a tarefa de localização, propomos uma nova estrutura de seleção dinâmica de regressores, a qual visa selecionar apenas os modelos mais competentes individualmente para cada novo padrão de teste de modo a tornar as predições mais precisas e acuradas. Devido a indisponibilidade de bases públicas e a escassez de eventos reais de falhas, desenvolvemos e disponibilizamos publicamente dois novos conjuntos de dados com simulações diversificadas de faltas. Os resultados alcançados demonstram a versatilidade e a efetividade do nosso método de classificação para ambas as propriedades desejadas, validado inclusive a partir de eventos reais de falhas com acurácia em torno de 96%. Neste quesito, a engenharia de atributos que realizamos foi fundamental para que pudéssemos atingir nosso objetivo, além, claro, do emprego do algoritmo de Árvores Extremamente Aleatórias usado para classificar o tipo de falha. Por sua vez, o método de localização também demonstrou desempenhos satisfatórios especialmente para a flexibilidade perante diferentes taxas de amostragem, apesar da leve vantagem para as maiores frequências. Todavia, não foi possível alcançar a propriedade da generalização para a localização. A provável razão dessa inaptidão está associada as diferentes assinaturas de falhas entre as linhas por conta de suas extensões distintas, mesmo que parâmetros como resistência e tipo de falha sejam semelhantes. Apesar disso, a estrutura de seleção dinâmica que propomos manifestou desempenhos estatisticamente superiores frente a outras abordagens.

Palavras-chave: diagnóstico de falha; inteligência artificial; qualidade de energia elétrica; reconhecimento de padrões; sistemas inteligentes.



## ABSTRACT

The transmission lines are crucial elements of the electric power system, transporting the energy generated in the power plants to substations near the customers. As a result, these lines travel long distances and, consequently, are exposed to several kinds of disturbances that affect their normal operating state, such as the incidence of atmospheric discharges and fire under the lines. Such disturbances are known as faults, which can occur unpredictably at any point of the transmission line, resulting in abnormal flows in the electrical properties of the power system components. The effects of a failure are a considerable concern once they can lead to interruption of electricity supply and damage to equipment and components involved in the power transmission. So, fault analysis becomes a fundamental activity for the protection system, emphasizing two tasks: fault type classification and fault location. Therefore, this work aims to propose and evaluate a method to classify and locate faults in transmission lines by using machine learning algorithms, considering the following properties: flexibility of sampling rate and generalization capacity to transmission lines with different configurations. Above all, we explored the use of ensemble models to carry out both tasks due to the gap identified on the topic. Especially for the localization task, we propose a new dynamic regressor selection framework, which aims to select only the most competent models individually for each new test pattern in order to achieve more precise and accurate predictions. Due to the unavailability of public databases and the scarcity of real fault events, we developed and made two new datasets with diverse fault simulations publicly available to the scientific community. The results demonstrated the versatility and effectiveness of our classification method for both desired properties, validated even based on real failure events with an accuracy of around 96%. In this scenario, the feature engineering that we carried out, in addition to using the Extremely Randomized Trees algorithm to classify the fault type, was fundamental to achieving our objective. In turn, the localization method also demonstrated satisfactory performances, especially for flexibility with different sampling rates, despite the slight advantage for higher frequencies. However, it was not possible to achieve the generalization property for localization. The probable reason for this inability is the different fault signatures between the lines due to their different lengths, even if parameters such as resistance and fault type are the same. Nonetheless, the proposed dynamic selection structure demonstrated statistically superior performance compared to other approaches.

**Keywords:** artificial intelligence; electric power quality; fault diagnosis; intelligent systems; pattern recognition.



## LISTA DE FIGURAS

1.1	Visão geral da estrutura do SEP . . . . .	18
2.1	Principais componentes de uma linha de transmissão . . . . .	25
2.2	Exemplo de uma falta real do tipo AG com início logo após o tempo 0,08 segundo. Na imagem (a) está representado o sinal da corrente, enquanto na imagem (b) está retratada a forma de onda da tensão . . . . .	27
2.3	Classificação das faltas em linhas de transmissão . . . . .	28
2.4	Diagrama simplificado das ramificações do AM . . . . .	34
2.5	Exemplo de uma árvore de decisão junto as regiões de decisão no espaço de atributos ( $x^1$ e $x^2$ ) . . . . .	35
2.6	Representação da abordagem convencional e das principais técnicas <i>ensemble</i> . . . . .	37
2.7	Representações genéricas das estruturas para a seleção (a) estática e (b) dinâmica de preditores, sejam eles classificadores ou regressores . . . . .	38
2.8	Representação do método <i>Stacking</i> . . . . .	39
2.9	Representação do algoritmo ERT. . . . .	40
2.10	Ilustração do algoritmo KNN para dois tamanhos de vizinhança ( $K = 3$ e $K = 7$ ). Considere a intensidade das cores dos círculos como uma representação de seus valores de atributo-alvo para a regressão, de modo que quanto mais escura a cor azul, maior é o seu valor numérico . . . . .	42
3.1	Diagrama simplificado das principais atividades de análise de faltas . . . . .	46
4.1	Sistema IEEE 9-barras utilizado para a geração dos dados de falta . . . . .	60
4.2	Exemplo de dois ciclos pós-falta demarcados em uma simulação de falta do tipo ABG para os sinais de (a) corrente e (b) tensão . . . . .	64
5.1	Ambiente de operação do método proposto para o diagnóstico de falhas em linhas de transmissão. . . . .	66
5.2	Fluxograma do componente proposto para a classificação de falta. . . . .	67
5.3	Fluxograma do componente proposto para a localização de falta . . . . .	69
5.4	Visão geral da estrutura SDR proposta . . . . .	70

6.1	Representação do processo de normalização considerando um evento de falha ABG. Na imagem (a) está retratado um trecho de uma simulação original do conjunto LT1, representando o estado normal de operação ( <i>i.e.</i> , sinal pré-falta) com correntes simétricas e equilibradas até o início da falha, a qual ocorre por volta do tempo 0,041 s. Conforme pode ser observado, a amplitude da corrente da forma de onda pré-falta permanece entre aproximadamente -300 A e 300 A. Portanto, o processo da normalização visa converter a oscilografia para uma escala menor, conforme apresentada pela imagem (b), porém conservando o comportamento pós-falta que excede a amplitude do nível do estado normal de operação, a qual após a normalização encontra-se no intervalo [0, 1]. Consequentemente, oscilografias com amplitudes distintas da primeira imagem (a) resultarão em amplitudes equivalentes a segunda imagem (b), permitindo que o método lide com qualquer oscilografia de maneira equivalente . . . . .	82
6.2	Sumarização dos resultados para os diversos cenários avaliativos realizados neste trabalho ( <i>ablation study</i> ) para a taxa de amostragem de 10 kHz . . . . .	87
6.3	Sinal trifásico da corrente para o evento de falha #7 . . . . .	89
6.4	Importância dos atributos para o algoritmo ERT . . . . .	95
6.5	Representação dos erros de predição para a frequência de 10 kHz. . . . .	102
6.6	Exemplo da dispersão das predições (a) sem o uso e (b) com a aplicação do módulo de filtro . . . . .	106
6.7	Representação de falhas para as LTs de (a) 500 kV e (b) 230 kV, cujos parâmetros de falta são iguais: tipo ABG, resistência de 50 $\Omega$ e ocorrência a 50% da extensão da linha . . . . .	113

## LISTA DE TABELAS

2.1	Levantamento das ocorrências de faltas . . . . .	29
3.1	Informações sumarizadas dos trabalhos selecionados para o estado da arte (parte 1)	52
3.2	Informações sumarizadas dos trabalhos selecionados para o estado da arte (parte 2)	54
4.1	Resumo dos atributos normalmente criados e usados para análise de falhas por trabalhos relacionados (em negrito estão os atributos propostos em nosso trabalho)	65
5.1	Hiperparâmetros da estrutura SDR proposta (em negrito valores utilizados como padrão neste trabalho). . . . .	73
6.1	Eventos reais de faltas em linhas de transmissão da rede da COPEL coletados entre Janeiro/2019 e Março/2020 . . . . .	77
6.2	Desempenho do método proposto para a classificação do tipo de falha (conjunto de validação) . . . . .	78
6.3	Desempenho do método proposto para a taxa de amostragem de 256 Hz (conjunto de teste) . . . . .	79
6.4	Desempenho do método proposto para a taxa de amostragem de 1 kHz (conjunto de teste) . . . . .	79
6.5	Desempenho do método proposto para a taxa de amostragem de 5 kHz (conjunto de teste) . . . . .	79
6.6	Desempenho do método proposto para a taxa de amostragem de 10 kHz (conjunto de teste) . . . . .	80
6.7	Média e desvio padrão da acurácia para a classificação do tipo de falha utilizando apenas os atributos de máximos e mínimos. . . . .	81
6.8	Média e desvio padrão da acurácia para a classificação do tipo de falha com a inclusão da etapa de normalização . . . . .	81
6.9	Matriz de confusão do experimento nº 4 para a taxa de amostragem de 1 kHz . . . . .	83
6.10	Média e desvio padrão da acurácia para a classificação do tipo de falha com a etapa de normalização e com a inclusão do atributo <i>groundDetection</i> . . . . .	84
6.11	Desempenho do método proposto para a taxa de amostragem de 256 Hz (conjunto LT2) . . . . .	85
6.12	Desempenho do método proposto para a taxa de amostragem de 1 kHz (conjunto LT2) . . . . .	85
6.13	Desempenho do método proposto para a taxa de amostragem de 5 kHz (conjunto LT2) . . . . .	85
6.14	Desempenho do método proposto para a taxa de amostragem de 10 kHz (conjunto LT2) . . . . .	86
6.15	Acurácia para a classificação do tipo de falta para eventos reais . . . . .	88

6.16	Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #20 do terminal 2. Os valores entre parênteses representam as probabilidades . . . . .	90
6.17	Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #21 do terminal 2. Os valores entre parênteses representam as probabilidades . . . . .	91
6.18	Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #29 do terminal 1. Os valores entre parênteses representam as probabilidades . . . . .	91
6.19	Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #30 do terminal 1. Os valores entre parênteses representam as probabilidades . . . . .	91
6.20	Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #21 do terminal 2 com o valor do atributo <i>groundDetection</i> estabelecido com o valor zero (0) – indicativo que não houve atuação da terra. Os valores entre parênteses representam as probabilidades	92
6.21	Importância dos atributos para o algoritmo ERT usado no método proposto para classificação de falhas. . . . .	93
6.22	Média e desvio padrão da acurácia para a classificação do tipo de falta para os algoritmos avaliados (conjunto de validação LT1) . . . . .	96
6.23	Média e desvio padrão da acurácia para a classificação do tipo de falta para os algoritmos avaliados (conjunto LT2) . . . . .	96
6.24	Comparação dos desempenhos pela acurácia para a classificação de falta para o conjunto de teste LT1 . . . . .	97
6.25	Comparação dos desempenhos pela acurácia para a classificação de falta para o conjunto de teste LT2 . . . . .	97
6.26	Comparação dos desempenhos para classificação de falta incluindo o processo de normalização para a capacidade de generalização (conjunto LT2) . . . . .	99
6.27	Média e desvio padrão dos erros em km para a localização de falha (LT1) . . . .	101
6.28	Distribuição dos erros para diferentes intervalos, considerando a frequência de 10 kHz. Os valores representam a média e o desvio padrão do número de ocorrências por intervalo ao longo das dez repetições . . . . .	102
6.29	Média e desvio padrão dos resultados em km comparando diferentes cenários de normalização . . . . .	103
6.30	Média e desvio padrão dos resultados em km comparando as diferentes métricas de distância . . . . .	104
6.31	Média e desvio padrão dos resultados em km comparando os diferentes critérios de seleção . . . . .	105
6.32	Média e desvio padrão dos resultados em km referentes ao emprego do módulo de filtro . . . . .	105
6.33	Listagem dos dez principais atributos ordenados por importância para o algoritmo ERT, considerando a média e o desvio padrão da medida de Gini . . . . .	107

6.34	Média e desvio padrão dos resultados em km para os métodos relacionados e o método proposto . . . . .	108
6.35	Média e desvio padrão dos resultados em km para cada algoritmo base . . . . .	109
6.36	Média e desvio padrão dos resultados em km para cada algoritmo <i>ensemble</i> . . .	109
6.37	Média e desvio padrão dos resultados em km para cada método SDR apenas para 10 kHz. . . . .	110
6.38	Média e desvio padrão dos erros em km para a capacidade de generalização (conjunto LT2) . . . . .	112
B.1	Características ordenadas por suas importâncias para o algoritmo ERT, considerando a média e o desvio padrão, entre parênteses, da medida de Gini calculada a partir dos dez modelos ERT pré-treinados . . . . .	133

## LISTA DE ACRÔNIMOS

AM	Aprendizado de Máquina
ATP	<i>Alternative Transient Program</i>
AUC	<i>Area Under the Curve</i>
Bi-GRU	<i>Bidirectional Gated Recurrent Unit</i>
CART	<i>Classification and Regression Tree</i>
CNN	<i>Convolutional Neural Network</i>
COPEL	Companhia Paranaense de Energia
CSAE	<i>Convolutional Sparse Autoencoder</i>
DFT	<i>Discrete Fourier Transform</i>
DL	<i>Deep Learning</i>
DS	<i>Dynamic Selection</i>
DW	<i>Dynamic Weighting</i>
DWS	<i>Dynamic Weighting Selection</i>
DWS-MS	<i>Dynamic Weighting Selection with Measure Selection</i>
DWT	<i>Discrete Wavelet Transform</i>
ELM	<i>Extreme Learning Machine</i>
ERT	<i>Extremely Randomized Trees</i>
FADb	<i>Fault Analysis Database</i>
FADbF	<i>Fault Analysis Database with Features</i>
FN	Falso Negativo
FP	Falso Positivo
GAF	<i>Gramian Angular Field</i>
IA	Inteligência Artificial
ID3	<i>Iterative Dichotomiser 3</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
KNN	<i>k-Nearest Neighbors</i>

LightGBM	<i>Light Gradient Boosting Machine</i>
LG	Falta fase-terra ( <i>line to ground</i> )
LL	Falta fase-fase ( <i>line to line</i> )
LLG	Falta fase-fase-terra ( <i>double line to ground</i> )
LLL	Falta fase-fase-fase ( <i>triple line</i> )
LLLG	Falta fase-fase-fase-terra ( <i>triple line to ground</i> )
LSTM	<i>Long Short-Term Memory</i>
LT	Linha de Transmissão
LT1	Linha de Transmissão de 500 kV, 414 km, 60 Hz
LT2	Linha de Transmissão de 230 kV, 106,646 km, 60 Hz
MAE	<i>Mean Absolute Error</i>
max	Máximo
min	Mínimo
MLP	<i>Multilayer Perceptron</i>
MSE	<i>Mean Squared Error</i>
N/I	Não informado
PCA	<i>Principle Component Analysis</i>
QP	Questão de Pesquisa
RC	Região de Competência
RDP	Registrador Digital de Perturbação
RF	<i>Random Forest</i>
RMS	<i>Root Mean Square</i>
RSM	<i>Random Subspace Method</i>
SDR	Seleção Dinâmica de Regressores
SEP	Sistema Elétrico de Potência
SG-ELM	<i>Summation-Gaussian Extreme Learning Machine</i>
SVD	<i>Singular Value Decomposition</i>
SVM	<i>Support Vector Machine</i>
SW-ELM	<i>Summation-Wavelet Extreme Learning Machine</i>
TKEO	<i>Teager-Kaiser Energy Operator</i>



VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
WPT	<i>Wavelet Packet Transform</i>
XGBoost	<i>eXtreme Gradient Boosting</i>

## LISTA DE SÍMBOLOS

$A$	Fase A
$B$	Fase B
$C$	Fase C
$G$	Terra ( <i>ground</i> )
$\mathcal{F}$	Conjunto ( <i>pool</i> ) de preditores
$\mathcal{F}'$	Subconjunto de preditores de $\mathcal{F}$
$\mathcal{F}_{ens}$	Conjunto de regressores selecionados dentro da estrutura SDR proposta
$\mathcal{F}'_{ens}$	Conjunto de regressores selecionados dentro da estrutura SDR proposta após a aplicação do módulo de “Filtro”
$N$	Número de estimadores
$\mathcal{T}$	Conjunto de treinamento
$\mathcal{W}$	Vetor de pesos
$x_j$	Padrão de teste
$x'_j$	Padrão de teste escalonado
$\mathcal{V}$	Conjunto de validação
$\mathcal{V}'$	Conjunto de validação normalizado junto ao padrão de teste
$\Psi$	Região de competência
$K$	Número de vizinhos para o algoritmo KNN
$T$	Série temporal
$\mathcal{D}$	Desempenhos, dados pela medida de competência empregada, dos regressores na estrutura de seleção dinâmica proposta
$\mathcal{D}'$	Desempenhos, dados pela medida de competência empregada, dos regressores selecionados dentro da estrutura SDR proposta
$\mathcal{P}'$	Predições dos regressores pré-selecionados na estrutura de seleção dinâmica proposta
$\mathcal{F}(x_j)$	Predição final do conjunto de preditores para um determinado padrão de teste
$f(x_j)$	Predição de um dado preditor para um determinado padrão de teste
$f_i$	Preditor unitário presente no <i>pool</i> $\mathcal{F}$

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
1.1	DESAFIOS	22
1.2	OBJETIVOS	22
1.3	HIPÓTESE	22
1.4	ASSERTIVAS	23
1.5	CONTRIBUIÇÕES	23
1.6	ESTRUTURA DO DOCUMENTO	24
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>25</b>
2.1	LINHAS DE TRANSMISSÃO	25
2.1.1	Faltas em Linhas de Transmissão	26
2.1.2	Proteção de Sistemas Elétricos de Potência	31
2.2	APRENDIZADO DE MÁQUINA	33
2.2.1	Árvores de Decisão	34
2.2.2	Modelos Múltiplos Preditivos	36
2.2.3	Árvores Extremamente Aleatórias	40
2.2.4	Seleção Dinâmica de Regressores	41
2.3	CONSIDERAÇÕES FINAIS	43
<b>3</b>	<b>ESTADO DA ARTE</b>	<b>45</b>
3.1	TRABALHOS RELACIONADOS	45
3.2	ANÁLISE CRÍTICA	56
3.3	CONSIDERAÇÕES FINAIS	58
<b>4</b>	<b>BASES DE DADOS PARA ANÁLISE DE FALHAS</b>	<b>59</b>
4.1	GERAÇÃO DOS EVENTOS DE FALTA	59
4.1.1	Linha de Transmissão de 500 kV	60
4.1.2	Linha de Transmissão de 230 kV	62
4.2	EXTRAÇÃO DE ATRIBUTOS	63
4.3	CONSIDERAÇÕES FINAIS	65
<b>5</b>	<b>MÉTODO PROPOSTO</b>	<b>66</b>
5.1	COMPONENTE CLASSIFICADOR DE FALTA	67
5.1.1	Normalização e Extração de Atributos	67
5.1.2	Predição	69
5.2	COMPONENTE LOCALIZADOR DE FALTA	69
5.2.1	Normalização e Extração de Atributos	70
5.2.2	Estrutura SDR Proposta	70

	5.2.3	Predição . . . . .	73
5.3		FERRAMENTAS . . . . .	74
5.4		CONSIDERAÇÕES FINAIS . . . . .	74
<b>6</b>		<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>75</b>
6.1		CLASSIFICAÇÃO DE FALTA . . . . .	75
	6.1.1	Protocolo Experimental . . . . .	75
	6.1.2	Desempenho do Método Proposto para a LT de 500 kV. . . . .	78
	6.1.3	Desempenho do Método Proposto para Generalização . . . . .	80
	6.1.4	Desempenho com Eventos Reais de Falta . . . . .	87
	6.1.5	Análise da Importância dos Atributos . . . . .	93
	6.1.6	Comparação com o Estado da Arte . . . . .	96
6.2		LOCALIZAÇÃO DE FALTA . . . . .	99
	6.2.1	Protocolo Experimental . . . . .	100
	6.2.2	Desempenho do Método Proposto . . . . .	101
	6.2.3	Análise da Importância dos Atributos . . . . .	107
	6.2.4	Comparação com o Estado da Arte . . . . .	108
	6.2.5	Desempenho do Método Proposto para Generalização . . . . .	112
	6.2.6	Complexidade . . . . .	114
6.3		CONSIDERAÇÕES FINAIS . . . . .	115
<b>7</b>		<b>CONCLUSÃO . . . . .</b>	<b>116</b>
7.1		LIMITAÇÕES . . . . .	117
7.2		TRABALHOS FUTUROS . . . . .	117
		<b>REFERÊNCIAS . . . . .</b>	<b>119</b>
		<b>APÊNDICE A – LEVANTAMENTO BIBLIOGRÁFICO . . . . .</b>	<b>131</b>
		<b>APÊNDICE B – ANÁLISE DOS ATRIBUTOS PARA A LOCALIZAÇÃO . . . . .</b>	<b>133</b>
		<b>APÊNDICE C – PUBLICAÇÕES . . . . .</b>	<b>136</b>
		<b>APÊNDICE D – GLOSSÁRIO . . . . .</b>	<b>137</b>

## 1 INTRODUÇÃO

O Sistema Elétrico de Potência (SEP) é uma estrutura complexa, sendo composta por um conjunto de instalações, elementos e equipamentos que permitem a geração, a transmissão e a distribuição de energia elétrica. Desse modo, a energia gerada nas usinas pode ser transportada até grandes, médios e pequenos consumidores, como indústrias, comércios e residências (Blume, 2017). Na Figura 1.1 está a representação simplificada de um SEP.

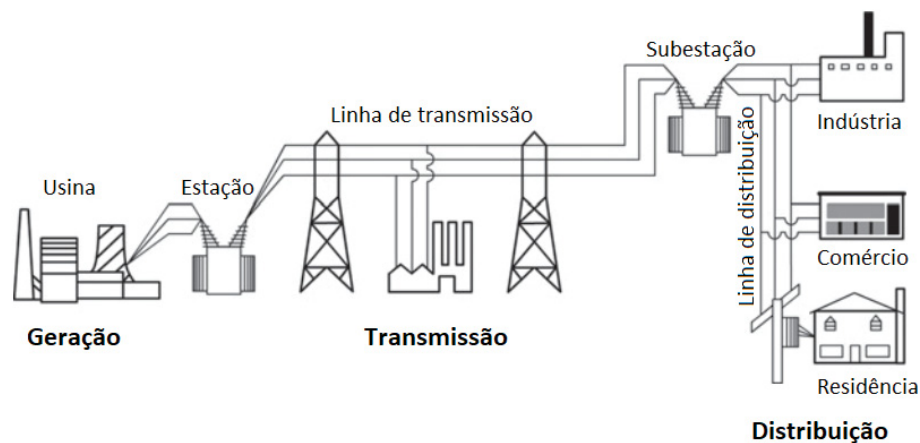


Figura 1.1: Visão geral da estrutura do SEP

Fonte: adaptado de Blume (2017)

Neste cenário, as linhas de transmissão (LTs) são elementos cruciais desta estrutura, conectando as usinas geradoras, localizadas em diferentes pontos do sistema, até às subestações mais próximas aos centros consumidores. No geral, essas linhas percorrem longas distâncias e, devido às suas características físicas, estão expostas às mais diversas situações adversas, como descargas atmosféricas, contato com vegetação e queimadas (ANEEL, 2020). Desse modo, por mais que possuam um sistema de proteção robusto e eficaz, as LTs estão suscetíveis a tais situações que podem acarretar em perturbações que afetam o seu estado normal de operação, as quais são denominadas de faltas (ou falhas). Cabe destacar que cerca de 89% das falhas no SEP ocorrem nas linhas de transmissão (Kindermann, 2018).

Uma falta pode ser definida como um fluxo anormal das grandezas elétricas nos componentes do sistema de potência, como o aumento no fluxo da corrente para uma ou mais fases (Yadav e Dash, 2014; Prasad et al., 2018; Furse et al., 2021). Em particular, sua incidência manifesta um transitório, cerca de alguns milissegundos, que provoca grandes variações na tensão e na corrente, ocasionando em um novo estado (regime) permanente com tais grandezas estáveis, simétricas ou assimétricas, de acordo com o tipo de falta (Bichels, 2018).

Suas consequências são de grande preocupação, visto que uma única falha pode impactar negativamente no fornecimento de energia elétrica para um grande número de consumidores, como a indisponibilidade de energia temporariamente ou permanentemente, ou ainda na ocorrência de *blackouts* (apagões) em efeito dominó. Logo, essa indisponibilidade pode acarretar em multas às concessionárias de energia e insatisfação dos consumidores, mas, principalmente, pode desabastecer locais críticos, como hospitais e unidades de saúde, nos quais o fornecimento de energia elétrica é indispensável.

Diante disto, os sistemas de proteção, localizados nas subestações, são componentes fundamentais que visam resguardar o SEP dos efeitos das faltas, detectando e isolando as partes afetadas, de modo a preservar a integridade e a estabilidade do sistema elétrico remanescente. Por meio do avanço da tecnologia digital, inúmeras melhorias e facilidades foram agregadas ao sistema de proteção, de modo que os dados de tensão e corrente, nos períodos pré e pós-falta, pudessem ser armazenados e empregados para diversas aplicações, como na análise dessas perturbações (Shakiba et al., 2023). Exemplo disso são os Registradores Digitais de Perturbações (RDPs), os quais permitem armazenar registros de oscilografias dessas grandezas elétricas.

Esses registros permitiram que diferentes métodos de análise fossem e sejam elaborados, com destaque para duas atividades fundamentais: classificação e localização de faltas. Em especial, a classificação permite a identificação das fases afetadas pela falta, enquanto a localização permite determinar o ponto aproximado na LT que tal perturbação aconteceu. A análise dessas atividades é importante para que os operadores possam realizar a tomada de decisão referente ao religamento da linha de transmissão com maior confiabilidade e segurança, assim como para o envio de equipes de manutenção até o local mais próximo da ocorrência da falha, se necessário, com um diagnóstico prévio do problema.

Neste contexto, diversos métodos foram propostos na literatura especializada para a realização dessas atividades a partir de três técnicas principais: ondas viajantes (*traveling waves*), impedância e Inteligência Artificial (IA). Porém, as abordagens convencionais (ondas viajantes e impedância) demonstram algumas desvantagens consideráveis (Gururajapathy et al., 2017; Mirzaei et al., 2018; Wang et al., 2023b). A técnica por ondas viajantes é complexa, exigindo uma alta frequência de amostragem e alto custo computacional (Dong et al., 2009; Saha et al., 2010; Mishra e Ray, 2018; Furse et al., 2021; Shakiba et al., 2023). Já a técnica da impedância pode ser afetada pela variação dos parâmetros de falta, principalmente para falhas com resistências elevadas, acarretando na diminuição do desempenho do sistema preditivo (Tang et al., 2000; Mallaki e Dashti, 2012; Patel et al., 2018; Panda et al., 2018; Furse et al., 2021).

Atualmente, os métodos baseados em IA têm ganhado maior atenção devido principalmente ao seu bom desempenho e a sua adaptabilidade para as diferentes condições e parâmetros de falta (Pradhan, 2010; Chen et al., 2016; Raza et al., 2020; Shakiba et al., 2023). No entanto, as abordagens que se fundamentam por esta técnica demandam de uma quantidade significativa de dados para o treinamento de seus algoritmos (Ferreira et al., 2016; Raza et al., 2020).

A indisponibilidade de bases de dados públicas com falhas de linhas em transmissão reflete em um problema crítico na literatura. Geralmente, cada autor gera os seus próprios casos de falta para treinar e avaliar suas propostas, não disponibilizando esses dados para acesso por outros pesquisadores. Esses conjuntos de dados são predominantemente gerados com configurações específicas para cada cenário considerado pelos autores, diferindo entre eles não apenas em termos de níveis de tensão e comprimento da LT, mas também toda a configuração do sistema elétrico considerado no estudo. Uma exceção é a base de dados UFPAFaults disponibilizada por Morais et al. (2010). Contudo, todos os dados correspondem aos sinais de tensão e corrente após a ocorrência da falha e representativos para um único terminal, além de conterem um

processamento prévio nesses dados. Ainda, essa base não permite a realização das atividades de detecção e localização de falta, limitando a análise apenas para a tarefa de classificação.

Também identificamos outras limitações nos trabalhos do estado da arte, sobretudo devido às técnicas de processamento de sinais empregadas, as quais necessitam de quantidades expressivas de amostras, principalmente quando mais níveis de decomposição são utilizados (Gafoor e Rao, 2006; Chen et al., 2018a; Haq et al., 2021; Mukherjee et al., 2021). Entre as técnicas de processamento mais utilizadas estão as *wavelets* (Mallat, 2009), que possibilitam a decomposição do sinal e a extração de coeficientes.

Conforme destacado por Silva et al. (2007) e Ye et al. (2020), as taxas de amostragem podem variar entre os dispositivos registradores produzidos por diferentes fabricantes, o que, por consequência, tende a afetar o desempenho de métodos de análise de falhas, especialmente aqueles que utilizam de técnicas de processamento de sinais (Saha et al., 2010). Por exemplo, na Companhia Paranaense de Energia (COPEL) os equipamentos de aquisição de dados variam a frequência de amostragem com taxas entre 256 Hz e 15,36 kHz. Desse modo, diversos dos métodos encontrados na literatura, por exemplo, não podem ser empregados para a menor taxa devido as limitações de suas propostas, sendo restringidos apenas para os sistemas com a instrumentação apropriada para as suas respectivas abordagens. Inclusive, alguns métodos relatam piora no desempenho para taxas de amostragem mais inferiores (Chen et al., 2018a; Haq et al., 2021), além de não realizarem a generalização de seus métodos para analisar linhas com configurações distintas daquelas utilizadas para treinar os algoritmos.

Essa generalização é essencial pois as subestações podem supervisionar simultaneamente linhas de transmissão com diferentes características entre si, como distintos níveis de tensão, comprimentos das linhas e variações nas potências dos geradores, resultando em amplitudes desiguais de tensão e corrente. Neste cenário, os métodos encontrados na literatura que empregam IA analisam cada linha de transmissão individualmente, treinando e testando seus algoritmos apenas com dados provenientes da mesma linha. Em outras palavras, as abordagens, no geral, não são avaliadas para linhas de transmissão diferentes daquela utilizada para treinar os algoritmos, *i.e.*, os autores não analisam a real capacidade de generalização de seus métodos. Por consequência, é necessário o desenvolvimento de modelos que possuam a generalização entre domínios diferentes (*domain generalization*) (Wang et al., 2023a), visto que cada configuração de LT (comprimento da LT, nível de tensão e detalhes do próprio sistema de potência) pode acarretar em uma distribuição específica das formas de onda de corrente e tensão.

Assim sendo, o desenvolvimento de uma abordagem que seja flexível em relação a taxa de amostragem dos dispositivos registradores é crucial, bem como a habilidade de analisar faltas independentemente da linha de transmissão sob supervisão, mesmo que o algoritmo tenha sido treinado para uma LT distinta.

Outro ponto interessante que merece destaque refere-se aos algoritmos utilizados para as predições, tanto para a classificação quanto para a localização de falhas. Mishra e Ray (2018) e Mukherjee et al. (2021) apresentaram revisões abrangentes sobre o diagnóstico de falhas, relatando que a maioria dos trabalhos relacionados utiliza algoritmos individuais (monolíticos), especialmente Máquina de Vetores de Suporte (*Support Vector Machine – SVM*) e redes neurais.

No entanto, vários estudos demonstram as vantagens e a superioridade do desempenho de modelos múltiplos preditivos (*ensemble*) sobre algoritmos monolíticos para diferentes áreas do conhecimento (Mendes-Moreira et al., 2012; Cruz et al., 2018a). O uso de múltiplos preditores, por exemplo, pode resultar em predições diferentes entre os modelos devido às propriedades específicas de cada preditor e, quando considerados em conjunto, podem resultar em desempenhos mais acurados e precisos (Krogh e Vedelsby, 1994; Ueda e Nakano, 1996).



Neste cenário, um tópico crescente na abordagem *ensemble* é a seleção dinâmica de preditores, sejam eles classificadores ou regressores (Moura, 2019; Sani et al., 2022). A estratégia da seleção dinâmica realiza a seleção de um subconjunto de modelos básicos a partir do conjunto de preditores original (usualmente referenciado como pelo termo *pool*) individualmente para cada novo padrão de teste a ser predito. Um ponto chave dessa estratégia é a diversidade entre os modelos, onde espera-se que diferentes modelos sejam especialistas em uma região específica do espaço de características, mais conhecida como Região de Competência (RC) (Cruz et al., 2018b). Apesar da seleção dinâmica para a classificação ser bem explorada e coberta pela bibliografia, ainda há poucos trabalhos voltados para a tarefa de regressão. Em nosso trabalho, a regressão representa a atividade de localização de falta.

Em resposta aos problemas e oportunidades relatados, nosso trabalho disponibiliza duas novas bases de dados com simulações de falta de duas LTs para a comunidade científica: uma com séries temporais dos sinais de corrente e tensão, e outra com atributos extraídos a partir destes mesmos sinais. Propomos ainda um *framework* de análise que compreende as atividades de classificação e localização de falhas em linhas de transmissão, com foco em duas propriedades: flexibilidade perante diferentes taxas de amostragem, e capacidade de generalização dessas tarefas para linhas com configurações distintas daquela empregada para o treinamento dos algoritmos.

Particularmente, buscamos explorar o uso de modelos múltiplos preditivos para ambas as tarefas previamente citadas. Acreditamos que a engenharia de atributos que desenvolvemos beneficiarão o uso desses algoritmos. Além disso, propomos neste trabalho uma nova arquitetura para a seleção dinâmica de regressores, de modo a atacarmos a lacuna identificada sobre o tema, especificamente sobre a pouca variedade de métodos destinados para a regressão.

Nossa estrutura de seleção dinâmica destaca-se em quatro aspectos principais: (i) variação da métrica de distância utilizada para definir a RC; (ii) variação do critério para seleção dos melhores regressores do conjunto; (iii) inclusão do processo de escalonamento de dados dentro da arquitetura (normalização incorporada); (iv) inclusão de um processo de filtro como etapa de pós-processamento. A relevância dos itens (i) e (ii) se deve ao fato de que referem-se a hiperparâmetros ainda inexplorados pelo estado da arte da seleção dinâmica de regressores. Logo, explorar outros valores para estes hiperparâmetros poderá acarretar na melhoria do desempenho do sistema preditivo.

Por sua vez, o item (iii) diz respeito a um módulo inédito adicionado em nossa estrutura para a normalização incorporada, que dimensiona o padrão de teste e o conjunto de validação juntos, individualmente para cada novo exemplo para definir a RC. A intuição para este processo *online* está relacionado ao fato de que algoritmos baseados em distância, predominantemente usados para definir a RC, são sensíveis à escala dos dados, o que pode afetar os exemplos retornados para comporem esta região. Em outras palavras, características com escalas discrepantes dominarão as demais e poderão levar a uma RC imperfeita (Barboza et al., 2023).

Já o item (iv), também adicionado como um módulo inédito em nossa estrutura, procura tornar as predições ainda mais precisas ao remover as predições que sejam discrepantes (*outliers*) entre os modelos selecionados, servindo como uma espécie de filtro. Cabe destacar que este trabalho é o primeiro a propor uma etapa de pós-processamento no estado da arte da Seleção Dinâmica de Regressores (SDR). Demonstraremos empiricamente na Subsubseção 6.2.2.4 o impacto que a inclusão deste módulo causa no desempenho do nosso método.

## 1.1 DESAFIOS

O principal desafio deste trabalho é desenvolver um método eficiente capaz de classificar e localizar faltas em linhas de transmissão independentemente da taxa de amostragem e que possa generalizar tais atividades para linhas com diferentes configurações. Outros desafios importantes a serem citados são:

- Indisponibilidade de bases de dados representativas e públicas de eventos de falta na literatura especializada;
- Indisponibilidade de uma quantidade significativa de dados reais de falta para treinamento e teste do método;
- Erro associado ao sistema utilizado para classificar e localizar as faltas reais disponibilizadas pela concessionária de energia elétrica, visto que a localização determinada, por exemplo, pode não corresponder a real posição da falha na linha de transmissão.

## 1.2 OBJETIVOS

O objetivo principal deste trabalho é propor e avaliar um método para classificar e localizar faltas em linhas aéreas de transmissão a partir de algoritmos de aprendizado de máquina. Os objetivos secundários são:

- Desenvolver e disponibilizar uma base de dados com casos diversificados de eventos de falta de linhas de transmissão, de modo que a base possa permitir a realização de tarefas como a detecção, a classificação e a localização de falhas, incluindo a capacidade de avaliação de métodos para a generalização de domínio (*domain generalization*);
- Propor um método eficiente, acurado e preciso para as tarefas de classificação e localização de falta independentemente da taxa de amostragem;
- Permitir a generalização da classificação e/ou da localização de faltas para linhas de transmissão com configurações distintas daquela empregada para treinar o algoritmo (*domain generalization*) (Wang et al., 2023a);
- Propor uma nova arquitetura para a seleção dinâmica de regressores para desempenhar a tarefa de localização;
- Avaliar e validar o método proposto a partir de experimentos abrangentes.

## 1.3 HIPÓTESE

A hipótese principal deste trabalho estabelece que os modelos múltiplos preditivos superem o desempenho de algoritmos monolíticos, tanto para classificar quanto para localizar faltas em linhas aéreas de transmissão. Outra hipótese refere-se que a engenharia de atributos desenvolvida habilite o método para que seja flexível em relação a taxa de amostragem utilizada para a coleta de dados para ambas as tarefas. Em especial para a classificação, acreditamos que o método irá ainda generalizar seu funcionamento para linhas de transmissão com configurações distintas daquela empregada para treinar o algoritmo, uma vez que a natureza de um tipo de falta prevalece independentemente das características da linha.

## 1.4 ASSERTIVAS

As assertivas deste trabalho são:

- As tarefas de classificação e localização de faltas são capazes de serem realizadas a partir de oscilografias de grandezas elétricas, como tensão e/ou corrente;
- Os dados de tensão e corrente são obtidos pelos sistemas de proteção localizados nos terminais das linhas de transmissão;
- Os registradores digitais de perturbações podem apresentar diferentes frequências de amostragem;
- A classificação do tipo de falta e a sua localização na LT permitem aos operadores do sistema maiores informações das falhas, de modo a encaminhar eficientemente a equipe de manutenção para a realização do reparo, resguardar a integridade de equipamentos e, principalmente, preservar a segurança de empregados e reduzir a indisponibilidade do fornecimento de energia elétrica;
- Os algoritmos de modelos múltiplos preditivos (*ensemble*) manifestam superioridade de desempenho perante algoritmos monolíticos para diversas aplicações em diferentes áreas do conhecimento;
- A inclusão da estratégia de seleção dinâmica no contexto dos algoritmos *ensemble* tende a repercutir no ganho de desempenho, fato já evidenciado para outras áreas do conhecimento.

## 1.5 CONTRIBUIÇÕES

Os esforços aplicados neste trabalho resultaram nas seguintes contribuições para a comunidade científica:

- Construção e disponibilização de duas bases de dados públicas com eventos diversificados de falhas em linhas de transmissão, com variação de parâmetros essenciais, como o tipo, a resistência, o tempo de início e a localização. A primeira base corresponde a séries temporais dos sinais de tensão e corrente, enquanto a segunda base contém atributos extraídos destes sinais. Estas bases beneficiarão, em especial, as áreas da engenharia elétrica e da computação, de modo que pesquisadores poderão utilizar estes dados para finalidades que incluem: análise de falhas de um modo amplo, detecção de anomalias (*i.e.*, falhas), *domain generalization*, classificação e regressão;
- Desenvolvimento de um método eficiente, acurado e preciso para as tarefas de classificação e localização de falta:
  - Independente da taxa de amostragem;
  - Com capacidade para atuar em diferentes linhas de transmissão com configuração distinta daquela utilizada para treinar o algoritmo.
- Uma nova arquitetura para a seleção dinâmica de regressores, a qual inclui módulos inéditos e importantes ao longo de sua estrutura. Apesar de avaliarmos esta estrutura apenas para a localização de faltas, ela também pode ser usada para qualquer outra aplicação de regressão:

- Um módulo denominado “Normalização Incorporada”, o qual objetiva realizar o processo de normalização dos dados dentro da estrutura SDR, o que permite definir uma RC mais precisa, repercutindo em melhores resultados;
- Um módulo denominado “Filtro”, o qual atua como uma etapa de pós-processamento ao remover as predições que sejam discrepantes (*outliers*) entre os modelos selecionados.

Além disso, as contribuições descritas anteriormente podem ser agrupadas nas seguintes categorias:

- **Contribuição social:** o diagnóstico eficiente de uma falha permite que os operadores do SEP realizem a tomada de decisão com maior confiança e segurança, o que tende a reduzir o tempo de indisponibilidade no fornecimento de energia elétrica para residências, indústrias e estabelecimentos comerciais;
- **Contribuição econômica/empresarial:** além dos pontos citados no item anterior, em especial no que se refere as indústrias e empresas, o rápido reestabelecimento da operação da linha acarretará na diminuição dos valores de eventuais multas cujas concessionárias de energia elétrica poderão estar submetidas (Resolução Normativa nº 729/2016 da ANEEL).

As publicações realizadas durante o projeto, as quais amparam as contribuições mencionadas previamente, estão citadas no Apêndice C.

## 1.6 ESTRUTURA DO DOCUMENTO

A organização do restante do documento está estruturada da maneira que segue. No Capítulo 2 está apresentada a fundamentação teórica com os conceitos básicos relacionados a temática deste trabalho, em especial sobre faltas em linhas de transmissão e modelos múltiplos preditivos, com destaque para a seleção dinâmica. No Capítulo 3 descrevemos sobre os trabalhos relacionados encontrados no estado da arte. Já no Capítulo 4, relatamos sobre as bases de dados criadas e disponibilizadas em nosso projeto, enquanto no Capítulo 5 descrevemos o método proposto para a classificação e para a localização de falhas. No Capítulo 6 estão expostos e discutidos os resultados alcançados ao longo do estudo a partir de experimentos abrangentes para a validação das nossas propostas. Por fim, no Capítulo 7 descrevemos a conclusão deste trabalho, pontuando as limitações e as perspectivas para trabalhos futuros. É digno de nota que no Apêndice D resumizamos em um breve glossário as definições dos principais termos que compõem esta tese.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo apresentamos uma introdução aos conceitos elementares que envolvem este projeto. Na Seção 2.1 descrevemos uma visão geral sobre as linhas de transmissão, com ênfase para as faltas que podem ocorrer nestas linhas (Subseção 2.1.1), além de uma contextualização sobre o sistema de proteção que objetiva garantir, entre outras funções, a segurança e a operação eficiente do SEP frente a ocorrência de falhas (Subseção 2.1.2). É digno de nota que o método proposto neste trabalho se enquadra como uma ferramenta no contexto do sistema de proteção para a análise de faltas, em particular para as tarefas de classificação e localização. Já na Seção 2.2 relatamos os princípios referentes ao aprendizado de máquina, especialmente aqueles que integram nosso método, o qual será apresentado mais adiante neste documento. Assim sendo, os tópicos que destacamos são as árvores de decisão (Subseção 2.2.1), os modelos múltiplos preditivos (Subseção 2.2.2), as Árvores Extremamente Aleatórias (Subseção 2.2.3) e, por fim, a seleção dinâmica de regressores (Subseção 2.2.4).

### 2.1 LINHAS DE TRANSMISSÃO

As linhas de transmissão (Figura 2.1) são elementos importantes do sistema elétrico de potência, permitindo que a energia gerada em usinas hidrelétricas, term nucleares, termelétricas, solares e eólicas, por exemplo, seja transportada até subestações para então ser distribuída aos consumidores (Fuchs, 1977; Pinto, 2014; Blume, 2017). Essas linhas, em sua maioria, percorrem longas distâncias e vastas regiões, e, conseqüentemente, estão vulneráveis a todos os tipos de perturbações naturais, ambientais e operacionais (Sato e Freitas, 2015).

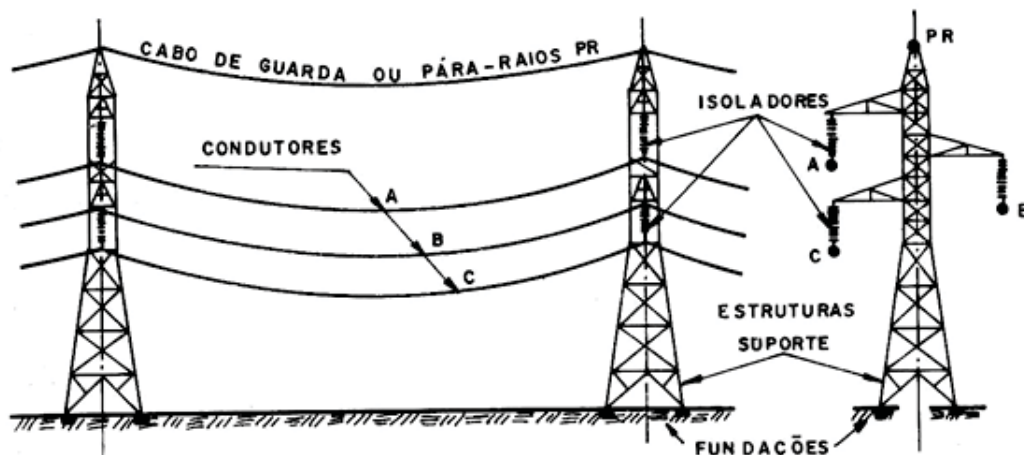


Figura 2.1: Principais componentes de uma linha de transmissão

Fonte: Labegalini et al. (1992)

Por esse motivo, a construção dessas linhas leva em consideração dados históricos, meteorológicos e geotécnicos da região para escolher a melhor alternativa a ser implantada (Moura et al., 2019a). Outro aspecto relevante está relacionado às escolhas das diferentes características dos componentes, como os materiais, os tipos e as configurações desses componentes.

Em geral, as linhas de energia são classificadas em (Sato e Freitas, 2015):

- **Distribuição** (2,2 kV a 34,5 kV);
- **Subtransmissão** (13,8 kV a 138 kV);
- **Transmissão** (69 kV a 765 kV):
  - Alta Tensão (69 kV a 220 kV);
  - Extra-Alta Tensão (345 kV a 765 kV);
  - Ultra-Alta Tensão (acima de 765 kV).

É digno de nota que os níveis de tensão do sistema de transmissão são definidos com base na distância percorrida e na quantidade de energia a ser transportada, de modo a evitar perdas elétricas ao longo de sua transmissão (de Vasconcelos, 2017). Além disso, as LTs funcionam com cargas predominantemente trifásicas (fases A, B e C), simétricas e equilibradas, ou seja, as amplitudes das tensões e das correntes são iguais entre cada uma das três fases, com defasagem de  $120^\circ$  entre elas (Hedman, 1979; Bichels, 2018).

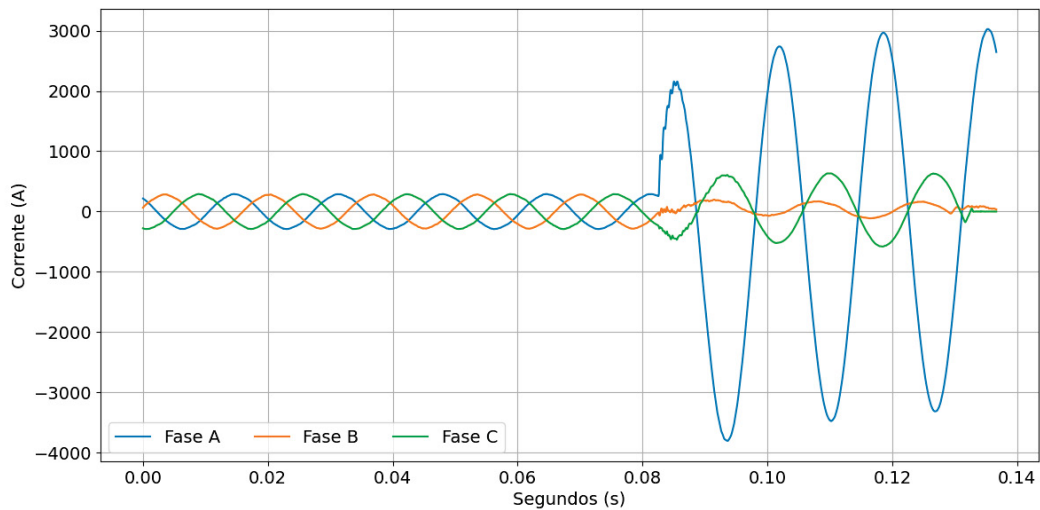
### 2.1.1 Faltas em Linhas de Transmissão

O projeto de uma linha de transmissão visa determinar as melhores configurações e os materiais mais adequados a serem utilizados nos componentes para a construção da linha. Apesar de todos os cuidados, as linhas aéreas estão expostas aos problemas mais diversos e imprevisíveis que podem afetar o seu estado normal de operação, como as condições ambientais adversas (Kindermann, 2018; Liang et al., 2022).

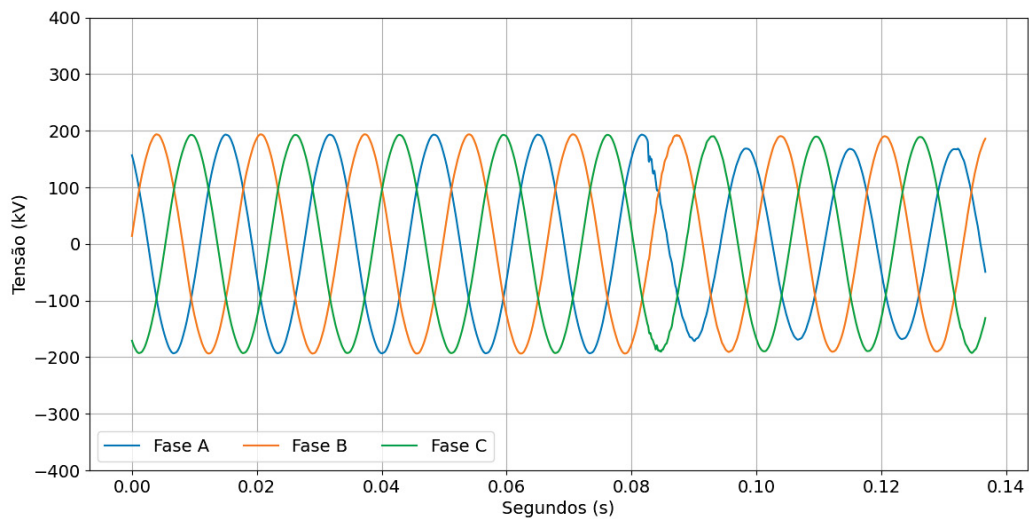
Logo, a ocorrência de intempéries nas LTs pode ocasionar em faltas, as quais impactam diretamente no fornecimento da energia elétrica devido a alterações repentinas e bruscas na operação normal do SEP (Yadav e Dash, 2014; Sato e Freitas, 2015), acarretando, eventualmente, na interrupção da transmissão de energia pelas linhas. Uma falta pode ser definida como um fluxo anormal das grandezas elétricas, como tensão e corrente, nos componentes de um sistema de potência (Prasad et al., 2018).

Na Figura 2.2 está ilustrado um exemplo de uma falta real coletada da rede da COPEL com a exibição dos sinais da corrente (Figura 2.2(a)) e da tensão (Figura 2.2(b)). Nota-se que, inicialmente, ambos os sinais manifestam o comportamento normal de operação da LT, com as amplitudes simétricas e equilibradas (pré-falta). Entretanto, logo após o instante 0,08 segundo ocorre a incidência de uma falta na linha, afetando o comportamento de ambos os sinais (pós-falta), em especial para o sinal da corrente. Em particular, as letras A, B e C caracterizam cada uma das três fases de uma linha de transmissão. Duas definições importantes que podem ser estabelecidas a partir deste exemplo são os sinais pré e pós-falta. Neste trabalho, denominamos sinal pré-falta o período que antecede a ocorrência da falha, enquanto o sinal pós-falta constitui os dados após a incidência do início da falha.





(a) Sinal da corrente



(b) Sinal da tensão

Figura 2.2: Exemplo de uma falta real do tipo AG com início logo após o tempo 0,08 segundo. Na imagem (a) está representado o sinal da corrente, enquanto na imagem (b) está retratada a forma de onda da tensão

Fonte: autoria própria

Neste cenário, as faltas podem ser classificadas em (Stevenson, 1986; Goh et al., 2017; Gururajapathy et al., 2017):

- **Falta em série (*series fault*)**: também conhecida como falha por circuito aberto (*open circuit fault*), ocorre devido à falha de um ou mais cabos condutores em um sistema elétrico. Essa falha pode ser ocasionada por um defeito na junção dos cabos ou quebra/dano nos condutores, por exemplo. Caracteriza-se pelo aumento da tensão e pela redução da corrente nas fases com falha. Frequentemente são transformadas em curto-circuitos após uma sequência de eventos;



- **Falta de derivação (*shunt fault*):** mais conhecida como falha por curto-circuito (*short circuit fault*), ocorre devido à conexão indesejada entre os próprios condutores de fase ou entre os condutores de fase e a terra em um sistema elétrico. Essa falha pode ser ocasionada por diferentes motivos, como descargas atmosféricas e contato dos condutores com vegetação. Caracteriza-se pelo aumento da corrente (Figura 2.2(a)) e pela redução da tensão (Figura 2.2(b)) especialmente nas fases afetadas pela falha (Clark, 1979).

Os curtos-circuitos são as perturbações mais comuns e também as mais severas, visto que provocam a circulação de correntes elevadas, as quais podem causar danos aos cabos e aos demais componentes (Stevenson, 1986; Sato e Freitas, 2015). Esse tipo de falta pode ocorrer em qualquer local da linha de transmissão, entre quaisquer pontos energizados ou entre um ponto energizado e a terra. Uma falta por curto-circuito pode ser classificada ainda em (Bichels, 2018):

- **Assimétrica:** também conhecida como falta desbalanceada/desequilibrada devido ao fluxo desequilibrado das correntes entre as fases. É representada pelas faltas monofásicas (fase-terra<sup>1</sup>) e bifásicas (fase-fase<sup>2</sup> e fase-fase-terra<sup>3</sup>), com ou sem o contato com a terra;
- **Simétrica:** também conhecida como falta balanceada/equilibrada devido ao fluxo equilibrado das correntes entre as fases. É representada pelas faltas trifásicas (fase-fase-fase<sup>4</sup> e fase-fase-fase-terra<sup>5</sup>), com ou sem contato com a terra.

Na Figura 2.3 está representada a classificação das faltas anteriormente descritas. Vale lembrar que as letras A, B e C caracterizam cada uma das três fases de uma linha de transmissão, enquanto a letra G corresponde a atuação da terra/solo (*ground*) na respectiva falta. Por exemplo, a sigla AG retrata a falta monofásica da fase A com a terra, enquanto AB representa a falta bifásica entre as fases A e B sem o contato com a terra.

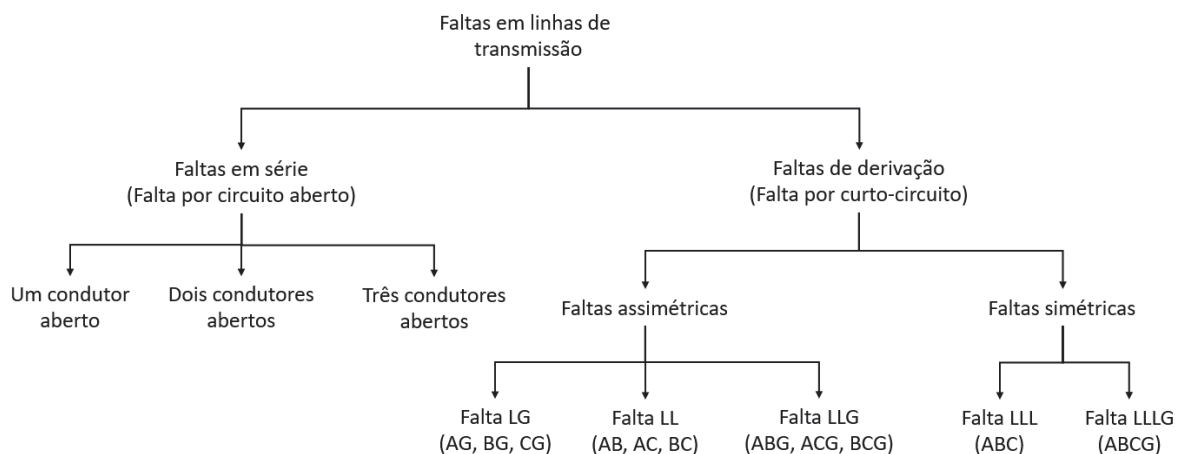


Figura 2.3: Classificação das faltas em linhas de transmissão

Fonte: adaptado de Prasad et al. (2018)

<sup>1</sup>Line to ground (LG).

<sup>2</sup>Line to line (LL).

<sup>3</sup>Double line to ground (LLG).

<sup>4</sup>Triple line (LLL).

<sup>5</sup>Triple line to ground (LLLG).

O curto-circuito fase-terra ocorre quando uma das fases da LT entra em contato com o solo. De modo semelhante ocorre para as faltas fase-fase-terra, com o agravante de envolver duas fases da linha. Por sua vez, o curto-circuito fase-fase ocorre quando dois condutores da LT, de fases diferentes, entram em contato um com o outro, sem o envolvimento do solo. Por fim, falhas trifásicas acontecem esporadicamente e ocorrem quando as três fases entram em contato entre si, sem necessariamente envolver o solo (Faulkenberry e Coffey, 1996).

Os curtos-circuitos que ocorrem com maior frequência são os assimétricos, principalmente os monofásicos. Já as faltas trifásicas simétricas são mais raras, porém as que causam os maiores danos ao SEP (Sato e Freitas, 2015). Na Tabela 2.1 estão apresentadas as frequências das ocorrências para cada tipo de falta.

Tabela 2.1: Levantamento das ocorrências de faltas

<b>Tipos de falta</b>	<b>Ocorrências</b>
Monofásica	64%
Bifásica	15%
Bifásica-terra	16%
Trifásica	5%

Fonte: Kindermann (2018)

Resumidamente, um curto-circuito pode ser de dois tipos (Stevenson, 1986; Kindermann, 2018):

- **Permanente:** representa 4% das ocorrências e, como o próprio nome indica, são do tipo irreversível. Após a abertura do disjuntor, a equipe de manutenção deve realizar o conserto na rede para que o sistema possa ser restabelecido. Essas falhas são provenientes de causas como linhas caídas ao solo, ruptura de uma cadeia de isoladores e falhas de para-raios;
- **Temporário:** representa 96% das ocorrências e são aquelas que ocorrem sem a existência de defeito na rede. Após a atuação da proteção, o sistema pode ser restabelecido normalmente, como por meio do religamento automático ou pelos próprios operadores. Essas falhas são provenientes de diversas causas, como o contato com galhos de árvores e descargas atmosféricas nos condutores.

Uma propriedade fundamental no estudo das redes trifásicas são os componentes simétricos do sistema de potência, utilizados para calcular as condições de desequilíbrio de um sistema trifásico (Fortescue, 1918; Fleming, 1998). A análise desses componentes ocorre por meio de três grandezas:

- **Sequência-positiva:** presente durante as condições trifásicas equilibradas. As grandezas de sequência-positiva representam as correntes e as tensões normais observadas nos sistemas de potência durante condições típicas de regime;
- **Sequência-negativa:** mede a quantidade de desbalanço existente no sistema de potência;
- **Sequência-zero:** mais comumente associada ao fato do envolvimento da terra em condições de desbalanço.

As grandezas de sequência-negativa e zero somente estão presentes com valores substancialmente elevados durante condições de faltas desequilibradas (assimétricas) no sistema de potência (Hedman, 1979; Fleming, 1998). Portanto, elas são frequentemente utilizadas para determinar se existe uma condição de falta no sistema. A sequência-negativa pode ser usada para detectar faltas fase-fase, fase-terra e fase-fase-terra. Por sua vez, a sequência-zero pode ser usada para detectar faltas fase-terra e fase-fase-terra.

As ameaças capazes de acarretar em falhas nas LTs, provenientes de condições isoladas ou combinadas, podem ser originadas por (Kindermann, 2018; ANEEL, 2020):

1. **Problemas de isolamento** – causam tensões elevadas nos cabos condutores que levam a fugas para a terra ou entre os cabos:
  - Desenho inadequado dos isoladores, dos equipamentos ou das estruturas;
  - Material inadequado ou de má qualidade empregado na fabricação;
  - Envelhecimento ou dano do próprio material.
2. **Problemas mecânicos** – provenientes da natureza e que provocam ação mecânica:
  - Ação do vento;
  - Neve;
  - Contaminação;
  - Árvores.
3. **Problemas elétricos** – provenientes da natureza ou devido à operação do sistema:
  - Descargas atmosféricas diretas ou indiretas;
  - Surtos de chaveamento (manobra).
4. **Problemas de natureza térmica** – provenientes do aquecimento nos cabos e nos equipamentos, prejudicando tanto a isolamento quanto a diminuição de suas vidas úteis:
  - Sobrecorrentes devido a sobrecarga no sistema;
  - Sobretensão dinâmica.
5. **Problemas de manutenção** – provenientes de erros durante a manutenção das linhas:
  - Peças de reposição inadequadas;
  - Erro humano durante a manutenção.
6. **Problemas de outra natureza** – provenientes de causas externas:
  - Vandalismo;
  - Queimadas;
  - Animais;
  - Acidentes de qualquer natureza.

### 2.1.2 Proteção de Sistemas Elétricos de Potência

As subestações, tipicamente alimentadas por várias linhas de transmissão, são partes fundamentais de um SEP, sendo uma de suas principais funções incorporar as unidades de medição, proteção, interrupção e despacho de energia elétrica (Pérez-Arriaga et al., 2011).

Em particular, o sistema de proteção, localizado nas subestações, é responsável por detectar e isolar, o mais rápido possível, a parte atingida por uma falta do restante da rede elétrica, preservando a integridade e a estabilidade do sistema elétrico remanescente. Essa anomalia provoca a circulação de correntes elevadas que podem causar danos aos componentes da LT, além de propagar seus efeitos ao longo do sistema caso nenhuma ação seja realizada, afetando diretamente a qualidade do fornecimento de energia até mesmo em zonas distantes do ponto onde aconteceu a falta (Sato e Freitas, 2015).

Os principais equipamentos que compõem um sistema de proteção são (Clark, 1979; Cidrás et al., 2011):

- **Bateria:** equipamento que garante a continuidade do fornecimento de energia para o dispositivo de proteção operar, visto que este não pode ser energizado diretamente pela linha;
- **Transformador de medida (instrumento):** equipamento responsável por reproduzir, em pequena escala, as altas magnitudes de tensão e corrente aos relés. Também denominado de transdutor;
- **Relé de proteção:** equipamento responsável pela tomada de decisão do sistema de proteção. O relé pode realizar tarefas como a ativação de alarmes, o acionamento de disjuntores e o envio de informações aos centros de operação;
- **Disjuntor:** equipamento que permite que um circuito energizado seja aberto ou fechado, ou seja, possui capacidade para interromper correntes de curtos-circuitos, isolando a parte sob falta do restante do sistema.

No geral, os relés utilizam como entrada as magnitudes de tensão e corrente, processadas pelos transformadores de instrumento, e extraem uma série de medidas, como valor eficaz e valores máximos. Em seguida, essas medidas alimentam um sistema de decisões lógicas dos relés (lógica de acionamento) que compara os valores obtidos com valores previamente ajustados. A intervenção do relé ocorrerá sempre que um valor de uma medida exceder o respectivo valor armazenado, atuando sobre um disjuntor que permite abrir um circuito para isolar a parte afetada pela falta e evitar que seus efeitos se propaguem (Cidrás et al., 2011; Sato e Freitas, 2015). A tomada de decisão deve ser realizada em um curto período de tempo, tipicamente entre 8 e 40 ms, dependendo do projeto desses equipamentos (Stevenson, 1986). Quanto à localidade do relé, faltas próximas provocam correntes grandes e de baixa tensão, enquanto defeitos mais distantes provocam correntes menores e tensões não tão baixas (Clark, 1979).

Nesse aspecto, o relé deve ser tão sensível quanto possível dentro de uma faixa de ajuste de operação, além de responder com extrema rapidez às grandezas elétricas as quais está regulado, com tempo de atuação na ordem de poucos milissegundos. Ainda, esse equipamento deve apresentar elevado grau de confiabilidade, ou seja, deve atuar corretamente para todas as condições de perturbações do sistema para o qual foi dimensionado e ajustado (Mamede Filho e Mamede, 2020).

Desse modo, os requisitos básicos que compõem um sistema de proteção são (Cidrás et al., 2011; Sato e Freitas, 2015; Mamede Filho e Mamede, 2020):

1. **Seletividade:** capacidade de garantir que somente o elemento de proteção (relé) mais próximo do defeito desconecte a parte defeituosa do sistema elétrico. Sua importância se deve ao fato de que, ao ocorrer uma falta, diversos locais do sistema detectarão sua ocorrência, sendo importante que apenas o elemento de proteção mais próximo atue para que seus efeitos e suas consequências sejam minimizadas. Em outras palavras, o relé deverá atuar apenas quando uma falta ocorrer dentro de sua zona de proteção;
2. **Velocidade:** capacidade do elemento de proteção atuar para eliminar uma falta no menor tempo possível. Quanto menor o tempo para isolar uma falha, menores são os impactos e os danos causados aos equipamentos;
3. **Sensibilidade:** capacidade do elemento de proteção reconhecer com precisão a faixa de valores indicados para sua atuação e não atuação, ou seja, o relé, por exemplo, deve diferenciar precisamente condições de falta e condições normais de operação;
4. **Confiabilidade:** capacidade do elemento de proteção atuar com segurança e exatidão em todas as funções que lhe foram destinadas, não ocorrendo acionamentos desnecessários ou omissão de acionamentos necessários.

A eficiência da proteção está diretamente relacionada aos ajustes realizados nos relés. Neste cenário, deve-se determinar as correntes de faltas para as mais variadas condições possíveis e, por meio desses valores, ajustar cada relé para que eles possam atuar da maneira mais correta. Esses valores de ajuste são oriundos de cálculos das correntes de curto-circuitos (Kindermann, 2018; Bichels, 2018). Cabe destacar que esses ajustes não são o foco do presente trabalho, partindo-se do princípio de que o sistema está previamente configurado com o conjunto de valores mais adequados.

Nas usinas e nas subestações do SEP é comum a presença de outros equipamentos, além dos relés, para o monitoramento das grandezas elétricas (por exemplo, correntes e tensões). Neste quesito, os Registradores Digitais de Perturbação (RDPs) desempenham papel indispensável na análise de ocorrências em um SEP. A principal função destes equipamentos é registrar a operação do sistema elétrico e sua proteção durante eventos como faltas elétricas, oscilações de frequência e falhas operativas (Moreto e Rolim, 2010). Tais equipamentos gravam continuamente os valores amostrados das grandezas elétricas em uma memória cíclica e, ao identificar uma perturbação, armazenam o conteúdo da memória em um arquivo de oscilografia<sup>6</sup>.

Assim, uma vez constatada uma falta, um RDP pode armazenar o registro do respectivo evento com dados pré e pós-falta a partir de especificações de tempo previamente configuradas (Kezunovic e Rikalo, 1996; Moreto e Rolim, 2010). Desta maneira, caso haja atuação indevida do sistema proteção, o registro do RDP poderá ser utilizado para investigar esta atuação e promover eventualmente medidas corretivas, como o reajuste dos relés de proteção (Bichels, 2018).

O armazenamento desses registros permite ainda que outras tarefas possam ser realizadas, como a classificação e a localização de faltas (Luo e Kezunovic, 2005; Saha et al., 2010). A classificação objetiva identificar qual o tipo da falha, como aquelas listadas na Figura 2.3, ocorreu. Por sua vez, o propósito da localização é determinar o ponto da ocorrência da falha na LT, reportando a distância identificada da falta em relação a um dos terminais<sup>7</sup> da linha. Suas

<sup>6</sup>Oscilografia representa o registro das oscilações de grandezas elétricas do sistema, como tensões e correntes (Moreto e Rolim, 2010).

<sup>7</sup>Em sistemas de transmissão de energia elétrica, um terminal, também referenciado como extremidade, é o ponto onde a LT se conecta a uma subestação ou a outro equipamento elétrico.

relevâncias se devem ao fato de servirem como informação para que os operadores possam tomar decisões para manutenção e religamento das linhas, caso o religamento automático não seja possível, de modo a reduzir o tempo de indisponibilidade do sistema.

Para fins de contextualização, a etapa de detecção de falta, que em tese antecede as tarefas de classificação e localização, é desempenhada pelo próprio sistema de proteção no contexto de aplicações *online*. Já em aplicações *offline*, dedicadas à análise dos registros de tais perturbações, a fase de detecção é importante para determinar o instante de início da falha na oscilografia, servindo como filtro de informações relevantes para as análises subsequentes (*i.e.*, classificação e localização) (Ferreira et al., 2016).

Considerando as faltas permanentes, os operadores podem utilizar a classificação do tipo de falha e a sua localização na linha de transmissão para o deslocamento das equipes para manutenção e, assim, o restabelecimento da linha. Já em relação as falhas temporárias, essas informações podem ser empregadas no sentido de prevenção, de modo a identificar possíveis pontos fracos do sistema e, desse modo, evitar que futuras faltas permanentes possam ocorrer (da Silveira, 2007). Cabe enfatizar que o método proposto neste projeto está inserido dentro do ambiente do sistema de proteção.

## 2.2 APRENDIZADO DE MÁQUINA

A Inteligência Artificial (IA) é uma área que busca permitir que dispositivos computacionais constituam a habilidade de “pensar”, ou seja, investiga a elaboração de algoritmos com a capacidade de identificar restrições expostas pelas representações (dados) com o objetivo de prover suporte e/ou realizar a tomada de decisão. Assim sendo, a IA explora maneiras de habilitar um dispositivo computacional a realizar tarefas e resolver problemas nos quais, até o momento, o ser humano tem um desempenho melhor (Rich et al., 2009; Ertel, 2017; Russell e Norvig, 2020).

Neste cenário, destaca-se a subárea do Aprendizado de Máquina (AM), ramo importante da inteligência artificial que abrange o desenvolvimento de algoritmos que possam aprender por meio dos dados (Alpaydin, 2020). A palavra “aprendizado”, no contexto de AM, representa o processo de busca automática pelas melhores representações do conhecimento embutido nos dados (Chollet e Allaire, 2018). A definição clássica de Tom Mitchell estabelece que (Mitchell, 1997): diz-se que um programa de computador aprende com a experiência  $E$  no que diz respeito a alguma classe de tarefas  $T$  e medida de desempenho  $P$ , se o seu desempenho nas tarefas em  $T$ , conforme medido por  $P$ , melhora com a experiência  $E$ .

O AM, conforme representado pela Figura 2.4, pode ser dividido nas seguintes abordagens principais (Duda et al., 2000; Igual e Seguí, 2017; Russell e Norvig, 2020):

- **Aprendizado supervisionado:** algoritmos que aprendem a partir de um conjunto de treinamento com exemplos rotulados para realizar a tarefa desejada, dividindo-se em dois tipos:
  - **Classificação:** atributo-alvo (classe) é categórico;
  - **Regressão:** atributo-alvo (atributo de predição) é um valor numérico.
- **Aprendizado não supervisionado:** algoritmos que aprendem a partir de um conjunto de treinamento com exemplos não rotulados. Utilizado para explorar dados e determinar agrupamentos de indivíduos de acordo com algum critério de similaridade;
- **Aprendizado semi-supervisionado:** algoritmos que utilizam exemplos rotulados para obter informações sobre o problema e utilizá-las para guiar o processo de aprendizado a partir dos exemplos não rotulados;



- **Aprendizado por reforço:** algoritmos cujo treinamento é realizado por meio de recompensas e penalidades, devendo aprender por si só qual a melhor estratégia (política) para acumular o máximo de recompensas.

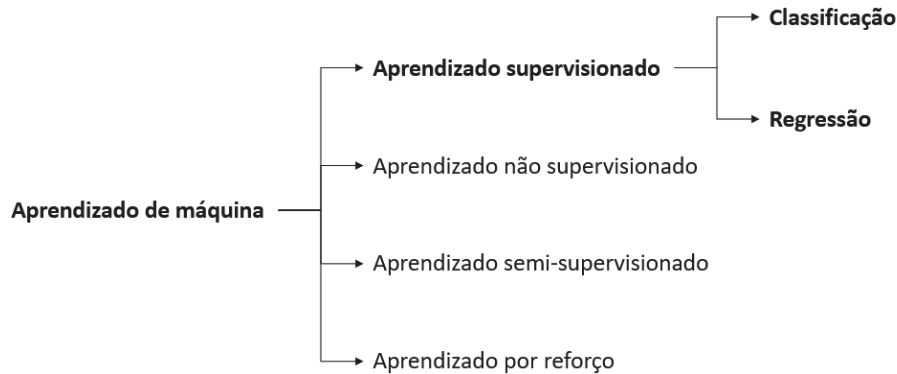


Figura 2.4: Diagrama simplificado das ramificações do AM

Fonte: autoria própria

O presente trabalho está inserido em ambas as tarefas do aprendizado supervisionado, conforme destacadas pela Figura 2.4. Em nosso estudo, a classificação representa a identificação do tipo de falta, enquanto a regressão corresponde a localização da falha na LT. Vale destacar que muitos dos algoritmos utilizados para a classificação, considerando as devidas adaptações, também podem ser empregados para problemas de regressão, e vice-versa. Exemplos desses algoritmos são *k*-Vizinhos mais Próximos (*k-Nearest Neighbors* – KNN), SVM, árvores de decisão e redes neurais (Alpaydin, 2020).

### 2.2.1 Árvores de Decisão

Uma árvore de decisão é um modelo que organiza o conhecimento extraído em uma estrutura hierárquica com refinamentos sucessivos até a definição final do arranjo, aprendendo regras de decisão inferidas a partir do conjunto de atributos (Loh, 2011). A estratégia adotada pelo algoritmo, denominada de dividir para conquistar, realiza a divisão de um problema complexo em problemas mais simples recursivamente, de modo que a estrutura final se assemelha a de uma árvore (Witten e Frank, 2017; Kubat, 2017; Alpaydin, 2020). Além de sua relativa simplicidade, este modelo destaca-se também pela sua capacidade de permitir a nós, seres humanos, a interpretação do conhecimento adquirido para a tomada de decisão tanto para a classificação quanto para a regressão, atendendo aos princípios de que os modelos de AM devem também ser “explicáveis” (Faceli et al., 2021). Na Figura 2.5 está exemplificado o funcionamento de uma árvore de decisão.



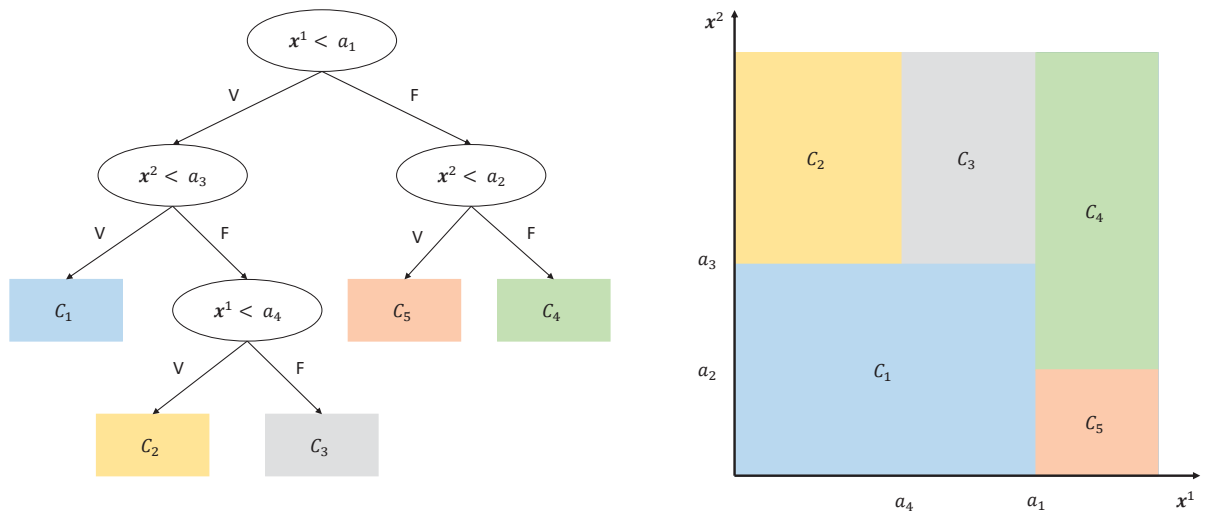


Figura 2.5: Exemplo de uma árvore de decisão junto as regiões de decisão no espaço de atributos ( $x^1$  e  $x^2$ )

Fonte: Faceli et al. (2021)

Uma árvore de decisão pode ser formalmente definida como um grafo direcionado acíclico com as seguintes propriedades (vide Figura 2.5):

- **Nó de divisão:** envolve testar um atributo específico. Normalmente, o teste em um nó compara um valor de atributo com uma constante. Entretanto, algumas árvores comparam dois atributos entre si ou usam alguma função com um ou mais atributos;
- **Nó folha:** fornece uma predição que se aplica a todas as instâncias que alcançam a folha, ou um conjunto de classificações, ou uma distribuição de probabilidade sobre todas as classificações possíveis.

Uma regra de divisão é orientada por uma medida que indica quão bem determinado atributo discrimina as classes, desejando-se selecionar o atributo que maximiza essa medida. Para a classificação, as regras de divisão baseadas no Ganho de Informação e no índice de Gini são as mais frequentemente utilizadas. Já para problemas de regressão, a função de custo a ser minimizada é, geralmente, o erro quadrático (Faceli et al., 2021).

Para prever uma instância desconhecida, percorre-se a árvore a partir do nó raiz em direção aos nós folhas, avaliando os sucessivos testes presentes nos nós de divisão, seguindo o ramo apropriado a partir de cada resultado. Ao atingir um nó folha, a instância é rotulada de acordo com o atributo-alvo atribuído à tal folha (Quinlan, 1986; Witten e Frank, 2017). Este modo de operação básico representa o funcionamento de diversos algoritmos de árvore de decisão disponíveis, como ID3 (*Iterative Dichotomiser 3*) (Quinlan, 1979), C4.5 (Quinlan, 1993) e CART (*Classification and Regression Tree*) (Breiman et al., 1984).

Uma etapa crítica no processo de construção das árvores de decisão é a realização da poda (*pruning*), a qual permite o combate ao sobreajuste (*overfitting*) (Russell e Norvig, 2020). Considerando uma árvore inicialmente completa, analisa-se, por exemplo, cada nó de teste que possui apenas nós folha como descendentes. Se o teste parecer irrelevante, detectando apenas ruído nos dados, então este teste é eliminado e substituído por um nó folha. Esse processo é repetido, considerando cada teste apenas com descendentes de folhas, até que cada um tenha sido podado ou aceito como está. A análise da relevância de um nó pode ser realizada a partir do

cálculo do seu ganho de informação. Em resumo, a poda funciona eliminando nós que não são claramente relevantes, trocando nós profundos por folhas (Breiman et al., 1984; Quinlan, 1987).

## 2.2.2 Modelos Múltiplos Preditivos

Uma estratégia interessante a ser adotada em busca de melhores resultados para as tarefas de classificação e regressão é a combinação de preditores, sendo o desempenho dessa combinação geralmente superior ao melhor modelo individual. Esse método é conhecido como modelos múltiplos preditivos, ou simplesmente como *ensemble* (Zhou, 2012; Kuncheva, 2014). Sua principal motivação está relacionada ao fato de que diferentes algoritmos podem realizar diferentes previsões devido às características específicas de cada algoritmo e, desse modo, resultar em melhores desempenhos quando considerados em conjunto (Aggarwal, 2015; Kubat, 2017; Russell e Norvig, 2020).

Esse método pode ser dividido em três fases (Britto et al., 2014; Cruz et al., 2018a): geração (Subsubseção 2.2.2.1), seleção (Subsubseção 2.2.2.2) e integração (Subsubseção 2.2.2.3). Cabe destacar que essa representação não é única, visto que as fases de seleção e integração podem ser facultativas, uma vez que todo o conjunto de preditores pode ser empregado sem a necessidade de seleção, assim como um único preditor pode ser selecionado do conjunto, sendo, neste caso, desnecessária a fase de integração (Mendes-Moreira et al., 2012; Britto et al., 2014).

### 2.2.2.1 Fase de Geração

Esta fase é responsável pelo treinamento de um conjunto de preditores  $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$ , em que  $N$  é o número de estimadores. Neste aspecto, o conjunto de preditores (classificadores ou regressores) pode ser gerado tanto por uma composição diversificada de algoritmos de indução distintos (heterogêneos) quanto por inúmeras instâncias de um mesmo algoritmo base (homogêneos). Neste último caso, o treinamento ocorre usualmente em diferentes subconjuntos aleatórios do conjunto de treinamento.

É digno de nota que uma condição necessária para que esse método seja útil é que os preditores individuais tenham um nível substancial de desacordo, ou seja, cometam erros independentes, visto que o emprego de avaliadores idênticos é inútil. Assim sendo, dois aspectos principais que merecem destaque são (Faceli et al., 2021):

- Combinação de preditores que cometem erros não correlacionados ou negativamente correlacionados;
- Cada avaliador deve ter um desempenho melhor do que uma escolha aleatória.

As principais técnicas de treinamento *ensemble* estão descritas na sequência e representadas na Figura 2.6 (Zhou, 2012; Kuncheva, 2014; Aggarwal, 2015):

- **Bagging**: acrônimo no inglês para *bootstrap aggregating*, realiza a repartição da base de treinamento em múltiplas partições, as quais são empregadas para o treinamento de cada preditor individualmente (Breiman, 1996; Domingos, 1997). Cabe mencionar que a amostragem é realizada com reposição (*bootstrapping*). Ou seja, uma vez que uma instância seja selecionada para compor o subconjunto de treinamento  $\mathcal{T}_i$ , ela retornará para o conjunto de treinamento original  $\mathcal{T}$ , o que significa que este exemplo estará disponível para ser selecionado por  $\mathcal{T}_{i+1}$  com a mesma probabilidade que qualquer outro exemplo de  $\mathcal{T}$  (Kubat, 2017). Portanto, a saída do *bagging* são  $N$  conjuntos de treinamento  $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_N$  responsáveis pelo treinamento de cada modelo  $f_i \in \mathcal{F}$  ( $1 \leq i \leq N$ );

- **Boosting**: realiza o treinamento sequencial de vários preditores fracos para gerar um preditor forte, de modo que cada um procura corrigir seu antecessor ao associar pesos para cada instância de treinamento (Schapire et al., 1998). Um preditor fraco é um modelo de aprendizado de máquina que tem um desempenho ligeiramente melhor do que uma escolha aleatória. Já um preditor forte é um modelo que tem um desempenho muito bom na tarefa em questão, geralmente capaz de fazer previsões com alta precisão. Assim, quanto maior o peso de um exemplo, maior é a importância atribuída a ele durante o treinamento. A ideia básica é focar nas instâncias incorretamente previstas de modo a aumentar o peso das mesmas. Para isso, a técnica inicia atribuindo pesos uniformes para todos os exemplos (*i.e.*, um conjunto de treinamento normal). Na sequência do treinamento, as instâncias que foram incorretamente previstas terão seus pesos aumentados enquanto os exemplos corretamente previstos terão um decréscimo em seus pesos (Russell e Norvig, 2020). O classificador/regressor final agrega os preditores aprendidos em cada iteração pela votação ponderada, sendo que o peso de cada preditor é uma função da sua precisão (Faceli et al., 2021);
- **Random Subspace Method (RSM)**: similar ao *bagging*, incluindo que os atributos também são selecionados aleatoriamente com reposição. Evita que atributos com alto poder de discriminação sejam sempre utilizados, resultando em uma maior diversidade dos preditores (Ho, 1998; Sammut e Webb, 2017);

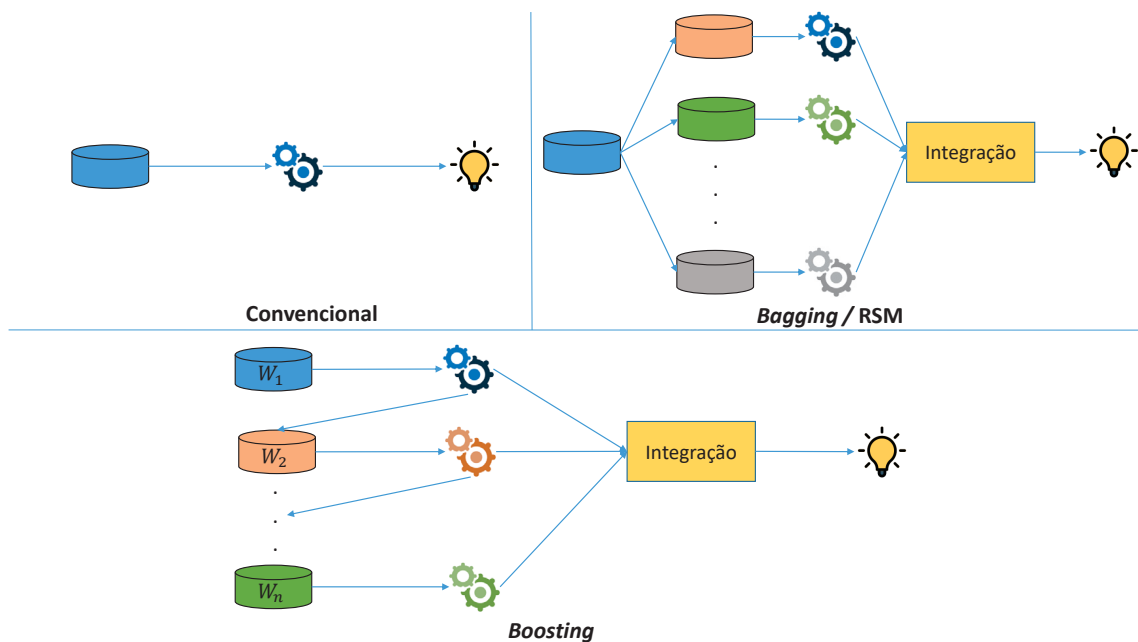


Figura 2.6: Representação da abordagem convencional e das principais técnicas *ensemble*

Fonte: autoria própria

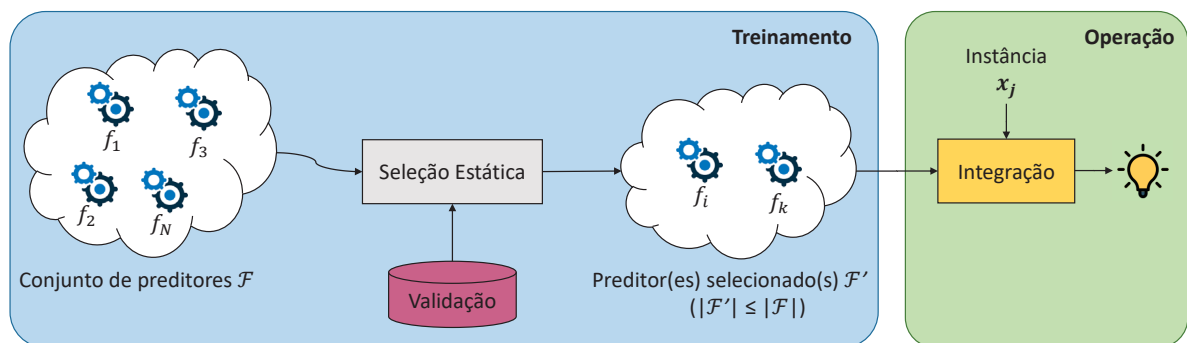
Apesar das técnicas *bagging* e RSM estarem comumente atreladas a construção de grupos de preditores homogêneos, elas também pode ser usadas para a geração de conjuntos heterogêneos. Em agrupamentos homogêneos, a diversidade é alcançada pela variação nos hiperparâmetros de configuração (valores iniciais durante a construção dos modelos), pelos diferentes subconjuntos de dados de treinamento (*bagging* e *boosting*) e pelos diferentes espaços de atributos (RSM). Por outro lado, a ideia por trás da abordagem heterogênea é a obtenção de especialistas que difiram entre si em termos de suas propriedades e conceitos em que são

baseados (Britto et al., 2014), como os simbólicos (árvores de decisão), os baseados em distância (KNN) e os conexionistas (redes neurais) (Kononenko e Kukar, 2007; Faceli et al., 2021).

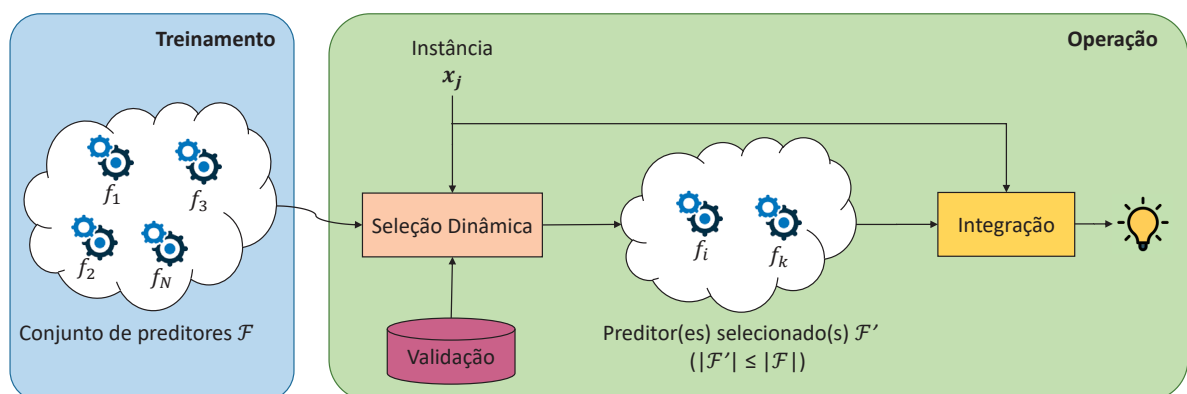
### 2.2.2.2 Fase de Seleção

Esta fase é responsável pela seleção de um subconjunto de modelos  $\mathcal{F}'$  do *ensemble* original  $\mathcal{F}$ , sendo que  $\mathcal{F}' \subseteq \mathcal{F}$ . Tal seleção pode ocorrer de duas maneiras (Figura 2.7):

- **Estática:** dado o conjunto inicial de preditores, um subconjunto é selecionado de modo que este mesmo agrupamento de modelos será usado para prever o rótulo para todas as amostras de teste  $x_j$  (Figura 2.7(a)). A solução ótima encontrada para o conjunto de validação é fixada e usada para a predição das instâncias desconhecidas (teste). Todo o esforço está voltado na extração da melhor combinação a partir dos desempenhos obtidos para o conjunto de validação rotulado (Ruta e Gabrys, 2005; Cruz et al., 2018a);
- **Dinâmica:** um único modelo ou um subconjunto de preditores do *ensemble* é selecionado em tempo real e separadamente para cada exemplo de teste (Figura 2.7(b)). Esta abordagem será descrita detalhadamente em uma subseção dedicada com foco na tarefa de regressão (Subseção 2.2.4) devido a sua relevância para o nosso trabalho.



(a) Seleção estática



(b) Seleção dinâmica

Figura 2.7: Representações genéricas das estruturas para a seleção (a) estática e (b) dinâmica de preditores, sejam eles classificadores ou regressores

Fonte: adaptado de Cruz et al. (2018a)

Vale salientar que, conforme mencionado, a fase de seleção pode ser opcional, visto que algoritmos *ensemble* tradicionais como Floresta Aleatória (*Random Forest* – RF) e Gradient Boosting não realizam tais tipos de seleção.

### 2.2.2.3 Fase de Integração

Por fim, esta fase é encarregada pela tomada de decisão final, realizada a partir da agregação das previsões dos modelos previamente selecionados. Caso apenas um modelo tenha sido selecionado na fase anterior, a integração não é necessária. As previsões podem ser unificadas (integradas) a partir de diferentes critérios, como voto majoritário (classificação) e média aritmética (regressão). Considerando o emprego da seleção dinâmica na fase anterior, é possível que estratégias mais flexíveis e arrojadas sejam desempenhadas, como a inclusão de um peso na previsão de cada modelo selecionado. Na Subseção 2.2.4 relatamos com mais detalhes este aspecto para o cenário da regressão.

Outra alternativa para a combinação é o uso do método Generalização em Pilha (*Stacked Generalization* ou simplesmente *Stacking*) (Wolpert, 1992; Rooney e Patterson, 2007). Todavia, este método não considera a Fase de Seleção, utilizando todos os modelos do *pool*, além de que seu emprego usualmente está atrelado ao uso de modelos heterogêneos (Faceli et al., 2021). Em complemento, o *Stacking* não determina como os modelos são gerados, apenas como são combinados.

O *Stacking* utiliza pelo menos dois níveis (camadas) de classificação/regressão, conforme ilustrada pela Figura 2.8. Neste caso, o conjunto de treinamento é dividido em dois subconjuntos: A e B. O subconjunto A é utilizado para treinar os preditores na primeira camada, os quais são na sequência empregados para realizar previsões no conjunto B. Os resultados das previsões de cada algoritmo são então usados como atributos de entrada para o algoritmo na segunda camada, denominado de *blender*, *meta-learner* ou, ainda, *meta-model* (Vilalta e Drissi, 2002; Witten e Frank, 2017; Géron, 2021). Para treinar o preditor *blender*, uma abordagem comum é usar um conjunto *hold-out*. O resultado desse último preditor é apresentado como a saída do modelo. Se desejado, mais níveis mesclados de *blenders* podem ser utilizados, conseqüentemente a base de treinamento deverá conter mais repartições a cada nova camada.

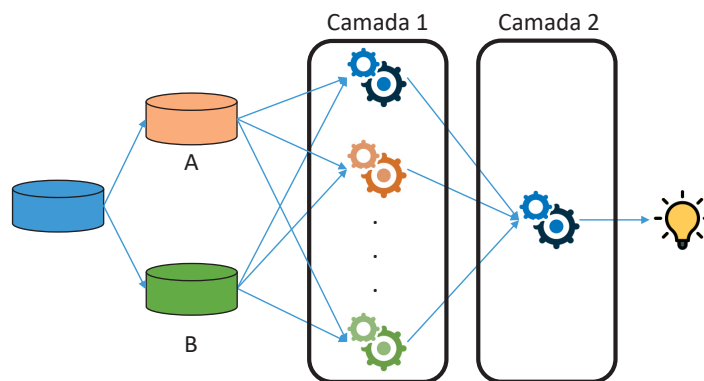


Figura 2.8: Representação do método *Stacking*

Fonte: autoria própria

### 2.2.3 Árvores Extremamente Aleatórias

O algoritmo Árvores Extremamente Aleatórias (Árvores Extras, *Extremely Randomized Trees* ou, ainda, *Extra Trees* – ERT) é um preditor constituído por um conjunto de árvores de decisão independentes, em que, por padrão durante a fase de geração, todo o conjunto de treinamento original é usado para construir individualmente cada árvore do *ensemble* (*i.e.*, sem *bootstrapping*) (Geurts et al., 2006). Em problemas de classificação, esse algoritmo realiza a integração das saídas das árvores ( $y_i, 1 \leq i \leq N$ ) em uma única predição ( $y$ ), para um determinado padrão de teste  $x_j$ , por meio do voto majoritário. Já em problemas de regressão, a estratégia utilizada compreende a média aritmética para combinar todas as predições. É digno de nota que, neste algoritmo, a fase de seleção não é realizada. Na Figura 2.9 está representada a estrutura do algoritmo ERT.

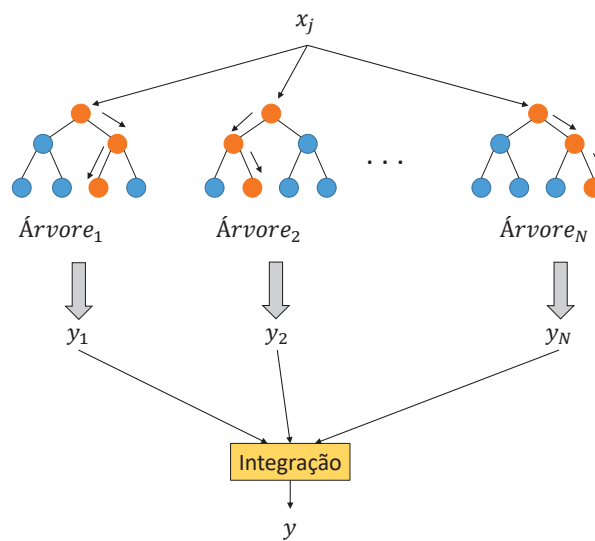


Figura 2.9: Representação do algoritmo ERT

Fonte: autoria própria

Assim como o algoritmo RF, o ERT introduz aleatoriedade na criação das árvores: em vez de buscar a melhor característica ao dividir um nó, ele busca entre um subconjunto aleatório dessas características, resultando em uma grande diversidade de árvores. Como resultado, evita que uma determinada característica predomine sobre as demais, uma vez que apenas parte dos atributos está disponível para cada árvore durante seu processo de criação. Ao mesmo tempo que o algoritmo busca minimizar o viés, ele também diminui a variância e controla o sobreajuste (*overfitting*) (Géron, 2021).

Um ponto chave que o distingue de outros métodos *ensemble* baseados em árvore, como o próprio RF, é o modo em que os limiares são definidos. Enquanto o RF busca determinar os melhores limiares possíveis, o ERT define os pontos de corte aleatoriamente, selecionando o melhor desses limites gerados aleatoriamente como regra de divisão. Consequentemente, essa propriedade torna o algoritmo ERT mais rápido de ser treinado comparado ao RF. Entre os principais parâmetros do algoritmo estão o número de estimadores (árvores) no *ensemble* e o número máximo de atributos (Igual e Seguí, 2017).



## 2.2.4 Seleção Dinâmica de Regressores

A Seleção Dinâmica de Regressores (SDR) realiza a seleção do(s) modelo(s) mais eficiente(s) a partir de um conjunto (*pool*) de regressores para estimar o valor do atributo-alvo individualmente para cada padrão (instância) de entrada, conforme ilustrada pela Figura 2.7(b). Para isso, espera-se que cada modelo seja um especialista em uma região do espaço de atributos, conhecido como região de competência (RC). Neste contexto, alguns problemas centrais da SDR que merecem destaque são: (i) determinação do conjunto de validação do qual será obtida uma RC; (ii) demarcação da RC; (iii) definição do critério para mensurar a competência de cada modelo do *ensemble* na RC; e (iv) definição do critério para selecionar o(s) estimador(es) mais adequado(s) para rotular uma dada instância desconhecida (Cruz et al., 2018a; Moura et al., 2019b). Outras duas questões que também merecem destaque dentro do cenário da SDR são a criação (treinamento) do *pool* de regressores e a combinação das previsões, as quais contemplam as fases de geração e integração, respectivamente.

Conforme discutido na Subsubseção 2.2.2.1, diferentes técnicas podem ser usadas para o treinamento do conjunto de preditores, como o *bagging* e o RSM. Uma alternativa é o uso dos modelos treinados a partir de algoritmos *ensemble* como RF (Breiman, 2001) e ERT (Geurts et al., 2006), conforme realizado por alguns trabalhos na literatura (Rodrigues et al., 2023; Ensina et al., 2024). Logo, considere um problema de regressão em que  $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$  seja um *ensemble* com  $N$  especialistas treinados a partir de uma dessas técnicas. A seleção dinâmica pode ser vista como uma divisão do espaço de atributos em  $Z > 1$  regiões de competência, denotado por  $R_1, R_2, \dots, R_Z$ . Desta maneira, para cada região  $R_j, j = 1, 2, \dots, Z$ , ao menos um preditor de  $\mathcal{F}$  manifesta desempenho superior do que os demais em  $R_j$  (Sergio et al., 2016).

Em síntese, a seleção dinâmica consiste nas seguintes etapas (Moura, 2019), as quais serão examinadas na sequência:

1. Dada a instância de teste  $x_j$ , determinar a RC com os  $K$  padrões mais semelhantes (*i.e.*,  $K$  vizinhos mais próximos considerando o emprego do algoritmo KNN) do conjunto de validação  $\mathcal{V}$  ou do conjunto de treinamento  $\mathcal{T}$ ;
2. Selecionar um subconjunto  $\mathcal{F}'$  do *ensemble*  $\mathcal{F}$  ( $\mathcal{F}' \subseteq \mathcal{F}$ ), de acordo com o desempenho individual de cada modelo na RC;
3. Obter a previsão  $f_i(x_j)$  de cada modelo selecionado  $f_i \in \mathcal{F}'$ ;
4. Obter a previsão final  $\mathcal{F}'(x_j)$ . Caso mais de um modelo tenha sido selecionado, o resultado será a combinação das previsões dos estimadores do subconjunto  $\mathcal{F}'$ .

A abordagem padrão para obter as similaridades entre os dados e determinar a RC é pelo algoritmo KNN utilizando a distância euclidiana (Cruz et al., 2018a). Considerando o padrão de teste  $x_j$ , sua RC é um agrupamento  $\Psi$  composto dos  $K$  vizinhos mais próximos de  $x_j$  no conjunto de validação (ou treinamento) em que  $\Psi \subset \mathcal{V}$ . Em outras palavras, a RC é a região de exemplos que cercam a instância  $x_j$ . A escolha pelo tamanho da vizinhança  $K$  é crucial para o desempenho do sistema. Mendes-Moreira et al. (2009) demonstraram por meio de seu estudo que o tamanho da vizinhança é dependente do problema ao analisarem diversos tamanhos em diferentes conjuntos de dados. Na Figura 2.10 está representada a definição da RC pelo uso do KNN, considerando dois valores distintos para o tamanho da vizinhança ( $K = 3$  e  $K = 7$ ).

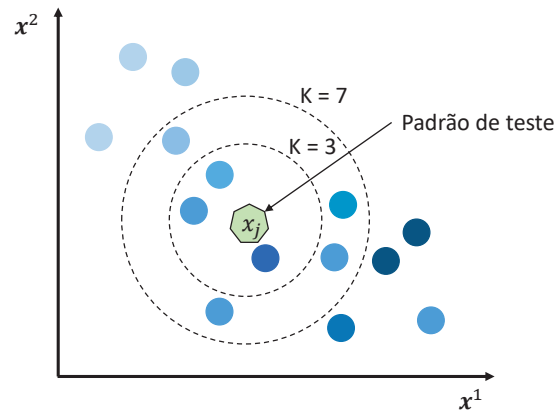


Figura 2.10: Ilustração do algoritmo KNN para dois tamanhos de vizinhança ( $K = 3$  e  $K = 7$ ). Considere a intensidade das cores dos círculos como uma representação de seus valores de atributo-alvo para a regressão, de modo que quanto mais escura a cor azul, maior é o seu valor numérico

Fonte: autoria própria

Uma outra alternativa ao KNN para a definição da RC é o emprego de algoritmos de clusterização, como o K-Means (Qiao e Wang, 2021). Por outro lado, Mendes-Moreira et al. (2009) propuseram um método que não é baseado em distância para definir a RC, o que flexibiliza a estratégia em relação a definição do tamanho da RC (hiperparâmetro  $K$  do algoritmo KNN). Sua abordagem seleciona as instâncias semelhantes usando o algoritmo CART, de modo que, para cada exemplo de teste, determina-se a folha da árvore em que o respectivo exemplo se enquadra. Assim, os exemplos de treinamento contidos nessa folha são os selecionados para a RC.

A estratégia SDR pode ser dividida em três diferentes técnicas conforme listadas a seguir (Rooney et al., 2004; Mendes-Moreira et al., 2009), as quais podem ou não incluir as fases de seleção e integração, dependendo do respectivo esquema:

- **Seleção Dinâmica (*Dynamic Selection – DS*):** seleciona apenas o regressor que apresentar o menor erro cumulativo na RC. Os erros podem ser ponderados pela distância entre a instância de teste e os exemplos da RC (Equação 2.1). Como um único modelo é selecionado, não há necessidade de combinação de modelos;
- **Ponderação Dinâmica (*Dynamic Weighting – DW*):** combina todos os regressores do *ensemble* usando a média ponderada, sem realizar qualquer seleção. Para cada padrão em  $\Psi$ , uma distância ponderada  $d_k$  é calculada utilizando a Equação 2.1, atribuindo maior importância para os vizinhos mais próximos:

$$d_k = \frac{\frac{1}{dist_k}}{\sum_{j=1}^K \left(\frac{1}{dist_j}\right)} \quad (2.1)$$

sendo que  $dist_k$  é a distância entre um padrão  $t_k \in \Psi$  e o padrão de teste  $x_j$ . O vetor  $\{d_1, d_2, \dots, d_k\}$  é usado para calcular o peso  $\alpha_i$  do regressor  $f_i$  pela Equação 2.2:

$$\alpha_i = \frac{\frac{1}{\sqrt{\sum_{k=1}^K (d_k \times err_{k,i})}}}{\sum_{n=1}^N \left(\frac{1}{\sqrt{\sum_{k=1}^K (d_k \times err_{k,n})}}\right)} \quad (2.2)$$



onde  $\sum \alpha_i = 1$ ,  $N$  é o tamanho do *ensemble* (número de modelos),  $k$  representa o índice do vizinho e  $err_{k,i}$  é o erro do  $i$ -ésimo regressor calculado usando o padrão  $t_k \in \Psi$ . Portanto, a predição final corresponde a Equação 2.3:

$$\mathcal{F}(x_j) = \sum_{i=1}^N f_i(x_j) \times \alpha_i \quad (2.3)$$

onde  $\mathcal{F}(x_j)$  é a decisão final do *ensemble*,  $f_i(x_j)$  e  $\alpha_i$  são a predição do regressor  $f_i$  e seu respectivo peso  $\alpha_i$  para o padrão de teste  $x_j$ . A utilização de pesos costuma apresentar melhores resultados em comparação com a utilização da média simples (Mendes-Moreira et al., 2009);

- **Seleção Dinâmica Ponderada (*Dynamic Weighting Selection – DWS*):** seleciona um subconjunto de modelos antes de executar os passos estabelecidos para DW. A literatura usualmente usa a abordagem padrão definida por Rooney et al. (2004): os modelos selecionados são aqueles com erro cumulativo inferior ao intervalo  $\frac{E_{max}-E_{min}}{2}$ , onde  $E_{max}$  e  $E_{min}$  são os maiores e os menores erros cumulativos de qualquer um dos regressores no *ensemble*  $\mathcal{F}$ , respectivamente. Uma alternativa é o critério usado por Mendes-Moreira et al. (2009), em que os modelos selecionados são aqueles com erro inferior a um limiar percentual relacionado ao melhor modelo (e.g., 10%, 30% e 90%). Contudo, esta última estratégia é mais complexa devido ao fato de que diversos limites percentuais devem ser analisados para determinar o melhor valor para uma dada base de dados (Mendes-Moreira et al., 2015).

Um fator chave na seleção dinâmica é a definição de uma métrica para representar o desempenho de cada modelo do *pool* na RC, conhecida como medida de competência. Esta métrica é importante para servir de referência às técnicas mencionadas anteriormente para a etapa de seleção (ou seja, DS e DWS), que selecionarão apenas o(s) modelo(s) que satisfaça(m) um determinado critério. Moura et al. (2019b) avaliaram oito medidas de competência utilizando as técnicas DS, DW e DWS, demonstrando que a escolha da melhor medida depende do problema.

Entretanto, a investigação de outros hiperparâmetros além da medida de competência e do tamanho da RC ainda é negligenciada pela literatura da SDR. Neste aspecto, é possível que a utilização de medidas de distância diferentes da euclidiana, bem como outros critérios para selecionar os regressores pela técnica DWS, também possam impactar positivamente no desempenho de um método SDR. Estas análises objetivam explorar a distribuição diversificada e particular dos dados de modo a atingir a RC mais adequada. Ao menos para o problema de classificação, existe uma correlação entre a qualidade dos vizinhos selecionados para compor a RC e os modelos selecionados para rotularem o atributo-alvo, afetando diretamente o desempenho final do sistema (Cruz et al., 2011, 2019; Wang e Mao, 2020).

### 2.3 CONSIDERAÇÕES FINAIS

O transporte de energia elétrica das usinas geradoras até os consumidores finais compreende uma série de elementos complexos, com destaque para as linhas de transmissão. Em particular, essas linhas percorrem longas distâncias e estão sujeitas à diferentes adversidades que podem afetar seu pleno funcionamento, como tempestades, queimadas e, inclusive, vandalismo.

Frente a isso, o sistema de proteção do SEP pode ser beneficiado pelo emprego de algoritmos de AM para o contexto de análise de falhas em LTs, como, em especial, a classificação do tipo de falta e a sua localização na linha. Ainda, esses algoritmos podem ser empregados tanto individualmente quanto combinados (*ensemble*), permitindo que uma diversidade de métodos sejam propostos e desenvolvidos no estado da arte.

Em síntese, apresentamos neste capítulo os conceitos básicos relacionados às linhas aéreas de transmissão (Seção 2.1). Sobretudo, elucidamos as propriedades e as características das faltas que podem ocorrer nas linhas, assim como uma visão geral do sistema de proteção do SEP, responsável por resguardá-lo dos efeitos que as falhas podem causar. Além disso, descrevemos os conceitos básicos relacionados ao aprendizado de máquina (Seção 2.2), identificando os algoritmos e as estratégias que serão utilizados neste trabalho para a análise de falhas.

No próximo capítulo, relatamos uma revisão sobre os trabalhos relacionados que estão disponíveis na literatura científica para o diagnóstico de faltas em LTs, particularmente sobre a classificação e a localização, tarefas de particular interesse do nosso projeto.

### 3 ESTADO DA ARTE

Nos capítulos anteriores, apresentamos a introdução e a motivação deste trabalho, bem como a fundamentação teórica referente aos principais conceitos envolvidos no tema do projeto. Em especial, contextualizamos sobre o SEP e os componentes básicos que fazem parte das linhas de transmissão, além da descrição detalhada sobre as falhas que podem afetar tais linhas. Por fim, mas não menos importante, discutimos conceitos fundamentais sobre algoritmos de aprendizado de máquina, amplamente utilizados atualmente para a análise de falhas. A explanação prévia desses aspectos é necessária para fornecer o embasamento teórico indispensável para o entendimento do escopo do problema tratado neste projeto: classificação e localização de falhas em linhas de transmissão.

Neste cenário, diversos métodos estão disponíveis na literatura científica sobre a análise de faltas, incluindo trabalhos que realizam revisões e mapeamentos bibliográficos sobre este tema (Yadav e Dash, 2014; Aleem et al., 2015; Singh e Vishwakarma, 2015; Chen et al., 2016; Ferreira et al., 2016; Mishra e Ray, 2018; Prasad et al., 2018; Raza et al., 2020; Mukherjee et al., 2021; Kanagasabapathy, 2021; Shakiba et al., 2023). Em complemento, realizamos uma busca abrangente por trabalhos relacionados a partir do protocolo estabelecido no Apêndice A.

Na primeira seção deste capítulo (Seção 3.1), apresentamos uma breve revisão dos métodos encontrados na literatura que aplicam algoritmos de AM para a classificação e/ou a localização de falhas em linhas de transmissão, de modo a exprimir as principais ideias já propostas e utilizadas no estado da arte. Ao final (Seção 3.2), discutimos as principais informações extraídas da literatura sobre o tema, destacando os problemas identificados e os pontos passíveis de serem explorados por novos trabalhos, de modo a justificar e contextualizar as contribuições que nosso projeto visa alcançar.

#### 3.1 TRABALHOS RELACIONADOS

Nesta seção estão apresentados os trabalhos encontrados no estado da arte que visam a classificação e a localização de faltas em linhas aéreas de transmissão. Em particular, esses trabalhos relatam diferentes abordagens para a realização dessas atividades, mas manifestam duas propriedades em comum: o uso de dados de oscilografias (tensão e/ou corrente) e a aplicação de algoritmos de aprendizado de máquina. Outro aspecto frequente é o emprego de dados de oscilografias provenientes de um único terminal (extremidade) da linha de transmissão.

Na Figura 3.1 está representado um diagrama simplificado com as principais atividades de análise de falhas. Primeiramente, os sinais de corrente e tensão são amostrados e fornecidos para o módulo de extração de atributos. Este módulo então extrai características usadas pelas tarefas de detecção, classificação e localização de falta. As saídas são o tipo e a estimativa da posição de ocorrência da falta na LT fornecidas pelas tarefas de classificação e localização, respectivamente. Alguns dos trabalhos cobrem todos os três aspectos, enquanto outros se concentram em um ou dois deles. Conforme pode ser observado, a tarefa de detecção da falta pode ser omitida uma vez que esta tarefa pode ser realizada informalmente pelos próprios autores. A detecção de falta destina-se a determinação do instante de início da falha na oscilografia, permitindo, então, que ciclos pré e/ou pós-falta sejam segmentados para a realização das tarefas de classificação e localização. Devido ao foco deste projeto estar relacionado a classificação e a localização de falhas, a tarefa de detecção não será abordada. Cabe ressaltar que diversas

obras na literatura também não realizam a detecção da falta, uma vez que tal problema já está, de maneira geral, bem consolidado (Ferreira et al., 2016).

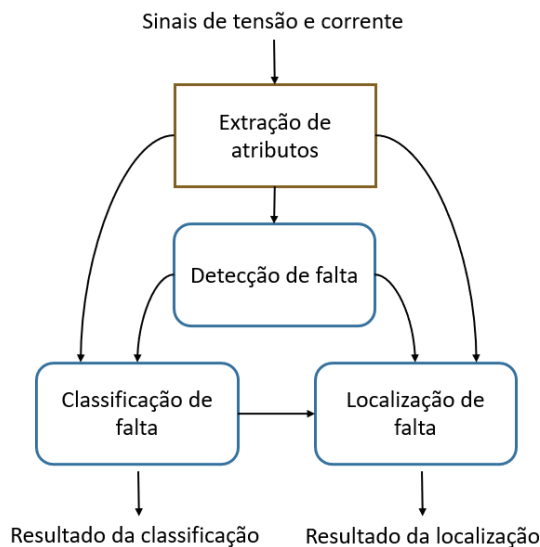


Figura 3.1: Diagrama simplificado das principais atividades de análise de faltas

Fonte: adaptado de Chen et al. (2016)

Os autores frequentemente propõem métodos individuais para o diagnóstico de faltas, ou seja, nas obras estão relatadas abordagens que desempenham apenas a classificação ou apenas a localização de falhas. Por outro lado, há trabalhos em que estão descritos métodos que realizam ambas as atividades simultaneamente. De modo a tornar as descrições destes trabalhos mais organizada e clara, inicialmente apresentamos as obras que tratam apenas da classificação, sendo na sequência relatados os trabalhos de localização e, ao final, as obras que realizam ambas as atividades conjuntamente.

Silva et al. (2006) e Silva et al. (2007) introduziram um método para classificação de falhas dividido em quatro etapas: (i) normalização, (ii) reamostragem, (iii) janelamento e (iv) classificação. Inicialmente, os autores (i) normalizaram as amostras dos sinais da tensão e da corrente de fase e de neutro pelos respectivos valores máximos encontrados em todo o registro, tornando os valores das formas de onda no intervalo entre  $[-1, 1]$ . Em seguida, os dados foram (ii) reamostrados para a frequência de 1,2 kHz e uma (iii) janela de dados de um quarto de ciclo foi fornecida como entrada para uma (iv) MLP (*Multilayer Perceptron*) classificar o tipo de falta. Essa janela foi deslizada ao longo de todo o registro após a incidência da falha com passos de uma amostra por vez, classificando cada janela individualmente. Assim, a falha que mais prevalecer ao longo das predições é aquela determina como a decisão final.

Morais et al. (2010) desenvolveram um *framework* para classificar faltas em linhas de transmissão, além de disponibilizarem publicamente o conjunto de dados utilizado no estudo. A base de dados, denominada UFPAFaults, contém 27.500 eventos de falha provenientes de três linhas de transmissão (18,5 km, 151,5 km e 325 km) de um sistema de 500 kV. Todos os parâmetros para cada simulação foram randomicamente gerados a uma taxa de amostragem de 40 kHz, sendo tais dados subamostrados para 2 kHz para constituírem a base de dados. Cada exemplo de falha é composto pelas formas de onda trifásicas de tensão e corrente apenas para o período pós-falha, os quais foram submetidos a um filtro passa-baixa e normalizados para o intervalo  $[-1, 1]$ , além de serem rotulados unicamente para o tipo de falta.

Já o método proposto para a classificação compreende a decomposição por *wavelets* e a concatenação de seus coeficientes (Mallat, 2009), denominado pelos autores como *waveletconcat*. No trabalho, quatro algoritmos (MLP, árvore de decisão J48, KNN e SVM) foram avaliados, sendo determinada a combinação *waveletconcat* e SVM como a de melhor desempenho.

Asadi Majd et al. (2017) propuseram um método para a classificação de falhas a partir do uso do algoritmo KNN e da comparação com limiares a fim de determinar o tipo da falha. Para isso, três-quartos do sinal pós-falta da corrente para as três fases foram divididos pelo valor eficaz (*Root Mean Square – RMS*) obtido de meio ciclo da corrente anteriormente ao instante do início da falta, denominando de janela de análise. Em seguida, o algoritmo KNN foi utilizado para determinar o quinto vizinho mais próximo de cada amostra da janela de análise, tendo em vista obter a distância entre cada amostra e o respectivo vizinho. Desse modo, os autores determinaram a distância máxima para cada fase e compararam essas distâncias com valores de limiares previamente definidos com a finalidade de classificar o tipo da falta. A frequência de amostragem avaliada pelos autores foi de 10 kHz.

Chen et al. (2018a) apresentaram um método para classificar falhas por meio da extração automática de atributos pelo próprio algoritmo, evitando a necessidade da definição das características pelos autores. Assim sendo, meio ciclo para as três fases dos sinais da tensão e da corrente foram combinados para um único sinal multicanal para a taxa de amostragem de 20 kHz. Em seguida, esse sinal foi utilizado como entrada para o algoritmo Autocodificador Esparso Convolutivo (*Convolutional Sparse Autoencoder – CSAE*) realizar a extração de atributos a fim de o classificador *softmax* determinar a classe (tipo) da falta.

É digno de nota que os autores avaliaram o método proposto para outras frequências de amostragem entre 1,25 kHz e 20 kHz, assim como utilizando os sinais da tensão e da corrente à parte. Os desempenhos observados foram inferiores aos resultados obtidos para a proposta original. Ao utilizar os sinais separadamente, o desempenho degradou de maneira considerável, principalmente para a forma de onda da tensão.

Ye et al. (2020) relataram um método para classificação de falhas em linhas de transmissão com taxa de amostragem de 1 kHz. Os autores utilizaram cerca de dois ciclos pós-falta do sinal da corrente conjuntamente com a sequência-zero da corrente, combinados em um mesmo vetor. Por sua vez, esse vetor foi utilizado de entrada para uma Rede Neural Convolutiva (*Convolution Neural Networks – CNN*) de uma dimensão realizar a tarefa de classificação (LeCun et al., 2015). Para lidar com as diferentes frequências de amostragem, os autores utilizaram a interpolação cúbica segmentada, tornando assim o sinal de entrada sempre do mesmo tamanho, independentemente da taxa de amostragem utilizada pelos registradores de perturbação. Apesar dos autores citarem bom desempenho do modelo proposto para outras frequências de amostragem, eles não reportaram os respectivos resultados numéricos.

O método também foi avaliado utilizando o sinal da tensão no lugar da corrente, sendo constatada baixa taxa de acerto para classificar os tipos de falhas. É importante observar ainda que o algoritmo foi treinado e testado com dados de duas linhas de transmissão distintas. No entanto, o modelo não foi avaliado para outras LTs diferentes daquelas utilizadas para treinar o algoritmo. Ou seja, os autores não examinaram a competência do método para generalizar a classificação de falhas para linhas cujos dados não foram usados durante o treinamento.

O trabalho desenvolvido por Haq et al. (2021) para a classificação de faltas decompõe os sinais trifásicos da tensão e da corrente pela técnica transformada discreta de *wavelet* (*Discrete Wavelet Transform – DWT*) de nove níveis (Mallat, 2009), sendo calculados os valores de energia e desvio padrão para cada coeficiente resultante da decomposição. O vetor de características, composto pelos atributos previamente citados, serviram de entrada para duas Máquinas de Aprendizado Extremo (*Extreme Learning Machine – ELM*) classificarem o tipo de falha (Huang

et al., 2006). Em particular, uma ELM foi designada para detectar o envolvimento das fases (A, B, C), enquanto a outra ELM identificava o envolvimento da terra na falha. A frequência de amostragem usada para a coleta dos dados foi de 20 kHz.

Em complemento, os autores avaliaram o método proposto para outras taxas de amostragem com valores entre 1,95 kHz e 20 kHz, além de empregar os sinais da tensão e da corrente separados. Consequentemente, os resultados obtidos demonstraram que o desempenho da proposta degradou para as frequências inferiores a 20 kHz para todos os cenários avaliados, sobretudo quando o sinal da tensão foi utilizado isoladamente.

Coban e Tezcan (2021) desenvolveram um método que emprega apenas meio ciclo trifásico pós-falta do sinal da corrente como vetor de atributos. Em seguida, os autores utilizaram a técnica de análise de componentes de vizinhança (*neighborhood component analysis*) para a seleção de atributos, de modo que, na sequência, o algoritmo SVM realizasse a classificação do tipo de falta. Os autores examinaram o método proposto com dados coletados a 20 kHz.

Bhadra et al. (2021) introduziram uma nova arquitetura de classificação ao converterem as séries temporais de tensão e corrente para uma imagem. Inicialmente, o método aplica a transformada *wavelet* nos sinais com o objetivo de eliminar ruídos. Em seguida, as séries temporais são apresentadas nesta análise como o campo angular Gramiano (*Gramian Angular Field* – GAF), significando uma representação no sistema de coordenadas polares dos dados da série temporal em vez de apenas no sistema cartesiano (Wang e Oates, 2015). Em seguida, essa imagem gerada é usada como entrada para uma CNN. Os autores analisaram sua proposta a uma frequência de amostragem de 20 kHz.

Fonseca et al. (2022) relataram um método de classificação de faltas que combina o filtro de *notch* com o algoritmo RF. Após a aplicação do filtro nos sinais trifásicos da tensão, os dados são normalizados para o intervalo entre [-1, 1]. Na sequência, o método segmenta apenas meio ciclo pós-falta para cada fase, sendo ao final concatenados para serem usados como entrada no modelo RF com 50 árvores. A taxa de amostragem foi de 15,36 kHz.

Na sequência desta seção, descrevemos trabalhos selecionados que objetivam a localização das faltas em linhas de transmissão. Ekici et al. (2008) utilizaram a técnica transformada de pacote de *wavelets* (*Wavelet Packet Transform* – WPT) para decompor os sinais de corrente e tensão (Mallat, 2009), considerando meio ciclo pré-falta e meio ciclo pós-falta provenientes de um único terminal. Em seguida, os coeficientes provenientes do WPT foram usados para calcular os atributos de energia e entropia para cada sinal separadamente. Essas características foram então utilizadas como entrada para uma MLP localizar a posição da falha na linha de transmissão. A frequência usada pelos autores para a coleta de dados foi de 10 kHz.

Ray et al. (2013) desenvolveram um método para localizar falhas em linhas de transmissão, considerando a coleta de dados para uma frequência de amostragem de 30 kHz. Inicialmente, ciclos únicos dos sinais trifásicos pós-falta da tensão e da corrente foram decompostos pelas técnicas de processamento de sinais DWT e WPT. Em seguida, os autores extraíram seis atributos de cada coeficiente resultante das técnicas previamente citadas: energia, média, desvio padrão, curtose, assimetria e entropia. Em complemento, o processo de seleção de atributos determinou as melhores características para cada técnica separadamente, sendo 21 atributos selecionados para a DWT e 32 para a WPT, os quais foram normalizados para valores entre [0, 1].

Por sua vez, esses conjuntos de características serviram de entrada para uma rede neural MLP localizar a falha na linha de transmissão. Os resultados obtidos demonstraram que o método composto por DWT-MLP apresentou os melhores resultados. Ainda, os autores relataram que no estudo outras taxas de amostragem foram analisadas, porém a frequência deve ser de no mínimo 30 kHz para atingir um bom desempenho.



Swetapadma e Yadav (2015) propuseram um método para a localização de falhas por meio de dados coletados de um único terminal para a taxa de amostragem de 1 kHz. Para isso, os autores utilizaram um ciclo pré-falta e dois ciclos pós-falta dos sinais da tensão e da corrente para as três fases, os quais foram primeiramente submetidos ao filtro Butterworth para então serem processados pela técnica DWT (Zumbahlen, 2008). Em seguida, o método calcula os valores de desvio padrão dos coeficientes de aproximação de terceiro nível resultantes da decomposição DWT, totalizando um conjunto de seis atributos.

Posteriormente, essas características foram usadas como entrada para uma rede neural com três saídas que correspondem a localização da falha para cada fase. Para as faltas monofásicas, a localização apresentada foi aquela indicada para a respectiva saída da rede neural. Já para os demais tipos de falta, a localização exibida foi aquela com menor distância determinada pela saída do algoritmo em relação às fases envolvidas na falha.

Em outro estudo, Swetapadma e Yadav (2017) mostraram um método que utiliza uma árvore de decisão para estimar a distância da falta a partir de dados coletados de uma única extremidade da linha. A extração de características foi realizada considerando uma técnica de processamento de sinais para um ciclo pré-falta e um ciclo pós-falta para as formas de onda de corrente e tensão, calculando o desvio padrão dos sinais processados para obter a Decomposição de Valores Singulares (*Singular Value Decomposition – SVD*) de ambos os sinais, fornecida então como entrada para a árvore de decisão. Para tanto, os autores avaliaram duas técnicas de processamento de sinais: transformada discreta de Fourier (*Discrete Fourier Transform – DFT*) e DWT (Smith, 2002). Os resultados demonstraram que a combinação DWT com árvore de decisão teve um desempenho ligeiramente melhor.

É importante mencionar que o método exige quatro módulos de localização de faltas, um para cada classe de falta (LG, LL, LLG, LLL). De maneira semelhante, mas utilizando valores de desvio padrão como recursos, os mesmos autores apresentaram um método combinando o algoritmo KNN com a técnica DFT, também necessitando de quatro módulos (Swetapadma e Yadav, 2018). A demanda por mais de um modelo é uma desvantagem das abordagens, uma vez que exigem uma classificação do tipo de falta como pré-requisito para selecionar o módulo mais apropriado para estimar a localização da falta. A frequência de amostragem usada por ambos os estudos foi de 4 kHz.

Fei et al. (2018) desenvolveram um esquema para a localização de falhas baseado no algoritmo SVM, empregando apenas o sinal pós-falta da tensão de uma única extremidade do sistema. Uma janela de dados contendo um oitavo de ciclo foi submetida a um filtro passa-baixa para eliminar o ruído e então usada diretamente como vetor de características para o SVM. A amostragem considerada no estudo foi de 20 kHz.

Zhang et al. (2020) propuseram um método para localização de falhas em linhas de transmissão para uma frequência de amostragem de apenas 100 Hz. Em particular, os autores utilizaram dados trifásicos da corrente de ambos os terminais da linha entre 0,1 segundo antes e 0,1 segundo após a ocorrência da falta, formando assim uma matriz. Antes de utilizá-la de entrada para o algoritmo proposto, os dados dessa matriz foram padronizados visando a distribuição normal. Todavia, a respectiva técnica de pré-processamento não foi reportada no trabalho. Em seguida, essa matriz padronizada foi aplicada como entrada para o algoritmo Unidade Recorrente Fechada Bidirecional (*Bidirectional Gated Recurrent Unit – Bi-GRU*) (Schuster e Paliwal, 1997), cuja saída representava a percentagem da extensão da linha em que a falha foi identificada.

Cabe destacar que o algoritmo foi treinado e avaliado para linhas de transmissão de diferentes comprimentos, provenientes do sistema IEEE de 39 barramentos. Porém, os autores não avaliaram a capacidade do método para localizar falhas em linhas de transmissão distintas daquelas empregadas para treinar o algoritmo.

Valabhoju et al. (2021) desenvolveram um método usando um algoritmo *ensemble* para estimar a distância da falha a partir de uma série de etapas. Em síntese, a abordagem inicialmente aplica o filtro Butterworth para suavizar um ciclo pós-falta das formas de onda de corrente e tensão, sendo então usada a técnica DWT para a decomposição dos sinais. Na sequência, características de desvio padrão foram obtidos a partir dos sinais decompostos para usá-las como entrada para um conjunto de árvores de regressão (*bagging*), cujo número de regressores não é explicitamente definido pelos autores. Cabe citar ainda o emprego da técnica análise de componentes principais (*Principle Component Analysis – PCA*) para a redução da dimensionalidade e da normalização do conjunto de atributos. A amostragem usada para a coleta de dados foi de 1 kHz.

Bhatnagar et al. (2023) apresentaram um esquema de localização de faltas baseado em RF e *Teager-Kaiser Energy Operator* (TKEO) (Boudraa e Salzenstein, 2018). A energia Teager dos sinais de tensão e corrente de um sinal de ciclo pré-falta e de um sinal de ciclo pós-falta foi fornecida como entrada aos módulos regressores de RF, um para cada classe de falta, de modo semelhante aos trabalhos mencionados anteriormente. O número de árvores nos módulos RF varia dependendo do tipo de falta: 300 para LG, 100 para LL, 200 para LLG e 200 para LLL. O método foi avaliado considerando o sistema de potência da IEEE com 14 barramentos, sem especificar os detalhes da LT, para uma taxa de amostragem de 1 kHz.

Por fim, descrevemos a seguir as obras que propõem métodos que realizam tanto a tarefa de classificação quanto a de localização. Ray e Mishra (2016) utilizaram um ciclo trifásico pós-falta do sinal da corrente extraído para um único terminal da linha em análise com uma frequência de amostragem de 30 kHz. Cabe ressaltar que os autores reportam que outras taxas foram avaliadas, mas os resultados foram melhores para 30 kHz do que qualquer outra frequência.

Em seguida, esse ciclo pós-falta foi decomposto pela técnica de processamento de sinais WPT para extrair uma série de coeficientes, a partir dos quais foram calculados dois atributos para cada coeficiente: energia e entropia. Na sequência, o processo de seleção de atributos (*forward feature selection*) determinou as duas características mais relevantes do conjunto: energia para o coeficiente ADAD4 e entropia para o coeficiente DDDD4. Esses atributos ainda foram normalizados para a faixa de valores entre [-1, 1].

Em complemento, o método compreendeu o uso de cinco SVMs, sendo quatro para a classificação e uma para a localização da falta. Assim, a atividade de classificação contemplou o uso de uma SVM para cada fase (A, B e C) e uma quarta SVM conectada entre a fase e a terra em conjunto com um valor de índice. Desse modo, a saída combinada de cada SVM permitiu classificar o tipo de falta. Já para a localização, uma única SVM foi utilizada com os mesmos atributos citados no parágrafo anterior.

Chen et al. (2018b) propuseram um *framework* para a classificação e a localização de falhas com a extração de atributos integrada ao algoritmo utilizado para a realização de ambas as tarefas. Primeiramente, o método obtém ciclos únicos pós-falta do sinal trifásico da corrente de ambos os terminais para em seguida calcular a diferença entre eles. Esse sinal resultante foi então empregado como entrada para o algoritmo *Summation-Wavelet Extreme Learning Machine* (SW-ELM) (Huang et al., 2006; Javed et al., 2014), assim como para uma variação do SW-ELM denominada *Summation-Gaussian ELM* (SG-ELM) que foi proposta pelos autores. A frequência de amostragem usada foi de 3,6 kHz.

Em particular, esses algoritmos introduzem uma segunda função de ativação em cada neurônio, *Wavelet Morlet* para o SW-ELM e Gaussiana para o SG-ELM, sendo a saída representada pela média das funções de ativação. Cabe destacar que as saídas dos algoritmos consistem em um único valor, integrando o tipo da falta (valor inteiro entre 1 e 11) e a sua localização (valor decimal em relação a percentagem da distância em relação ao barramento



de envio). Os autores também avaliaram a proposta para diferentes configurações de linhas de transmissão. Contudo, os algoritmos foram treinados e avaliados separadamente para cada LT, ou seja, eles não generalizaram as atividades de classificação e localização.

Ferreira et al. (2020) apresentaram um método para classificação e localização de faltas utilizando dados de um único terminal da linha de transmissão, com uma taxa de amostragem de 100 kHz. Para isso, os autores extraíram seis atributos (RMS, máximo, mínimo, média, energia e integral numérica) a partir de três ciclos pré-falta e três ciclos pós-falta para as três fases dos sinais da tensão e da corrente, individualmente para cada combinação, totalizando 72 características. Na sequência, esses atributos serviram como entrada para duas redes neurais efetuarem as tarefas de classificação e localização.

Belagoune et al. (2021) apresentaram um método para classificação e localização de falhas para uma frequência de amostragem de apenas 60 Hz. Para isso, os autores utilizaram os dados trifásicos, magnitudes e ângulos, da tensão e da corrente provenientes de ambos os terminais da LT. É digno de nota que os autores utilizaram todo o conjunto de simulação, ou seja, eles não realizam a extração de ciclos pré e/ou pós-falta conforme outros trabalhos na literatura realizam. Cada simulação ocorreu por 10 segundos, sendo que a falha inicia 1 segundo após o início da simulação.

Logo, cada simulação contém 601 amostras para as quais foram empregados os dados de ângulos e magnitudes previamente citados. Desse modo, os autores construíram um vetor de atributos para cada ensaio de falta e o utilizaram como entrada para dois algoritmos do tipo Memória Longa de Curto Prazo (*Long Short-Term Memory* – LSTM) realizarem as atividades de classificação e localização.

Outros trabalhos na literatura também têm utilizado amplamente modelos de aprendizado profundo (*Deep Learning* – DL) para a classificação e/ou a localização de faltas (Kanagasabapathy, 2021). Em geral, esses métodos usam dados brutos de corrente e tensão como entrada para um modelo baseado em algoritmos como LSTM e também CNN (Fan et al., 2019; Swaminathan et al., 2021; Wang et al., 2022; Bon e Dai, 2022; Ensina et al., 2022a; Shakiba et al., 2022; Fang et al., 2023).

Sumarizamos as informações dos trabalhos descritos neste capítulo nas tabelas que estão dispostas a seguir. Na Tabela 3.1, destacamos as características das linhas de transmissão analisadas pelas respectivas obras, os tipos de sinais utilizados provenientes de oscilografias (*i.e.*, tensão e/ou corrente), o tamanho da janela considerada para extrair dados a serem processados pelos respectivos métodos, além da tarefa cujo trabalho foi projetado (classificação e/ou localização). Já na Tabela 3.2 relatamos informações voltadas as abordagens de predição em si, com foco na realização ou não de técnicas de processamento de sinais e dos atributos utilizados pelos autores a partir da janela de dados citada na tabela anterior, juntamente com o algoritmo de predição usado e de algumas eventuais observações sobre os métodos.

Tabela 3.1: Informações sumarizadas dos trabalhos selecionados para o estado da arte (parte 1)

<b>Referência</b>	<b>Configuração da linha de transmissão</b>	<b>Tipo de sinal utilizado</b>	<b>Tamanho da janela de dados</b>	<b>Tarefa realizada</b>
Silva et al. (2006)	188 km, 230 kV, 60 Hz	tensão e corrente (fase e neutro)	$\frac{1}{4}$ de ciclo pós-falta	classificação
Ekici et al. (2008)	360 km, 380 kV	tensão e corrente	$\frac{1}{2}$ ciclo pré-falta e $\frac{1}{2}$ ciclo pós-falta	localização
Morais et al. (2010)	18,5 km, 500 kV 151,5 km, 500 kV 325 km, 500 kV	tensão e corrente	N/I	classificação
Ray et al. (2013)	300 km, 400 kV, 50 Hz	tensão e corrente	1 ciclo	localização
Swetapadma e Yadav (2015)	100 km, 400 kV, 50 Hz	tensão e corrente	1 ciclo pré-falta e 2 ciclos pós-falta	localização
Ray e Mishra (2016)	300 km, 400 kV, 50 Hz	corrente	1 ciclo pós-falta	classificação e localização
Asadi Majd et al. (2017)	200 km, 20 kV, 50 Hz	corrente	$\frac{3}{4}$ de ciclo	classificação
Swetapadma e Yadav (2017)	100 km, 400 kV, 50 Hz	tensão e corrente	1 ciclo pré-falta e 1 ciclo pós-falta	localização
Swetapadma e Yadav (2018)	100 km, 400 kV, 50 Hz	tensão e corrente	1 ciclo pré-falta e 1 ciclo pós-falta	localização
Chen et al. (2018a)	200 km, 220 kV, 50 Hz	tensão e corrente	$\frac{1}{2}$ de ciclo	classificação
Chen et al. (2018b)	100 km, 230 kV, 60 Hz 100 km, 400 kV, 60 Hz	corrente	1 ciclo pós-falta	classificação e localização
Fei et al. (2018)	300 km, 400 kV, 50 Hz	tensão	$\frac{1}{8}$ de ciclo	localização
Ye et al. (2020)	80 km, 220 kV 450 km, 500 kV	tensão e corrente	cerca de 2 ciclos	classificação
Ferreira et al. (2020)	181 km, 500 kV, 60 Hz	tensão e corrente	3 ciclos pré-falta e 3 ciclos pós-falta	classificação
Zhang et al. (2020)	IEEE-39 barras	corrente	0,1 segundo pré-falta e	localização

Continua na próxima página

Tabela 3.1: Continuação da página anterior

Referência	Configuração da linha de transmissão	Tipo de sinal utilizado	Tamanho da janela de dados	Tarefa realizada
Haq et al. (2021)	(linha não detalhada) 100 km, 60 Hz	tensão e corrente	0,1 segundo pós-falta	classificação
Coban e Tezcan (2021)	71,3 km, 154 kV, 50 Hz	corrente	N/I	classificação
Bhadra et al. (2021)	100 km, 220 kV, 50 Hz	tensão e corrente	$\frac{1}{2}$ de ciclo	classificação
Valabhoju et al. (2021)	198 km, 400 kV, 50 Hz	tensão e corrente	N/I	localização
Belagoune et al. (2021)	300 km, 400 kV, 50 Hz	tensão e corrente	1 ciclo pós-falta	classificação e localização
Fonseca et al. (2022)	300 km, 500 kV, 60 Hz	tensão	todo o registro	classificação e localização
Bhatnagar et al. (2023)	IEEE-14 barras (linha não detalhada)	tensão e corrente	$\frac{1}{2}$ de ciclo	classificação
			1 ciclo pré-falta e 1 ciclo pós-falta	localização

N/I - Não Informado

Tabela 3.2: Informações sumarizadas dos trabalhos selecionados para o estado da arte (parte 2)

Referência	Técnica de processamento de sinais	Atributos	Algoritmo	Observações
Silva et al. (2006)	nenhum	amostras dos sinais	MLP	requer a classificação de inúmeras janelas de dados, usando o voto majoritário como decisão final
Ekici et al. (2008)	WPT	energia e entropia	MLP	
Morais et al. (2010)	<i>wavelet</i>	amostras dos sinais	SVM	
Ray et al. (2013)	DWT	energia, média aritmética, desvio padrão, curtose, assimetria e entropia	MLP	
Swetapadma e Yadav (2015)	DWT	desvio padrão	MLP	utiliza o filtro <i>Butterworth</i>
Ray e Mishra (2016)	WPT	energia e entropia	SVM	requer cinco SVMs: quatro para a classificação e uma para a localização
Asadi Majd et al. (2017)	nenhum	amostras dos sinais	KNN	dividi o sinal pós-falta pelo valor RMS obtido de $\frac{1}{2}$ ciclo pré-falta
Swetapadma e Yadav (2017)	DWT	SVD	árvore de decisão	requer quatro módulos, um para cada classe de falta (LG, LL, LLG, LLL)
Swetapadma e Yadav (2018)	DFT	desvio padrão	KNN	requer quatro módulos, um para cada classe de falta (LG, LL, LLG, LLL)
Chen et al. (2018a)	nenhum	amostras dos sinais	CSAE+ <i>softmax</i>	
Chen et al. (2018b)	nenhum	amostras dos sinais	SG-ELM	requer dados de ambos os terminais, e o mesmo modelo realiza ambas as tarefas ( <i>i.e.</i> , classificação e localização)
Fei et al. (2018)	nenhum	amostras dos sinais	SVM	utiliza um filtro passa-baixa
Ye et al. (2020)	nenhum	amostras dos sinais	CNN	
Ferreira et al. (2020)	nenhum	RMS, máximo, mínimo,	MLP	utiliza um modelo para cada tarefa

Continua na próxima página

Tabela 3.2: Continuação da página anterior

Referência	Técnica de processamento de sinais	Atributos	Algoritmo	Observações
Zhang et al. (2020) Haq et al. (2021)	nenhum DWT	média aritmética, energia e integral numérica amostras dos sinais energia e desvio padrão	Bi-GRU ELM	requer dados de ambos os terminais utiliza duas ELMs, uma para detectar o envolvimento das fases e outra para a terra
Coban e Tezcan (2021) Bhadra et al. (2021) Valabhoju et al. (2021)	nenhum <i>wavelet</i> DWT	amostras dos sinais imagem desvio padrão	SVM CNN <i>Bagging</i> de árvores de decisão	converte os sinais em imagem utiliza o filtro <i>Butterworth</i> e PCA
Belagoune et al. (2021) Fonseca et al. (2022) Bhatnagar et al. (2023)	nenhum nenhum nenhum	amostras dos sinais amostras dos sinais sinal TKEO	LSTM RF RF	utiliza um modelo para cada tarefa utiliza o filtro de <i>notch</i> requer quatro módulos, um para cada classe de falta (LG, LL, LLG, LLL)

### 3.2 ANÁLISE CRÍTICA

A descrição dos trabalhos na seção anterior permite uma análise ampla e, ao mesmo tempo, crítica das abordagens já dispostas no estado da arte. Por conseguinte, podemos determinar tanto propriedades já solucionadas pela literatura especializada quanto oportunidades ainda em aberto para contribuições de novos métodos.

Um dos principais problemas identificados é a indisponibilidade de conjuntos de dados públicos de falhas em linhas de transmissão, condição crucial para os trabalhos que empregam algoritmos de AM em sua solução. Essa escassez está atrelada a diversos fatores, como competitividade, custo e confidencialidade dos dados, fazendo com que cada pesquisador gere seus próprios dados. Consequentemente, as avaliações dos métodos para diferentes conjuntos de dados impede que seus desempenhos sejam comparados diretamente, exigindo que cada deles seja reproduzido na íntegra para uma comparação justa e correta. Uma exceção é a base de dados UFPAFaults, desenvolvida e disponibilizada por Morais et al. (2010). Todavia, esse conjunto possui apenas dados pós-falta e processamento prévio nas oscilografias, além de restringir sua aplicabilidade apenas para a tarefa de classificação do tipo da falta. Portanto, métodos que realizam a detecção e a localização de faltas, ou qualquer outra atividade vinculada a análise de oscilografias de falhas, não podem usufruir plenamente de seus exemplos.

Considerando os aspectos que já estão estabelecidos e respaldados por uma variedade de experimentos em estudos relacionados, destaca-se a utilização do sinal de corrente como a principal alternativa para o diagnóstico de falhas (*i.e.*, classificação e localização). Entre os trabalhos que avaliam o uso da tensão e da corrente, individualmente ou em conjunto, estão Chen et al. (2018a), Ye et al. (2020) e Haq et al. (2021), os quais relataram que os piores resultados estão vinculados quando seus métodos empregam apenas o sinal da tensão. A justificativa para este comportamento no desempenho está amparada pela discussão presente na Subseção 2.1.1, cuja bibliografia já fundamenta que o sinal da corrente é a forma de onda mais significativa para as faltas por curto-circuito. Tal propriedade pode, inclusive, explicar em parte o desempenho relativamente menor do trabalho de Fonseca et al. (2022) em relação às demais obras para a classificação de falhas, uma vez que os autores utilizam apenas o sinal da tensão em sua proposta.

Em resumo, o uso apenas do sinal da corrente demonstra ser suficiente para a realização da classificação do tipo de falta, desde que uma engenharia de atributos apropriada e um algoritmo de aprendizado de máquina adequado sejam utilizados. Por sua vez, as informações extraídas dos trabalhos relacionados para a localização de falhas expressam um melhor caminho para o uso dos sinais da tensão e da corrente em conjunto, reverberando, normalmente, em resultados mais precisos e acurados.

No geral, grande parte dos autores reportam altas taxas de acerto para os seus respectivos métodos, tanto para a classificação quanto para a localização a partir do uso de algoritmos de AM. Optamos por não reportar os desempenhos das obras na seção anterior devido ao fato de que tais desempenhos foram obtidos a partir de diferentes conjuntos de dados, de modo que uma análise comparativa direta entre eles poderia incutir em um viés ao leitor. De qualquer maneira, podemos identificar algumas limitações que são capazes de dificultar ou até mesmo impedir a aplicação de tais métodos em cenários reais.

Diversos autores utilizam técnicas de processamento de sinais, como aquelas baseadas em transformadas *wavelet* (DWT e WPT) e Fourier (DFT), para decompor os sinais originais de corrente/tensão e, assim, extrair coeficientes para serem utilizados diretamente como características ou manuseados para a extração de atributos (Ferreira et al., 2016; Chen et al., 2018a). Porém, essas técnicas podem demandar de uma quantidade considerável de amostras, geralmente exigindo uma frequência de amostragem maior, inviabilizando sua aplicação em alguns sistemas

de proteção reais que eventualmente possuem restrições na disponibilidade de dados, por exemplo, taxas de amostragem mais baixas (Gafoor e Rao, 2006; Silva et al., 2007; Godse e Bhat, 2020; Haq et al., 2021).

Em outras palavras, tais abordagens exigem instrumentação mínima do sistema de proteção para a coleta de dados, atrelando seu desempenho e até mesmo seu funcionamento a determinados valores de frequências (Ray et al., 2013; Ray e Mishra, 2016; Haq et al., 2021). Outro aspecto é o fato de que as técnicas baseadas em *wavelets* aumentam progressivamente a complexidade, especialmente para níveis crescentes de decomposição do sinal, bem como a determinação da *wavelet* mãe (Mukherjee et al., 2021), resultando numa maior complexidade computacional e na falta de generalização do modelo para outros sistemas de transmissão (Chen et al., 2018a).

Além disso, estes trabalhos não avaliam seus métodos com dados de linhas de transmissão diferentes daquelas utilizadas para treinar seus métodos mesmo para a tarefa de classificação que, em tese, é menos complexa que a localização. Ou seja, não avaliam a competência para generalização. Num cenário real, tais métodos necessitarão de treinamento individualizado para cada linha de transmissão supervisionada pelo sistema de proteção, o que poderia desencorajar seu uso, uma vez que o sistema de proteção pode supervisionar diversas linhas simultaneamente com características distintas, por exemplo, comprimento e nível de tensão.

Uma ressalva ocorre para o método proposto por Silva e colaboradores, publicado em duas oportunidades (Silva et al., 2006, 2007). Segundo os autores, sua etapa da normalização permite que o método seja independente de LT, ou seja, permite que o método, treinado para uma determinada linha, possa generalizar para outras LTs com diferentes níveis de tensão. Inclusive, os autores relatam a análise do método em eventos reais de falha coletados da rede da Companhia Hidro Elétrica do São Francisco (Chesf), cujo treinamento do modelo ocorreu a partir dos dados simulados de uma única LT.

Todavia, o método possui ao menos duas desvantagens. A primeira é a etapa de janelamento que exige que inúmeras janelas sejam classificadas, de modo que o tipo com mais predições torna-se a decisão final (voto majoritário). Já a segunda e principal desvantagem, também associada ao uso do janelamento, está relacionada ao fato de que o método requer a determinação prévia do instante final da falha na oscilografia, visto que registros reais de falha não contêm apenas sinais pré e pós-falha, mas também dados após a atuação do sistema de proteção. Logo, esses sinais não mais refletem o comportamento de uma falta e, portanto, não devem ser considerados para a classificação. Adicionalmente, a quantidade de ciclos representativos de falha em eventos reais não é uniforme, a qual pode variar dependendo do tempo de atuação do sistema de proteção perante cada falta. Por exemplo, considerando a rede da COPEL, há registros de falhas com apenas três ciclos após a sua incidência, enquanto outros registros contêm uma quantidade maior de ciclos.

Um último ponto digno de ser destacado refere-se ao fato de que o uso de algoritmos *ensemble* é uma metodologia praticamente inexplorada no diagnóstico de falhas, especialmente para a tarefa de localização de faltas, demonstrando uma lacuna que deve ser considerada com atenção. Esta motivação está relacionada com os argumentos descritos na Subseção 2.2.2, em que a combinação de diferentes regressores pode resultar em melhores desempenhos. Neste contexto, podemos aplicar a seleção dinâmica para obter os regressores mais competentes do conjunto para melhorar o sistema de predição. Até onde sabemos, este é o primeiro trabalho que aplica a estratégia SDR para localização de faltas.



### 3.3 CONSIDERAÇÕES FINAIS

Apesar da literatura científica abranger uma variedade de métodos para as atividades de classificação e localização de falhas, ainda há contribuições a serem exploradas em ambos os tópicos. Neste quesito, destaca-se a necessidade do desenvolvimento de métodos flexíveis perante diferentes taxas de amostragem. Além do mais, os trabalhos encontrados não avaliam a capacidade de seus métodos generalizarem a execução da classificação e da localização para linhas de transmissão distintas daquela empregada para o treinamento dos algoritmos (*domain generalization*). Apesar de alguns autores avaliarem suas abordagens para mais de uma LT, os respectivos algoritmos foram treinados com dados dessas linhas, ou seja, não analisaram a capacidade de generalização para linhas desconhecidas (não utilizadas durante o treinamento).

Outro aspecto crucial se deve ao fato de que tais métodos que empregam algoritmos de AM demandam de dados para o treinamento de seus modelos, repercutindo, assim, na necessidade de um conjunto representativo de eventos de faltas em LTs. A única base de dados de faltas encontrada disponível publicamente (UFPAFaults proposta por Moraes et al. (2010)) é limitada, uma vez que apenas a tarefa de classificação pode ser realizada a partir dela, além de outros fatores desfavoráveis que foram argumentados na seção anterior.

Em resposta a essas limitações, apresentamos neste trabalho novos conjuntos de dados de faltas em linhas de transmissão, disponibilizando tanto as séries temporais de tensão e corrente quanto as características extraídas das oscilografias, sem a utilização de qualquer técnica de processamento de sinais. Além disso, propomos um método para a classificação de falhas com flexibilidade de taxa de amostragem e capacidade de generalização, visando superar os problemas técnicos relatados no parágrafo anterior. Em complemento, propomos um novo *framework* SDR que visa selecionar os modelos mais precisos de um conjunto para prever a localização da falta. Tanto o método de classificação quanto o de localização visam explorar o uso de algoritmos *ensemble*, conforme motivação anteriormente relatada.

Nos próximos capítulos descrevemos as bases de dados criadas e disponibilizadas publicamente no âmbito deste projeto, assim como o método proposto para as tarefas de classificação e localização.



## 4 BASES DE DADOS PARA ANÁLISE DE FALHAS

Apresentamos neste capítulo os novos conjuntos de dados elaborados neste trabalho, constituídos por diversos exemplos de faltas em linhas de transmissão. Primeiramente, descrevemos algumas informações sobre o ambiente considerado para gerar os eventos de falha, além dos parâmetros utilizados para simulá-los (Seção 4.1). Em especial, geramos dados para duas linhas de transmissão distintas, com níveis de tensão de 500 kV (Subseção 4.1.1) e 230 kV (Subseção 4.1.2). Posteriormente, relatamos o processo de extração de atributos a partir destes eventos (Seção 4.2).

O resultado destas ações foi a criação de duas novas bases de dados: *Fault Analysis Database* (FADb) e *Fault Analysis Database with Features* (FADbF). O primeiro conjunto compreende os dados de séries temporais de corrente e tensão de ambas as linhas (500 kV e 230 kV). Já o segundo conjunto representa uma coleção de atributos extraídos a partir das séries temporais da primeira base. Essa nossa contribuição permite sanar a lacuna identificada na literatura referente a escassez de bases públicas para a análise de falhas.

### 4.1 GERAÇÃO DOS EVENTOS DE FALTA

Nesta seção estão descritos os processos para a geração dos casos sintéticos de falhas em LTs. No âmbito deste trabalho, geramos eventos simulados de faltas para duas linhas de transmissão distintas, cada qual com suas próprias particularidades, como extensão e nível de tensão. Na Subseção 4.1.1 introduzimos o processo de geração de falhas para a linha de 500 kV, enquanto na Subseção 4.1.2 apresentamos os eventos gerados para a LT de 230 kV. Estes eventos de ambas as linhas constituem uma primeira base de dados que denominamos de FADb, cujos dados representam séries temporais de corrente e tensão.

As simulações basearam-se no modelo IEEE de 9-barras disponibilizado pela ferramenta ATPDraw v7.2 (Høidalen et al., 2019). Este modelo representa uma aproximação de um sistema elétrico de potência real (Kumar e Bhagat, 2019), sendo composto por nove barras, três geradores e seis linhas de transmissão. Na Figura 4.1 está representado o modelo considerado no estudo.

O ATPDraw<sup>1</sup> é um pré-processador gráfico para ATP<sup>2</sup> (*Alternative Transients Program*), usado para criar e editar arquivos de circuitos de maneira interativa. A saída do ATPDraw é um arquivo que é fornecido como entrada para o programa ATP, simulador consagrado na indústria de energia elétrica para simulações digitais de fenômenos transitórios de natureza eletromagnética e eletromecânica (Haginomori et al., 2016). Ambas as ferramentas são amplamente utilizadas na comunidade científica para estudos de circuitos elétricos, particularmente em sistemas de potência, considerando a investigação sobre análise de falhas e seus efeitos (Ye e Campbell, 2006; Resende e Silveira, 2016; Lopes et al., 2023).

---

<sup>1</sup><https://www.atpdraw.net/>

<sup>2</sup><https://www.atp-empt.org/>

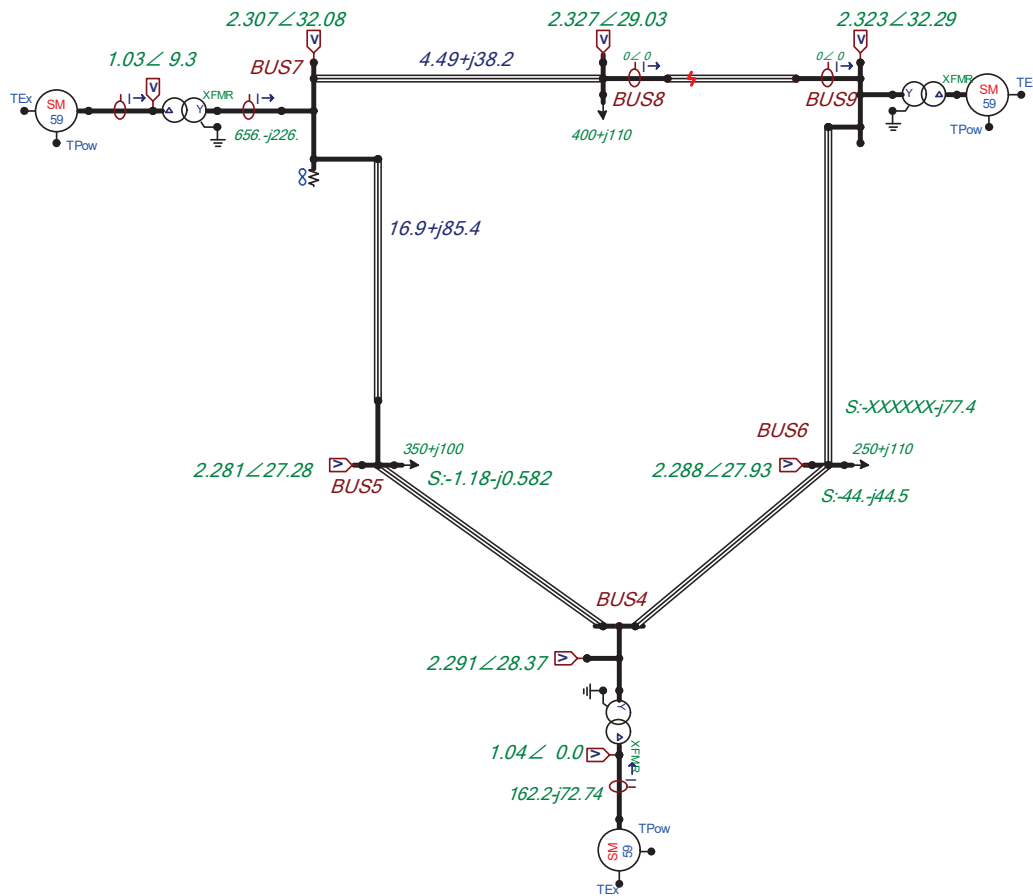


Figura 4.1: Sistema IEEE 9-barras utilizado para a geração dos dados de falta

Fonte: adaptado de Høidalen et al. (2019)

#### 4.1.1 Linha de Transmissão de 500 kV

Em particular, realizamos alterações pontuais no modelo original da IEEE de 9-barras para representar, em parte, o sistema de transmissão desejado, com uma tensão maior e uma extensão mais longa para uma das linhas de transmissão. O modelo utilizado para a realização das simulações e os respectivos arquivos de falta estão disponíveis publicamente em nosso repositório<sup>3</sup>. Cabe ressaltar que tal modelo foi validado junto a engenheiros eletricitas e pesquisadores do Instituto de Tecnologia para o Desenvolvimento<sup>4</sup> (Lactec).

As simulações das faltas ocorreram para uma única linha de transmissão de 414 km, 500 kV e 60 Hz, entre as barras 8 (BUS8) e 9 (BUS9) da Figura 4.1. Essa especificação visa representar a linha de transmissão mais longa da rede da Companhia Paranaense de Energia (COPEL), ligando as subestações de Itatiba (São Paulo) e Bateias (Paraná).

Ainda, cada ensaio iniciou sem a presença de falha na linha, sendo que o início da perturbação ocorre em diferentes instantes de tempo. Os dados coletados durante os ensaios compreenderam os sinais de corrente e tensão para cada uma das três fases (A, B e C) de ambas as extremidades (terminais) da linha de transmissão (BUS8 e BUS9). Cada ensaio compreendeu a duração de meio segundo (0,5 s) com taxa de amostragem de 10 kHz ( $\Delta t = 1 \times 10^{-4}$ ). Assim sendo, cada ciclo de tensão/corrente abrange cerca de 167 amostras (10 kHz dividido por 60 Hz).

<sup>3</sup><https://1drv.ms/u/s!ArMEeMx4MYDNimHVxiDx3b4CI3iL?e=8GfXg7>

<sup>4</sup><https://lactec.com.br/>

Os parâmetros de falta empregados nas simulações foram:

- **Tipo de falta:**
  - Monofásica: AG, BG e CG;
  - Bifásica: AB, AC e BC;
  - Bifásica-terra: ABG, ACG e BCG;
  - Trifásica: ABC/ABCG.
- **Localização de falta:** 1 a 100% da extensão da linha, com intervalos de 1%;
- **Resistência de falta:** 0,01 a 200  $\Omega$ , com intervalos de 10  $\Omega$ ;
- **Tempo de início de falta:** 0,091 s, 0,093 s, 0,095 s, 0,097 s, 0,099 s, 0,101 s, 0,103 s e 0,105 s.

O parâmetro tempo de início de falta representa a mesma função do parâmetro do ângulo de início da falta, usualmente utilizado em trabalhos relacionados. A importância dessa propriedade se deve a representação da incerteza associada com o instante da ocorrência da falta (Ferreira et al., 2016). Cada tempo de início de falta está inserido no mesmo ciclo, porém em posições distintas deste ciclo.

Vale reforçar que as letras A, B e C representam cada uma das três fases da linha de transmissão, enquanto a letra G corresponde a atuação da terra na falta. A combinação das iniciais indica as respectivas fases que estão envolvidas na falha. Por exemplo, AG representa uma falta envolvendo a fase A e a terra, assim como AB representa a falta entre as fases A e B sem o envolvimento do solo. Mais detalhes podem ser encontrados nas seguintes referências (Yadav e Dash, 2014; Aleem et al., 2015; Grainger et al., 2016).

Em complemento, as falhas trifásicas ABC e ABCG foram agrupadas como um mesmo tipo devido ao fato de apresentarem o mesmo comportamento para as formas de onda trifásicas para as faltas simuladas. Isto se justifica em razão de não haver componentes de sequência-negativa nem de sequência-zero para curtos-circuitos trifásicos equilibrados, conforme Teorema de Fortescue (Fortescue, 1918). Portanto, não há diferença no curto-circuito trifásico (ABC) e trifásico com terra (ABCG) (Fonseca et al., 2022).

Por outro lado, um curto-circuito bifásico à terra terá componentes de sequência-positiva, negativa e zero, o que gera uma impedância resultante diferente de um curto-circuito bifásico que terá apenas componentes de sequência-positiva e negativa. Consequentemente, resultará em comportamentos distintos entre os tipos de falta, mesmo que apenas o parâmetro do tipo de falta seja alterado. Discutimos anteriormente sobre esses componentes na Subseção 2.1.1.

Logo, a geração dos dados compreendeu a combinação entre cada um dos parâmetros previamente citados, resultando em um total de 168.000 casos de falta. Desse modo, garante-se que cada tipo de falta tenha o mesmo número de exemplos, além das mesmas combinações de valores de parâmetros de resistência, localização e instante da ocorrência de falta (conjunto balanceado). A partir dos parâmetros previamente descritos e do modelo disponibilizado, é possível reproduzir os eventos da base de dados, assim como a criação de novas simulações de falhas considerando outras linhas ou outros valores para os parâmetros. Neste caso, por exemplo, poderíamos gerar novas falhas com resistências superiores a 200  $\Omega$ .

Em nosso repositório, os dados da base FADb estão representados por séries temporais. Define-se série temporal como uma sequência de observações ao longo do tempo à respeito de um determinado fenômeno (Esling e Agon, 2012). Formalmente, uma série temporal  $T$  corresponde a uma sequência de amostras  $T = (T_0, T_1, T_2, \dots, T_{n-1})$  ordenada cronologicamente, com comprimento  $|T| = n$ , em que  $T_i$  é o valor da amostra na posição  $i$  ( $0 \leq i < n$ ) em  $T$  (Brillinger, 2001; Schneider et al., 2021). Na FADb, os fenômenos que são representados por séries temporais equivalem aos sinais de tensão e corrente, sendo cada grandeza considerada individualmente para cada terminal da LT.

Como resultado, o conjunto de dados FADb possibilita futuros *benchmarks*<sup>5</sup> e avaliações para diversas aplicações de engenharia elétrica e também da computação. Entre as tarefas passíveis de serem realizadas está a detecção de anomalias em séries temporais (detecção de falhas), além de problemas do aprendizado de máquina para a classificação (classificação de tipo de falta) e a regressão (localização de falhas).

#### 4.1.2 Linha de Transmissão de 230 kV

Os procedimentos adotados para a geração dos eventos para a LT de 230 kV são similares aos descritos na subseção anterior. Entretanto, estes eventos foram fundamentados a partir do modelo original IEEE de 9-barras fornecido pela ferramenta ATPDraw (Høidalen et al., 2019). Conseqüentemente, a configuração dos componentes do sistema de potência para essa linha de transmissão difere daquela apresentada na Subseção 4.1.1.

Para este cenário, utilizamos uma LT de 230 kV, 106,646 km e 60 Hz, empregando uma taxa de amostragem de 100 kHz. Adicionalmente, cada simulação teve duração de 1 s. Inclusive, seu modelo no ATPDraw também está disponível em nosso repositório (mesmo endereço eletrônico informado na subseção anterior).

Os parâmetros de falta considerados foram:

- **Tipo de falta:**
  - Monofásica: AG, BG e CG;
  - Bifásica: AB, AC e BC;
  - Bifásica-terra: ABG, ACG e BCG;
  - Trifásica: ABC/ABCG.
- **Localização de falta:** 1 a 100% da extensão da linha, com intervalos de 1%;
- **Resistência de falta:** 0,01 a 200  $\Omega$ , com intervalos de 5  $\Omega$ ;
- **Tempo de início de falta:** 0,1 s.

A combinação de cada um destes parâmetros permitiu a criação de 41.000 novos eventos de falha. Apesar dos dados gerados para esta linha de 230 kV não estarem atualmente em nenhum repositório devido ao tamanho necessária para o seu armazenamento (aproximadamente 469 GB), esta parte do conjunto FADb poderá ser compartilhada sob requisição.

<sup>5</sup>Em síntese, um *benchmark* é um ponto de referência para que métodos possam ser comparados ou avaliados. Serve como uma linha de base pelo qual o desempenho, a qualidade ou a eficiência de algo pode ser mensurado.

## 4.2 EXTRAÇÃO DE ATRIBUTOS

O conjunto de dados FADbF, originado a partir das simulações descritas na Seção 4.1, amplia os atributos comumente usados na literatura para a análise de falhas a partir de estratégias que empregam o AM (média aritmética, desvio padrão, energia, entropia, RMS, máximo e mínimo) (Ray et al., 2013; Chen et al., 2016; Mishra e Ray, 2018). Vale ressaltar que exploramos diversos mapeamentos, revisões e artigos relacionados, como os trabalhos descritos e referenciados na Seção 3.1, visando identificar as principais características utilizadas na literatura para compor nosso conjunto de dados. Além dessas características, adicionamos atributos inéditos no conjunto FADbF referente ao diagnóstico de falta, como média geométrica, distância e covariância, entre outros que listaremos a seguir.

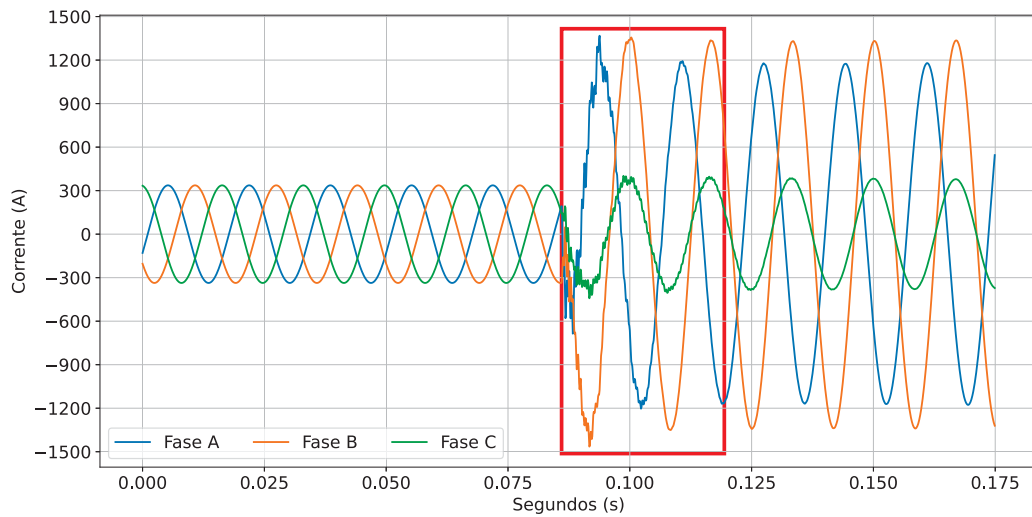
No total, extraímos 21 características separadamente para cada uma das três fases das formas de onda de tensão e corrente ao longo de dois ciclos pós-falta de um único terminal, resultando em 126 atributos ( $21 \times 3 \times 2 = 126$ ) além dos dois alvos possíveis, ou seja, tipo de falha (tarefa de classificação) e localização da falha (tarefa de regressão). A lista de atributos inclui: tipo de falha, localização da falha dada em quilômetros (km), energia, distância, RMS, curtose, assimetria, frequência máxima, largura de banda espectral, área sob a curva (AUC), distância pico a pico, inclinação, entropia de Shannon, média aritmética, média geométrica, média harmônica, variância, desvio padrão, mediana, covariância entre as fases, correlação entre as fases, máximo e mínimo. A extração de características das séries temporais de corrente e tensão foi realizada com o suporte das seguintes bibliotecas em Python: TSFEL v0.1.4 (Barandas et al., 2020), PyInform v0.2.0 (Moore, 2019) e statistics v3.11.4 (Python, 2023).

A escolha do comprimento da janela com tamanho equivalente a dois ciclos pós-falta para a extração de atributos, conforme ilustrada na Figura 4.2, fundamenta-se pelo estado da arte que expressa que dois ciclos fornecem informações suficientes para o diagnóstico de faltas (Mishra e Ray, 2018; Mukherjee et al., 2021; Shakiba et al., 2022). O uso de mais de dois ciclos pode implicar na indisponibilidade de dados, uma vez que o sistema de proteção atua o mais rápido possível para identificar o início de uma falha e, então, isolá-la do restante do sistema.

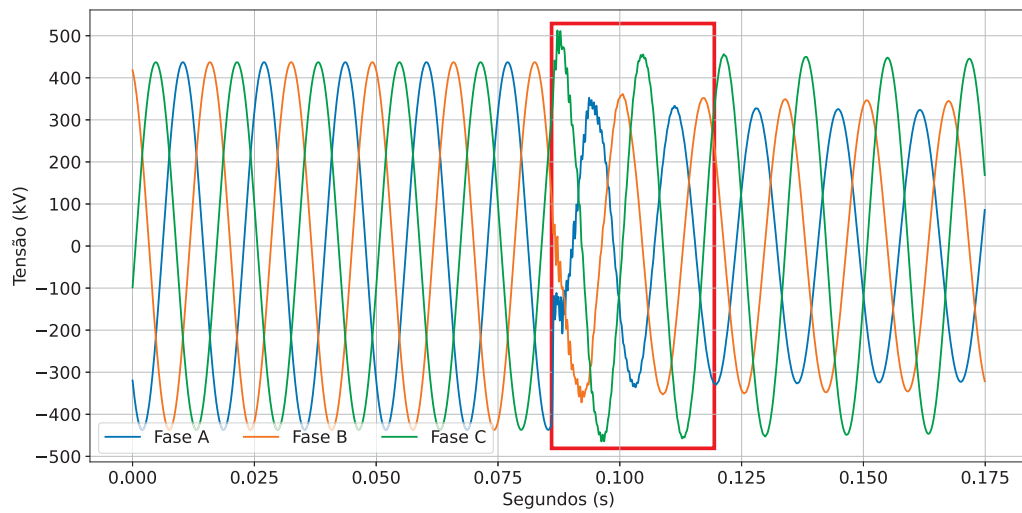
É digno de nota que disponibilizamos publicamente o conjunto FADbF<sup>6</sup>, junto a uma lista dos atributos extraídos e todos os seus detalhes (Ensina et al., 2023). De modo semelhante ao relatado para a base FADb na Subseção 4.1.1, o FADbF também possibilita futuros *benchmarks*, especialmente para a subárea do AM ao considerar as tarefas de classificação (classificação de tipo de falha) e regressão (localização de falha).

A Tabela 4.1 resume os atributos comumente extraídos e usados pelos trabalhos relacionados. Vale ressaltar que nenhum destes conjuntos de dados está disponível para a comunidade científica, exceto a nossa base de dados. Não incluímos a base UFPAFaults na tabela, apesar de publicamente acessível, devido aos autores não extraírem, de fato, atributos dos sinais de corrente e tensão. Observa-se que parte destes trabalhos utilizou técnicas de processamento de sinais como etapa prévia a extração de atributos. Consequentemente, eles são afetados pelos mesmos problemas de processamento de sinais discutidos anteriormente na Seção 3.2. Por outro lado, o nosso conjunto de dados não depende de nenhuma destas técnicas, apesar de que nada impede que trabalhos futuros executem alguma técnica de processamento de sinais previamente a extração de atributos dos sinais de tensão e corrente da base FADb para a geração de um novo conjunto de dados. Fica evidente também que nossas características são mais diversificadas e interessantes de serem exploradas.

<sup>6</sup><https://dx.doi.org/10.5281/zenodo.10275032>



(a) Sinal da corrente



(b) Sinal da tensão

Figura 4.2: Exemplo de dois ciclos pós-falta demarcados em uma simulação de falta do tipo ABG para os sinais de (a) corrente e (b) tensão

Fonte: autoria própria

Tabela 4.1: Resumo dos atributos normalmente criados e usados para análise de falhas por trabalhos relacionados (em negrito estão os atributos propostos em nosso trabalho)

Referência	Atributos	Processamento de sinal
Ray et al. (2013)	energia, desvio padrão, média aritmética, curtose, assimetria e entropia	DWT
Ekici et al. (2008)	energia e entropia	WPT
Swetapadma e Yadav (2015)	desvio padrão	DWT
Ray e Mishra (2016)	energia e entropia	WPT
Swetapadma e Yadav (2018)	desvio padrão	DFT
Ferreira et al. (2020)	RMS, máximo, mínimo, média aritmética, energia e integral	nenhum
Haq et al. (2021)	energia e desvio padrão	DWT
Valabhoju et al. (2021)	desvio padrão	DWT
Nosso trabalho	energia, <b>distância</b> , RMS, curtose, assimetria, <b>largura de banda espectral</b> , AUC, <b>declive frequência máxima</b> , <b>distância pico a pico</b> , entropia, média aritmética, <b>média harmônica</b> , <b>média geométrica</b> , <b>variância</b> , desvio padrão, <b>mediana</b> , <b>covariância entre as fases</b> , <b>correlação entre as fases</b> , máximo e mínimo	nenhum

### 4.3 CONSIDERAÇÕES FINAIS

Apresentamos ao longo deste capítulo o processo de geração de eventos simulados de falhas em LTs (Seção 4.1). Resumidamente, propomos uma nova base de dados que denominamos de FADb, composta por casos diversificados de faltas de duas linhas de transmissão: 500 kV (Subseção 4.1.1) e 230 kV (Subseção 4.1.2). Além disso, relatamos os procedimentos realizados para a extração de atributos destas bases (Seção 4.2), gerando, assim, um novo conjunto de dados: FADbF. Ambas as bases que foram geradas neste trabalho estão publicamente disponíveis em nossos repositórios, salvo os dados para a linha de 230 kV que podemos fornecer sob requisição.

No próximo capítulo descrevemos o método proposto para a análise de falhas em linhas de transmissão, em particular para as tarefas de classificação e localização de faltas. Os experimentos de avaliação do método utilizam os dados gerados e relatados neste capítulo.



## 5 MÉTODO PROPOSTO

O presente trabalho visa o desenvolvimento de um método capaz de classificar e localizar falhas independentemente da frequência de amostragem, uma vez que os registradores digitais de perturbação podem apresentar variadas taxas de amostragem. Objetivamos ainda, quando possível, a capacidade de generalização do método para analisar linhas de diferentes configurações, mesmo que os dados destas não tenham sido utilizados para treinar o algoritmo.

Na Figura 5.1 está ilustrado o ambiente operacional do método proposto para o diagnóstico de falhas, juntamente com a demarcação de três cenários correlacionados. Uma vez que uma perturbação representativa de uma falta, originada de uma descarga atmosférica, por exemplo, seja identificada pelo sistema de proteção do SEP (Figura 5.1 (1)), equipamentos como os descritos na Subseção 2.1.2 irão armazenar os respectivos registros de oscilografia em centrais de análise (Figura 5.1 (2)). Logo, estes registros podem ser fornecidos para os componentes de análise de nosso método, em especial os componentes de classificação e localização de falhas (Figura 5.1 (3)).

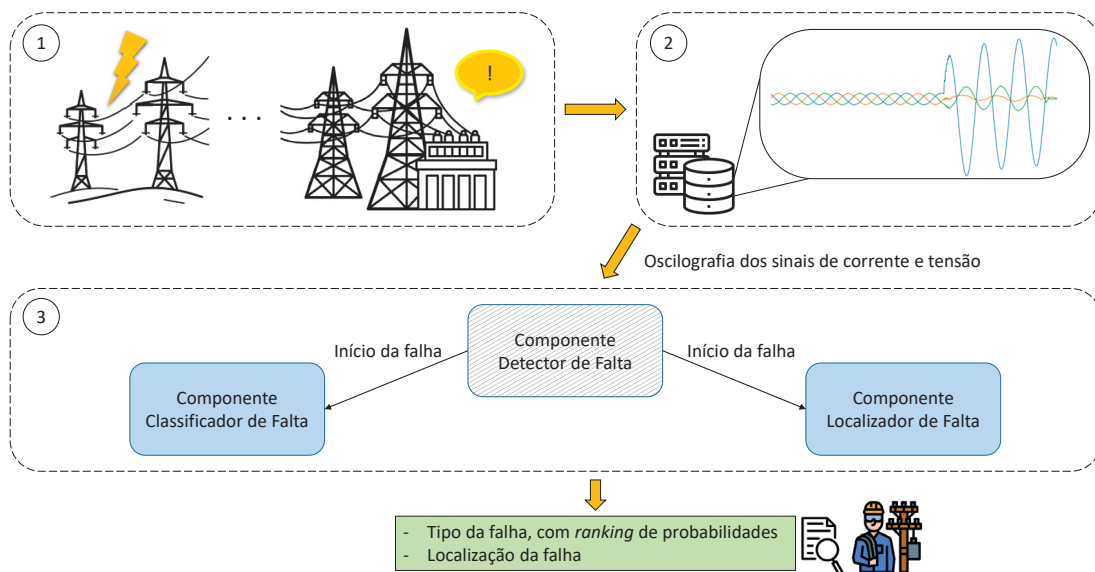


Figura 5.1: Ambiente de operação do método proposto para o diagnóstico de falhas em linhas de transmissão

Fonte: autoria própria

Cabe evidenciar que, neste trabalho, assumimos que a tarefa de detecção de falta (Componente Detector de Falta) é realizada por um método terceiro, visto que essa tarefa já está bem coberta e resolvida pela literatura, incluindo a flexibilidade para distintas taxas de amostragem e a generalização. Diversas técnicas estão dispostas no estado da arte para esta finalidade, incluindo métodos baseados em AM e estatísticos por meio de janelas deslizantes (Gilbert e Morrison, 1997; Raza et al., 2020). Inclusive, entre estas técnicas está uma nova abordagem não supervisionada que propomos para a detecção de falta em um trabalho relacionado no âmbito deste projeto (Tamalu et al., 2023).

Assim sendo, este capítulo inicia com a descrição dos componentes propostos para a classificação (Seção 5.1) e para a localização de falhas (Seção 5.2), sendo, ao final, destacadas as ferramentas utilizadas para o desenvolvimento deste trabalho (Seção 5.3).

## 5.1 COMPONENTE CLASSIFICADOR DE FALTA

Nesta seção descrevemos o método proposto para a classificação do tipo de falha, cujo fluxograma está retratado na Figura 5.2. Observe que a saída do módulo de classificação disponibiliza não apenas o tipo de falha previsto pelo algoritmo, mas também um *ranking* de probabilidades de cada classe corresponder aos dados de entrada. Tal funcionalidade é fornecida pelo próprio algoritmo utilizado para a classificação, o qual será mencionado mais adiante.

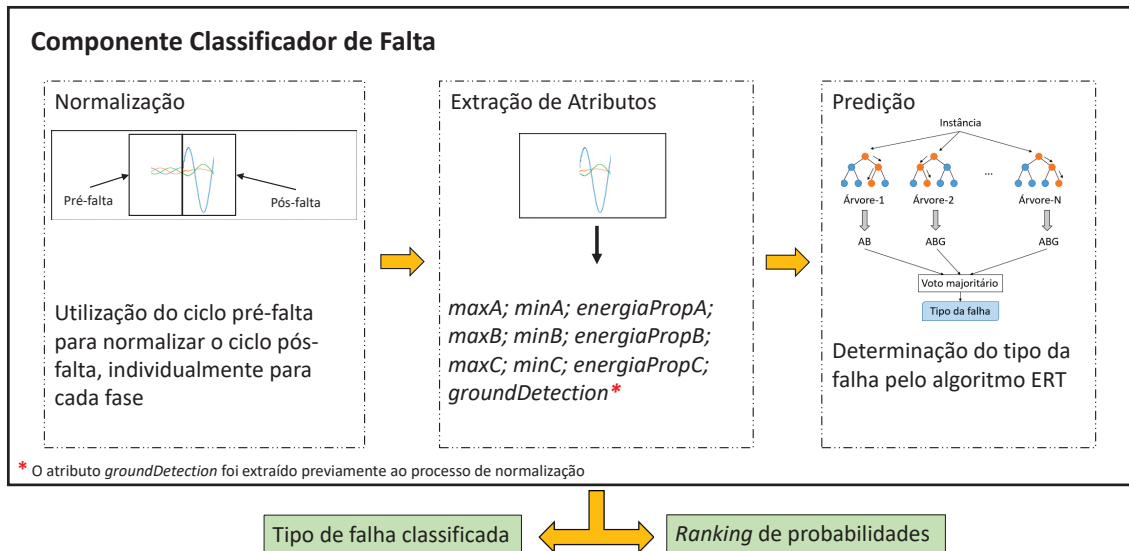


Figura 5.2: Fluxograma do componente proposto para a classificação de falha

Fonte: autoria própria

A seção está organizada da seguinte maneira. Inicialmente apresentamos o processo para extração de atributos (Subseção 5.1.1), seguido pela descrição do algoritmo classificador usado em nosso estudo (Subseção 5.1.2). Cabe mencionar que os atributos extraídos e utilizados neste componente diferem das características listadas para a base FADbF na Seção 4.2, as quais são empregadas no componente de localização.

### 5.1.1 Normalização e Extração de Atributos

O processo de extração de atributos ocorre a partir de uma sequência de cinco etapas. A primeira etapa compreende a obtenção de um ciclo pré-falta e um ciclo pós-falta do sinal da corrente para cada uma das três fases a partir do arquivo de oscilografia dado como entrada. Cabe mencionar que as entradas deste componente correspondem tanto ao arquivo de oscilografia quanto a informação do momento de início da falha fornecido pelo Componente Detector de Falta. Essa informação é importante para permitir que o presente componente possa segmentar os sinais nos ciclos previamente mencionados.

Na segunda etapa, o ciclo trifásico pós-falta é inicialmente utilizado para a obtenção do atributo que denominamos de *groundDetection*. Em particular, esse atributo binário tem o objetivo de auxiliar na identificação da ocorrência ou não da terra na falha. Assim sendo, calculamos um valor de índice para cada amostra instantânea do ciclo pós-falta da corrente, conforme definida pela Equação 5.1 e proposta originalmente por Akke e Thorp (1998). Em particular, a Equação 5.1 baseia-se no componente da corrente de sequência-zero, o qual é um indicador da presença da terra na falha (vide Subseção 2.1.1).

$$Index_i = \frac{|I_{A_i} + I_{B_i} + I_{C_i}|}{média(|I_{A_i}|, |I_{B_i}|, |I_{C_i}|)} \quad (5.1)$$

sendo  $Index_i$  o valor do índice,  $I_{A_i}$ ,  $I_{B_i}$  e  $I_{C_i}$  os sinais da corrente para as fases A, B e C, respectivamente, para o instante  $i$ . A média desses índices foi então comparada ao limiar de 0,1, o qual foi definido experimentalmente neste trabalho com base em um subconjunto de faltas reais coletadas da rede da COPEL. Desse modo, o valor do atributo binário *groundDetection* foi:

- **0**: caso a média fosse menor ou igual à 0,1. Indica que a terra não teve participação na falta;
- **1**: caso a média fosse maior que 0,1. Indica que a terra teve participação na falta.

Já na terceira etapa, os valores de máximo (max) e mínimo (min) foram obtidos do ciclo pré-falta, separadamente para cada fase, para normalizar o ciclo pós-falta pela técnica MinMax da respectiva fase, ilustrada pela Equação 5.2. Vale ressaltar que cada exemplo (instância) foi normalizado individualmente pelos seus próprios dados.

$$\hat{X}_i = \frac{X_i - \min(X_{pré-falta})}{\max(X_{pré-falta}) - \min(X_{pré-falta})} \quad (5.2)$$

sendo  $\hat{X}_i$  o dado normalizado de  $X_i$  no ciclo pós-falta para o instante  $i$  para uma determinada fase (A, B ou C), enquanto  $X_{pré-falta}$  representa um ciclo pré-falta para a respectiva fase de  $X_i$ , ciclo o qual utilizamos para coletar os valores de máximo e mínimo para usarmos nesta equação. Observe que esse valores de máximo e mínimo do ciclo pré-falta foram empregados apenas nesta equação para normalizar o ciclo pós-falta e não compõem o vetor de atributos.

Na quarta etapa, calculamos os valores da energia para cada fase do ciclo pós-falta normalizado da corrente, conforme demonstrada pela Equação 5.3.

$$E = \sum_{i=1}^N \hat{X}_i^2 \quad (5.3)$$

onde  $E$  é o valor de energia do ciclo pós-falta normalizado do sinal de corrente  $\hat{X}$  para uma fase em particular. Em seguida, calculamos a energia total (*energiaTotal*) que compreende a soma das energias das três fases, para então gerarmos os atributos que representam a proporção da contribuição da energia (*energiaProp*) de cada fase para a energia total (*energiaTotal*), conforme exemplificada pela Equação 5.4.

$$energiaProp = \frac{E}{energiaTotal} \times 100 \quad (5.4)$$

onde  $energiaProp$  é a proporção da contribuição da energia ( $E$ ) de uma determinada fase em relação a energia total ( $energiaTotal$ ). Conforme pode ser observado, o valor do atributo  $energiaProp$  é um valor que varia entre 1 e 100, de modo que a somatória dos atributos de cada fase (i.e.,  $energiaPropA$ ,  $energiaPropB$  e  $energiaPropC$ ) equivale a 100.

Na quinta e última etapa, os atributos de máximo e mínimo foram extraídos do ciclo pós-falta normalizado do sinal da corrente ( $\hat{X}$ ), individualmente para cada fase. Portanto, nosso vetor de características é constituído por dez atributos:  $groundDetection$ ,  $maxA$ ,  $minA$ ,  $energiaPropA$ ,  $maxB$ ,  $minB$ ,  $energiaPropB$ ,  $maxC$ ,  $minC$  e  $energiaPropC$ .

### 5.1.2 Predição

A classificação do tipo de falha foi realizado pelo algoritmo Árvores Extremamente Randomizadas (ERT) apresentado na Subseção 2.2.3. O número de estimadores no algoritmo foi estabelecido em 50 árvores neste trabalho. Além da tarefa de classificar o tipo da falha, o ERT também é responsável por prover as probabilidades de cada tipo corresponderem aos dados de entrada, fornecendo tais probabilidades como informação complementar aos operadores do SEP. O algoritmo de árvore de decisão utilizado no ERT é o CART.

## 5.2 COMPONENTE LOCALIZADOR DE FALTA

Nesta seção descrevemos o método proposto para a localização de falhas em LTs, cujo fluxograma está retratado na Figura 5.3. Vale mencionar que as entradas deste componente correspondem ao arquivo de oscilografia e a informação do instante de início da falha, similarmente ao descrito para o componente anterior.

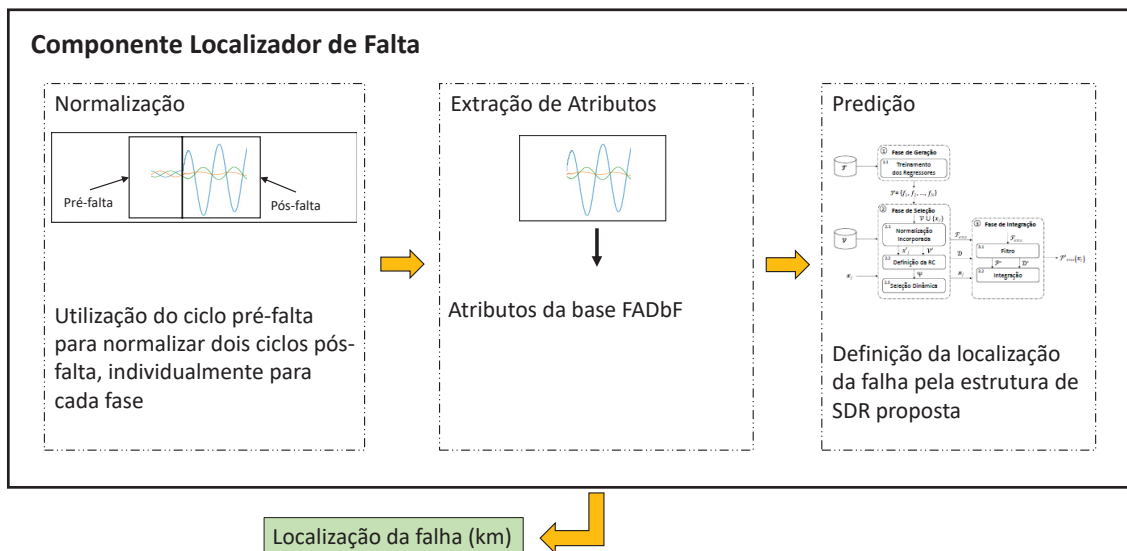


Figura 5.3: Fluxograma do componente proposto para a localização de falha

Fonte: autoria própria

A seção está organizada na seguinte sequência. Na Subseção 5.2.1 relatamos sobre os atributos extraídos e utilizados para este componente, enquanto na Subseção 5.2.2 introduzimos uma nova estrutura de seleção dinâmica de regressores empregada para estimar a distância do ponto de incidência da falha na LT. Por fim, apresentamos na Subseção 5.2.3 o algoritmo utilizado para treinar os modelos que serão usados como *pool* de regressores para a SDR.

### 5.2.1 Normalização e Extração de Atributos

O processo de extração de atributos ocorre a partir de uma sequência de três etapas, com encargos, de modo geral, semelhantes aos descritos para o componente de classificação (Seção 5.1). A primeira etapa compreende a obtenção de um ciclo pré-falta e dois ciclos pós-falta dos sinais da corrente e da tensão para cada uma das três fases a partir do arquivo de oscilografia dado como entrada.

A segunda etapa, responsável pela normalização dos ciclos pós-falta, ocorre de maneira análoga ao reportado para o componente de classificação mediante a Equação 5.2. Todavia, há duas particularidades pontuais que divergem entre ambas as etapas dos componentes. Para a localização, a normalização do sinal pós-falta ocorre para dois ciclos em vez de apenas um. Além disso, este componente atua tanto na forma de onda da corrente quanto da tensão. Conseqüentemente, o ciclo pré-falta utilizado para normalizar os dois ciclos pós-falta da tensão corresponde ao próprio sinal da tensão.

Por sua vez, na terceira e última etapa extraímos os mesmos atributos que descrevemos na Seção 4.2. Em outras palavras, as características obtidas para este componente representam a base FADbF, incluindo apenas a etapa de normalização que antecede a realização deste processo.

### 5.2.2 Estrutura SDR Proposta

Uma visão geral da arquitetura SDR proposta está apresentada na Figura 5.4. Nas subseções a seguir, descrevemos cada fase que compõe nossa estrutura.

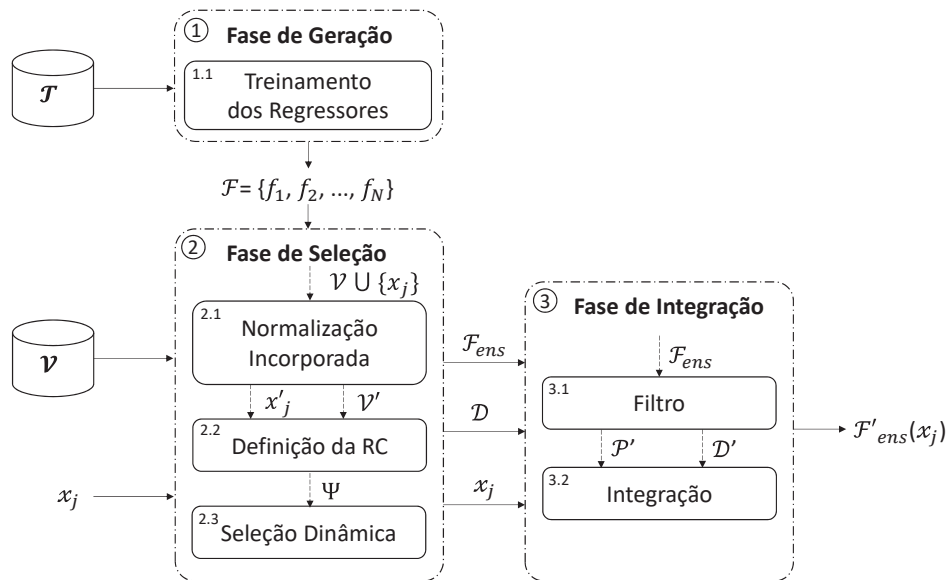


Figura 5.4: Visão geral da estrutura SDR proposta

Fonte: autoria própria

### 5.2.2.1 Fase de Geração

Esta fase, composta por um único módulo denominado “Treinamento dos Regressores”, é responsável pela geração do conjunto de estimadores  $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$ . Para esta finalidade, as mesmas propriedades discutidas na Subsubseção 2.2.2.1 são válidas para o treinamento dos modelos do conjunto. Em síntese, podemos utilizar as técnicas *bagging* e RSM para treinar os regressores e gerar a diversidade necessária, especialmente quando utilizado um conjunto homogêneo de algoritmos base. No caso de um conjunto heterogêneo, cada algoritmo base pode ser treinado ou com todos os exemplos da base de treinamento ou com diferentes partições de tal base. Alternativamente, podemos usar os modelos gerados por algoritmos como RF e ERT. Ao final, esta fase provê ao estágio seguinte o *pool* de regressores treinados  $\mathcal{F}$ .

### 5.2.2.2 Fase de Seleção

A segunda fase é responsável por selecionar dinamicamente os regressores mais competentes do conjunto de modelos  $\mathcal{F}$ , constituído por três módulos: “Normalização Incorporada”; “Definição da RC”; e “Seleção Dinâmica”. Esta fase contém os valiosos hiperparâmetros que nosso método explora para obter melhores desempenhos, os quais estão adequadamente identificados ao longo da descrição dos módulos a seguir. Cabe esclarecer que ajustar os hiperparâmetros automaticamente para definir a melhor configuração não faz parte do método. Portanto, os pesquisadores deverão avaliá-los empiricamente a partir da comparação de múltiplas configurações de modo a determinar o melhor arranjo de valores.

**Normalização Incorporada.** Este primeiro módulo corresponde a normalização do conjunto de validação  $\mathcal{V}$  que virá a ser utilizado para determinar a RC  $\Psi$ . Usualmente, os métodos do estado da arte realizam a etapa da normalização como uma etapa de pré-processamento de modo externo a estrutura SDR, além de que os próprios algoritmos do *pool* são treinados com tais dados escalonados (Moura et al., 2020; Silva et al., 2021), fornecendo o conjunto já transformado para a respectiva estrutura. A realização dessa etapa é essencial principalmente para algoritmos baseados em distância, como o KNN usado para determinar a RC (Singh e Singh, 2020).

Antes de executar o processo de normalização, adicionamos o padrão de teste  $x_j$  ao conjunto de validação original ( $\mathcal{V}' = \mathcal{V} \cup \{x_j\}$ ) para, então, realizarmos a normalização deste conjunto. Em seguida, removemos o padrão de teste escalonado  $x'_j$  do conjunto  $\mathcal{V}'$ , de modo a usá-lo como entrada ao algoritmo KNN determinar os exemplos mais próximos do padrão de teste escalonado  $x'_j$  no conjunto de validação transformado  $\mathcal{V}'$ . Em nossa estrutura SDR, qualquer algoritmo de normalização pode ser usado para transformar os dados originais em um mesmo intervalo de valores, como as técnicas MinMax e Z-score. Esta independência do procedimento de escalonamento é crucial, uma vez que também pode depender do problema (Amorim et al., 2023).

Vale ressaltar que esse processo é repetido individualmente para cada novo padrão de teste a ser predito. Portanto, utilizamos a etapa de normalização apenas para determinarmos os vizinhos em torno do padrão de teste escalonado  $x'_j$ , considerando o conjunto de validação transformado  $\mathcal{V}'$ , durante o segundo módulo (“Definição da RC”). No presente módulo, não há hiperparâmetro a ser ajustado. Deve-se enfatizar que este tipo de normalização integrada ao método SDR é proposta e exclusiva do nosso *framework*.



**Definição da RC.** Utilizamos o algoritmo KNN para definir a RC  $\Psi$  de modo semelhante ao já realizado no estado da arte. Todavia, avaliamos um hiperparâmetro inexplorado na literatura da SDR: a função para o cálculo da distância. Embora os trabalhos relacionados usem a métrica Euclidiana como padrão, a escolha desta medida não deve ser negligenciada visto que este hiperparâmetro também pode repercutir em uma melhoria no desempenho. Em resposta a esta questão, também avaliamos outras medidas de distância: Cityblock, Geodésica e Cosseno (Barboza et al., 2023).

Em resumo, o segundo módulo é encarregado de obter os índices dos vizinhos mais próximos de  $x'_j$ , determinados em  $\mathcal{V}'$  pelo algoritmo KNN, para então recuperar as respectivas amostras no conjunto de validação original  $\mathcal{V}$ , criando, assim, a RC  $\Psi$ . Como pode ser observado, a normalização no primeiro módulo foi utilizada apenas para determinar as amostras mais próximas do padrão de teste escalonado  $x'_j$ , enquanto a RC é composta pelas amostras originais. Os hiperparâmetros deste módulo são a função de distância e o valor de  $K$  do algoritmo KNN que define o tamanho da RC  $\Psi$ .

**Seleção Dinâmica.** Por sua vez, o terceiro módulo realiza a seleção dos modelos mais competentes a partir de seus desempenhos na RC  $\Psi$ . Para tanto, diversas medidas de competência podem ser utilizadas para determinar o desempenho de cada modelo do *pool*, como o erro quadrático máximo e a variância, entre outros possíveis indicadores (Moura et al., 2019b). Em seguida, emprega-se uma regra/critério para selecionar os melhores regressores.

O critério para seleção dos melhores regressores pode ser exercido de diferentes maneiras, mas a regra padrão corresponde àquela definida por Rooney et al. (2004), conforme mencionado na Subseção 2.2.4, a qual seleciona apenas os regressores cujos erros cumulativos estão abaixo do intervalo  $\frac{E_{max} - E_{min}}{2}$ . Em contrapartida, também podemos utilizar outras estratégias para estabelecer a regra de seleção, como mediana, média harmônica, média aritmética e média geométrica. Neste cenário, podemos calcular o valor mediano/médio entre os desempenhos com base na medida de competência considerada e utilizar esse valor mediano/médio como limite, selecionando apenas os regressores que apresentarem erro inferior a este limiar. Os hiperparâmetros deste módulo incluem a medida de competência e o critério de seleção.

A saída desta fase corresponde ao conjunto de modelos selecionados  $\mathcal{F}_{ens}$  junto aos respectivos desempenhos  $\mathcal{D}$ , além do padrão de teste  $x_j$ .

### 5.2.2.3 Fase de Integração

Nesta última fase, constituída por dois módulos, ocorre a combinação das predições dos regressores selecionados na fase anterior. Entretanto, antes das predições serem agregadas de fato, realizamos uma etapa de pós-processamento que denominamos de filtro, sendo esta uma de nossas principais contribuições no método SDR proposto.

**Filtro.** A ideia básica por trás do filtro é identificar e remover as predições discrepantes (*outliers*) provenientes dos regressores  $\mathcal{F}_{ens}$  que foram selecionados na fase anterior. Para isso, calcula-se inicialmente a média e o desvio padrão das predições dos regressores  $\mathcal{F}_{ens}$  para o padrão de teste  $x_j$ . Em seguida, obtêm-se o limite superior (média somada ao desvio padrão multiplicado por um fator) e inferior (média subtraída pelo desvio padrão multiplicado por um fator). O fator utilizado possui o valor três (3). Por fim, as predições com valores maiores ao limite superior e menores ao limite inferior são removidas. O único hiperparâmetro deste módulo é o valor do fator utilizado para a detecção de *outliers*.

**Integração.** A primeira responsabilidade deste módulo é a definição dos pesos  $\mathcal{W} = \{\alpha_1, \alpha_2, \dots, \alpha_M\}$  das  $M$  predições  $\mathcal{P}'$  selecionadas após a aplicação do filtro para a predição final do método. Esta tarefa ocorre similarmente ao descrito na Subseção 2.2.4 a partir das Equações 2.1 e 2.2, considerando os desempenhos  $\mathcal{D}'$  do subconjunto de modelos



selecionados  $\mathcal{F}'_{ens}$  responsáveis pelas predições filtradas  $\mathcal{P}'$ . Por sua vez, a predição final advém de acordo com a Equação 2.3, utilizando as predições  $\mathcal{P}'$  do subconjunto de modelos selecionados  $\mathcal{F}'_{ens}$  e seus respectivos pesos  $\mathcal{W}$ . Neste módulo não há hiperparâmetro a ser ajustado.

Sumarizamos na Tabela 5.1 os hiperparâmetros do *framework* SDR proposto junto a exemplos de valores que tais hiperparâmetros podem assumir.

Tabela 5.1: Hiperparâmetros da estrutura SDR proposta (em negrito valores utilizados como padrão neste trabalho)

Nome	Valores, mas que não se limitam a
Função de distância	Euclidiana, Cosseno, Geodésica, <b>Cityblock</b>
$K$	5, 10, 15, 20, <b>30</b> , 40, 50
Medida de competência*	variância, soma do erro absoluto, soma do erro quadrático, erro quadrático mínimo, <b>erro quadrático máximo</b> , semelhança do vizinho, erro quadrático da soma raiz, erro quadrático mais próximo
Critério de seleção	critério padrão, mediana, média aritmética, <b>média geométrica</b> , média harmônica
Fator do filtro	1, 2, <b>3</b> , 4, 5

\*Acesse Moura et al. (2019b) para detalhes sobre cada medida de competência.

### 5.2.3 Predição

A tarefa de localização de falha, conforme previamente descrito, ocorre a partir da estrutura SDR que propomos. Para isso, treinamos o algoritmo ERT com  $N = 700$  estimadores e usamos as suas árvores para comporem nosso conjunto (*pool*) de regressores  $\mathcal{F}$ . Assim como mencionado para o componente classificador de falta (Subseção 5.1.2), o algoritmo de árvore de decisão utilizado no ERT é o CART. Identificamos em negrito na Tabela 5.1 a combinação de valores determinados como a melhor configuração para este estudo, conforme será experimentalmente justificado na Seção 6.2.

Além disso, o modelo foi treinado com o atributo-alvo em termos percentuais do comprimento da LT. Consequentemente, a saída do modelo consiste em um valor real entre [1%, 100%], correspondente à porcentagem do comprimento da linha onde a falta foi localizada. Este valor é convertido para distância em quilômetros (km) após a predição pela Equação 5.5. Assim, todos os resultados relatados e discutidos na Seção 6.2 estão em quilômetros.

$$distância_{km} = \frac{LT_{comprimento} * distância_{porcentagem}}{100} \quad (5.5)$$

onde  $distância_{km}$  corresponde a localização da falha em km,  $LT_{comprimento}$  é o comprimento (extensão) da linha de transmissão em km, e  $distância_{porcentagem}$  é a localização da falha em termos percentuais da extensão da linha.

### 5.3 FERRAMENTAS

Os experimentos realizados em nosso trabalho contemplaram a linguagem de programação Python<sup>1</sup>, em conjunto com as implementações dos algoritmos de AM (*e.g.*, ERT) predominantemente fornecidas pela biblioteca scikit-learn<sup>2</sup> (Pedregosa et al., 2011). Duas exceções são os algoritmos XGBoost<sup>3</sup> (*eXtreme Gradient Boosting*) e LightGBM<sup>4</sup> (*Light Gradient Boosting Machine*) (Chen e Guestrin, 2016; Ke et al., 2017), os quais também analisamos ao longo do projeto de modo a determinar as melhores técnicas e justificar o emprego do ERT como preditor base, tanto para a classificação quanto para a localização.

### 5.4 CONSIDERAÇÕES FINAIS

Apresentamos neste capítulo o método proposto para a classificação do tipo de falta e para a localização da falha em linhas de transmissão. Entre as propriedades que buscamos para o nosso método está a realização dessas atividades independentemente da taxa de amostragem, além da capacidade de generalizá-las para linhas de transmissão com configurações distintas. Inclusive, esses aspectos foram identificados como problemas nas abordagens encontradas no estado da arte, uma vez que apresentam baixo desempenho ou nem chegam a considerar essas particularidades. Cabe enfatizar que o método SDR proposto pode ser aplicado em qualquer problema de regressão, não se limitando apenas a tarefa de localização. No próximo capítulo descrevemos os protocolos experimentais adotados para a avaliação do nosso método, seguidos pela apresentação e discussão dos respectivos resultados alcançados.

---

<sup>1</sup>Versão 3.11.1, <https://www.python.org/>

<sup>2</sup>Versão 1.2.0, <https://scikit-learn.org/stable/>

<sup>3</sup>Versão 4.0.0, <https://lightgbm.readthedocs.io/en/latest/>

<sup>4</sup>Versão 1.7.6, <https://xgboost.readthedocs.io/en/stable/index.html>

## 6 RESULTADOS E DISCUSSÃO

Neste capítulo, dividido em duas seções, apresentamos e discutimos os resultados alcançados ao longo do projeto. Na Seção 6.1 examinamos os resultados referentes a abordagem proposta para a classificação de falhas em linhas de transmissão, enquanto na Seção 6.2 tratamos dos resultados perante a localização de faltas.

Os diversos experimentos abrangentes realizados foram comparados estatisticamente, visando fornecer conclusões sólidas sobre as melhores estratégias e configurações. Utilizamos o teste de normalidade de D'Agostino & Pearson ( $p\text{-value} < 0,05$ ) para todos os grupos de dados para então selecionarmos os testes de hipótese mais apropriados com base na distribuição dos dados (paramétrico ou não-paramétrico). Os respectivos testes de hipótese usados estão citados ao longo das Seções 6.1 e 6.2, todos com intervalo de confiança de 95%. Assim, utilizamos diferentes testes estatísticos ao longo das análises referentes aos três aspectos essenciais para a escolha do melhor teste individualmente para cada cenário de comparação: número de grupos, distribuição dos dados e grupos pareados/não-pareados.

### 6.1 CLASSIFICAÇÃO DE FALTA

Esta seção divide-se em seis subseções com experimentos abrangentes de avaliação de nosso método, nas quais analisamos as seguintes questões de pesquisa (QP):

- **QP1:** a engenharia de atributos desenvolvida habilita o método proposto a generalizar para LTs desconhecidas (*i.e.*, linhas não usadas para o treinamento do modelo)?
- **QP2:** o método treinado com dados simulados (sintéticos) é capaz de classificar eficientemente eventos reais de falta?

Iniciamos com a descrição do protocolo experimental na Subseção 6.1.1, cujos detalhes sustentam as investigações que seguem. Na Subseção 6.1.2 estão apresentados os desempenhos do esquema proposto quando o método é treinado e avaliado com dados provenientes da uma mesma linha de transmissão. Na Subseção 6.1.3, inicialmente reportamos os resultados de experimentos complementares para justificar e ressaltar a importância da engenharia de atributos realizada para a classificação. Em seguida, nesta mesma subseção, apresentamos os desempenhos definitivos do nosso método para a generalização, ou seja, a capacidade de classificar eventos de faltas para uma LT distinta daquela empregada para treinar o algoritmo. Na Subseção 6.1.4 estão relatados os resultados quando empregamos eventos reais de falhas em linhas de transmissão coletados da rede da COPEL. Já na Subseção 6.1.5 analisamos as importâncias dos atributos para o algoritmo ERT usado como classificador, complementando as análises dispostas nas subseções anteriores. Por fim, na Subseção 6.1.6 analisamos o desempenho de diversos algoritmos de AM (CART, SVM, KNN, MLP, LightGBM, XGBoost, RF e ERT) e comparamos o desempenho do nosso método frente a trabalhos encontrados na literatura.

#### 6.1.1 Protocolo Experimental

A classificação do tipo de falha foi realizado pelo algoritmo ERT, vide Subseção 5.1.2, com o número de árvores estabelecido em 50 árvores. Outros valores para esse parâmetro também foram avaliados, mas o montante de 50 estimadores demonstrou os melhores resultados

em experimentos preliminares. Além da tarefa de classificar o tipo da falta, o ERT também é responsável por prover as probabilidades de cada tipo corresponderem aos dados de entrada, fornecendo tais probabilidades como informação complementar aos operadores do SEP.

O treinamento e a avaliação do algoritmo ocorreu com exemplos provenientes de ambos os terminais conjuntamente. Assim, as simulações da base de dados FADb para a LT de 500 kV, descrita na Subseção 4.1.1, foram aleatoriamente divididas nos conjuntos de treinamento, validação e teste. O conjunto de treinamento compreende 50% de todos os exemplos (168.000 simulações, sendo 84.000 exemplos de cada terminal), enquanto os conjuntos de validação e teste contêm 25% de todos os eventos cada (84.000 simulações, sendo 42.000 exemplos de cada lado da linha). A escolha por essa divisão em subconjuntos (método *hold-out*) ocorreu devido ao grande volume de dados da nossa base (Stapor et al., 2021). O particionamento em conjuntos de treinamento, validação e teste foi realizado dez vezes, resultando em diferentes conjuntos com instâncias distintas entre cada repetição. Os resultados apresentados no Capítulo 6 retratam a média e o desvio padrão referentes a essas dez repetições.

Além disso, os dados originais foram reamostrados para outras três taxas de amostragem. Portanto, avaliamos nosso método para quatro cenários diferentes: 256 Hz, 1 kHz, 5 kHz e 10 kHz. Os desempenhos com o uso do sinal da corrente foram estatisticamente comparados por meio de testes apropriados com base nas características dos dados, com o objetivo de verificar a existência de diferença estatisticamente significativa entre eles. Os respectivos testes estão apresentados ao longo das próximas subseções.

Também analisamos a capacidade de generalização do nosso modelo para classificar faltas para linhas de transmissão com configuração distinta daquela empregada para treinar o algoritmo. Em um estudo de caso, utilizamos os dados da LT de 230 kV descritos na Subseção 4.1.2. É digno de nota que o algoritmo foi treinado com o conjunto de treinamento previamente mencionado (linha de transmissão de 500 kV, 414 km, 60 Hz) e testado com todos os exemplos da LT de 230 kV. Para o melhor entendimento do leitor, vamos identificar a linha de 500 kV (Subseção 4.1.1) como LT1 (linha de transmissão 1), enquanto esse novo conjunto de dados para a linha de 230 kV será nomeado como LT2 (linha de transmissão 2).

Além dos dados sintéticos (simulados), nós também avaliamos o método proposto frente a eventos reais de falta coletados da rede da COPEL, conforme retratados pela Tabela 6.1. É digno de nota que esses eventos são provenientes de diferentes linhas de transmissão, cada qual com suas respectivas características. Todas essas falhas foram coletadas pelos equipamentos de proteção a uma taxa de amostragem de 15,36 kHz, entre Janeiro/2019 e Março/2020.

Observe que alguns eventos possuem apenas dados para um dos terminais da linha. Apesar de serem 35 ocorrências ao todo, consideramos os dados de cada terminal como um caso a parte. Assim, o número total de eventos corresponde a 62, já desconsiderando os terminais cujos dados estão indisponíveis. Vale ressaltar que esses eventos são raros, os quais resultam em uma dificuldade no estabelecimento de um conjunto de treinamento representativo composto apenas por casos reais. Logo, os eventos simulados visam fornecer um conjunto de treinamento com parâmetros e condições de falha diversificados.

O algoritmo foi treinado com o mesmo conjunto de treinamento anteriormente mencionado (linha de transmissão de 500 kV, 414 km e 60 Hz) para um total de dez repetições por evento. Repare ainda que os dados de treinamento representam taxa de amostragem de 10 kHz, enquanto os casos de falhas reais permanecem com sua taxa de amostragem original de 15,36 kHz, sem a realização de subamostragem.

Tabela 6.1: Eventos reais de faltas em linhas de transmissão da rede da COPEL coletados entre Janeiro/2019 e Março/2020

ID	Tensão	Linha de transmissão	Comprimento	Causa da falha	Tipo	Disponibilidade dos dados
#1	230 kV	CASCADEL - FIGUACUNORTE	124 km	Indeterminada	AG	2 terminais
#2	525 kV	CASCADEL OEST - FIGUACU	117 km	Indeterminada	BG	2 terminais
#3	525 kV	LONDRINA ESU - ASSIS	122 km	Descarga atmosférica	CG	2 terminais
#4	230 kV	KLABIN - PONTA G NORTE	139 km	Descarga atmosférica	AG	2 terminais
#5	230 kV	MARINGA - SARANDI	19 km	Vegetação	AG	2 terminais
#6	230 kV	STA QUITERIA - UMBARA	25 km	Vegetação	AG	1 terminal
#7	230 kV	CURIT.LESTE - POSTO FISCAL	61 km	Descarga atmosférica	CG	2 terminais
#8	230 kV	FIGUEIRA - JAGUARIAIVA	83 km	Pássaro Curicaca	AG	2 terminais
#9	230 kV	CASCADEL N - CASCADEL O	37 km	Indeterminada	CG	2 terminais
#10	230 kV	BATEIAS - JAGUARIAIVA	137 km	Descarga atmosférica	BCG	1 terminal
#11	230 kV	LONDRINA ESU - FIGUEIRA	100 km	Descarga atmosférica	CG	1 terminal
#12	230 kV	FIGUEIRA - KLABIN CELULOSE	85 km	Indeterminada	CG	2 terminais
#13	500 kV	ARARAQUARA 2 - TAUBATE	334 km	Indeterminada	AG	2 terminais
#14	230 kV	UBERABA - UMBARA	37 km	Indeterminada	BG	2 terminais
#15	230 kV	LONDRINA ESU - FIGUEIRA	100 km	Indeterminada	CG	1 terminal
#16	230 kV	FIGUEIRA - KLABIN CELULOSE	85 km	Pássaro Curicaca	CG	2 terminais
#17	230 kV	BATEIAS - JAGUARIAIVA	138 km	Descarga atmosférica	ABC	1 terminal
#18	500 kV	ARARAQUARA 2 - TAUBATE	334 km	Descarga atmosférica	AG	2 terminais
#19	230 kV	LONDRINA ESU - FIGUEIRA	92 km	Indeterminada	CG	2 terminais
#20	230 kV	CASCADEL - S. OSORIO	80 km	Queimada	AG	1 terminal
#21	230 kV	CURIT.LESTE - POSTO FISCAL	61 km	Descarga atmosférica	ABC	1 terminal
#22	230 kV	PARIGOT SOUZA - POSTO FISCAL	65 km	Descarga atmosférica	ACG	2 terminais
#23	230 kV	FIGUEIRA - KLABIN CELULOSE	85 km	Pássaro Curicaca	CG	2 terminais
#24	500 kV	ARARAQUARA 2 - TAUBATE	334 km	Queimada	CG	2 terminais
#25	500 kV	ARARAQUARA 2 - TAUBATE	334 km	Queimada	CG	2 terminais
#26	230 kV	FIGUEIRA - CHAVANTES	105 km	Pássaro Curicaca	BG	2 terminais
#27	230 kV	FIGUEIRA - CHAVANTES	105 km	Indeterminada	AG	2 terminais
#28	230 kV	LONDRINA ESU - FIGUEIRA	92 km	Descarga atmosférica	AG	2 terminais
#29	525 kV	CASCADEL OEST - FIGUACU	117 km	Indeterminada	AG	2 terminais
#30	525 kV	CASCADEL OEST - FIGUACU	117 km	Vegetação	AG	2 terminais
#31	230 kV	JAGUARIAIVA - MAUA	106 km	Indeterminada	BC	2 terminais
#32	230 kV	CASCADEL - S. OSORIO	80 km	Máquina agrícola	BG	1 terminal
#33	230 kV	LONDRINA ESU - FIGUEIRA	92 km	Descarga atmosférica	BG	2 terminais
#34	525 kV	CASCADEL OEST - FIGUACU	117 km	Indeterminada	BG	2 terminais
#35	230 kV	KLABIN - PONTA G NORTE	139 km	Pivô central (irrigação)	AG	2 terminais

De modo a justificar a importância de nossa engenharia de atributos, realizamos experimentos abrangentes para retratar os resultados tanto sem o processo de normalização quanto sem os atributos *groundDetection* e aqueles baseados na energia (*energiaPropA*, *energiaPropB* e *energiaPropC*). Os experimentos foram avaliados pelas métricas da acurácia (Equação 6.1), precisão (Equação 6.2) e F1-Score (Equação 6.3). A escolha por estas métricas justifica-se por elas serem as medidas usadas pela literatura para analisar o desempenho dos métodos de classificação de faltas, principalmente a acurácia.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (6.1)$$

$$Precisão = \frac{VP}{VP + FP} \quad (6.2)$$

$$F1 - Score = \frac{VP}{VP + \frac{FN+FP}{2}} \quad (6.3)$$

sendo VP, VN, FP e FN os valores de verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo, respectivamente.

#### 6.1.2 Desempenho do Método Proposto para a LT de 500 kV

O desempenho do método proposto para o conjunto de validação está apresentado na Tabela 6.2 considerando as quatro taxas de amostragem avaliadas. Os valores para todas as métricas de avaliação utilizadas são superiores à 99,99% para todas as frequências.

Tabela 6.2: Desempenho do método proposto para a classificação do tipo de falha (conjunto de validação)

Taxa de amostragem	Acurácia	Precisão	F1-Score
<b>256 Hz</b>	99,9980% ± 0,0024%	99,9980% ± 0,0024%	99,9980% ± 0,0024%
<b>1 kHz</b>	99,9993% ± 0,0010%	99,9993% ± 0,0010%	99,9993% ± 0,0010%
<b>5 kHz</b>	99,9989% ± 0,0012%	99,9989% ± 0,0012%	99,9989% ± 0,0012%
<b>10 kHz</b>	99,9992% ± 0,0013%	99,9992% ± 0,0013%	99,9992% ± 0,0013%

Em seguida, avaliamos o método utilizando o conjunto de teste para mensurar o desempenho real da proposta, com exemplos diferentes daqueles empregados no conjunto de validação. Cabe destacar que o conjunto de validação foi usado primordialmente com o objetivo de determinar a melhor configuração (atributos e algoritmo) para a nossa abordagem, conforme será detalhado ao longo das próximas subseções. Nas Tabelas 6.3, 6.4, 6.5 e 6.6 estão reportados os desempenhos do método para o conjunto de teste, individualmente para cada tipo de falha e separadamente para cada frequência de amostragem considerada.

Tabela 6.3: Desempenho do método proposto para a taxa de amostragem de 256 Hz (conjunto de teste)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	99,9988% ± 0,0037%	99,9964% ± 0,0057%	99,9976% ± 0,0031%
BG	99,9988% ± 0,0037%	100,0000% ± 0,0000%	99,9994% ± 0,0019%
CG	100,0000% ± 0,0000%	99,9976% ± 0,0050%	99,9988% ± 0,0025%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	99,9988% ± 0,0038%	99,9928% ± 0,0116%	99,9958% ± 0,0057%
ACG	99,9893% ± 0,0117%	99,9988% ± 0,0038%	99,9941% ± 0,0056%
BCG	99,9988% ± 0,0037%	99,9988% ± 0,0038%	99,9988% ± 0,0025%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9985% ± 0,0010%	99,9985% ± 0,0010%	99,9985% ± 0,0010%

Tabela 6.4: Desempenho do método proposto para a taxa de amostragem de 1 kHz (conjunto de teste)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	100,0000% ± 0,0000%	99,9988% ± 0,0038%	99,9994% ± 0,0019%
BG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
CG	99,9988% ± 0,0038%	99,9988% ± 0,0037%	99,9988% ± 0,0025%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,9988% ± 0,0038%	99,9994% ± 0,0019%
ACG	99,9976% ± 0,0050%	99,9988% ± 0,0038%	99,9982% ± 0,0029%
BCG	99,9988% ± 0,0038%	100,0000% ± 0,0000%	99,9994% ± 0,0019%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9995% ± 0,0006%	99,9995% ± 0,0006%	99,9995% ± 0,0006%

Tabela 6.5: Desempenho do método proposto para a taxa de amostragem de 5 kHz (conjunto de teste)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BG	100,0000% ± 0,0000%	99,9976% ± 0,0050%	99,9988% ± 0,0025%
CG	100,0000% ± 0,0000%	99,9977% ± 0,0074%	99,9988% ± 0,0037%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,9988% ± 0,0037%	99,9994% ± 0,0019%
ACG	99,9976% ± 0,0075%	100,0000% ± 0,0000%	99,9988% ± 0,0038%
BCG	99,9964% ± 0,0058%	100,0000% ± 0,0000%	99,9982% ± 0,0029%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9994% ± 0,0008%	99,9994% ± 0,0008%	99,9994% ± 0,0008%



Tabela 6.6: Desempenho do método proposto para a taxa de amostragem de 10 kHz (conjunto de teste)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BG	100,0000% ± 0,0000%	99,9988% ± 0,0037%	99,9994% ± 0,0019%
CG	100,0000% ± 0,0000%	99,9965% ± 0,0079%	99,9982% ± 0,0040%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ACG	99,9964% ± 0,0080%	100,0000% ± 0,0000%	99,9982% ± 0,0040%
BCG	99,9988% ± 0,0038%	100,0000% ± 0,0000%	99,9994% ± 0,0019%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9995% ± 0,0008%	99,9995% ± 0,0008%	99,9995% ± 0,0008%

De maneira similar aos resultados apresentados para o conjunto de validação, o método apresentou alto desempenho com as médias da acurácia superiores à 99,99% para os casos de teste. É digno de nota que a falta trifásica ABC, o tipo de falha mais perigoso para o sistema de potência (Prasad et al., 2018), foi corretamente classificado para todos os casos de teste com acurácia de 100%.

Nós também realizamos análises estatísticas para comparar os desempenhos entre as diferentes frequências a partir da principal métrica utilizada pela literatura para avaliar a classificação do tipo de falha: acurácia. Os desempenhos entre os quatro grupos (256 Hz, 1 kHz, 5 kHz e 10 kHz) foram comparados por meio do teste pareado de Friedman com intervalo de confiança de 95%. Os resultados revelaram que não há diferença estatisticamente significativa entre os resultados ( $p$ -valor = 0,1609). O interesse nessa análise se deve ao fato da constatação da efetividade do método independentemente da taxa de amostragem, sem a ocorrência de perdas significativas para frequências menores, condição a qual afeta outros métodos encontrados no estado da arte (Chen et al., 2018a; Haq et al., 2021).

Conforme pode ser identificado, nossa abordagem utiliza instâncias de ambos os terminais da linha de transmissão apenas para o treinamento do algoritmo. Em vista disso, uma vez que o algoritmo esteja treinado, pode ser usado para classificar falhas para qualquer terminal individualmente. Assim, não há a necessidade de obtenção simultânea e sincronização dos dados coletados entre os terminais. Cabe mencionar que ambas as propriedades (*i.e.*, flexibilidade para diferentes taxas de amostragem e uso de dados de um único terminal) não são exclusivas do nosso trabalho, de modo que também podem ser encontradas em outras obras do estado da arte. O grande diferencial do nosso método está na capacidade de generalização, a qual abordaremos nas subseções seguintes.

### 6.1.3 Desempenho do Método Proposto para Generalização

Com o intuito de analisar a capacidade da nossa abordagem para generalizar a classificação para uma linha de transmissão com configuração diferente daquela para a qual foi treinada, um segundo conjunto de dados foi gerado sinteticamente para uma linha de transmissão de 230 kV, 106,646 km, 60 Hz (LT2). Um total de 41.000 eventos de falha, cujos parâmetros de falha foram descritos na Subseção 4.1.2, foram simulados e usados apenas como um conjunto de teste. O treinamento do algoritmo ocorreu por meio do uso dos dados provenientes exclusivamente do conjunto de treinamento da LT1, de modo que os desempenhos expostos a seguir correspondem ao uso do conjunto de validação da LT1 e todo o conjunto da LT2.

No decorrer desta subseção realizamos um *ablation study* em que apresentamos inicialmente os resultados sem o processo de normalização, sem o atributo *groundDetection* e sem os atributos baseados na energia (*energiaPropA*, *energiaPropB* e *energiaPropC*), de modo a respaldar a importância desses aspectos em nosso método. Em outras palavras, neste primeiro experimento utilizamos apenas os atributos de máximos e mínimos extraídos do ciclo pós-falta sem qualquer tratamento (*i.e.*, sem a normalização). Na Tabela 6.7 estão reportados os desempenhos do modelo seguindo este primeiro cenário avaliativo. Cabe salientar que, nesta subseção, as tabelas cujos resultados não repercutem o método proposto na íntegra, ou seja, sem a normalização ou alguma das características originais, identificam em seu rodapé quais procedimentos compõem tais análises (✓) e quais não as compõem (✗).

Tabela 6.7: Média e desvio padrão da acurácia para a classificação do tipo de falha utilizando apenas os atributos de máximos e mínimos

Taxa de amostragem	LT1	LT2
<b>256 Hz</b>	99,9138% $\pm$ 0,0055%	39,6527% $\pm$ 2,1842%
<b>1 kHz</b>	99,9400% $\pm$ 0,0094%	40,6251% $\pm$ 1,3810%
<b>5 kHz</b>	99,9400% $\pm$ 0,0202%	42,1098% $\pm$ 2,0169%
<b>10 kHz</b>	99,9440% $\pm$ 0,0171%	41,3361% $\pm$ 1,5396%

✗ Normalização  
✓ Atributos máximos e mínimos  
✗ Atributo *groundDetection*  
✗ Atributos baseados na energia

Observa-se que o modelo apresentou desempenhos satisfatórios para a classificação de exemplos quando a linha de transmissão com a mesma configuração foi utilizada para o treinamento e para a avaliação do algoritmo (LT1). Por outro lado, quando o modelo foi usado para classificar instâncias de uma linha de transmissão com configuração diferente daquela empregada para treinar o algoritmo (LT2), os resultados demonstraram baixa taxa de acerto. Esse comportamento pode ser justificado em parte pelo fato de que as escalas (amplitudes) dos sinais entre ambas as linhas (LT1 e LT2) são distintas, afetando o desempenho do modelo.

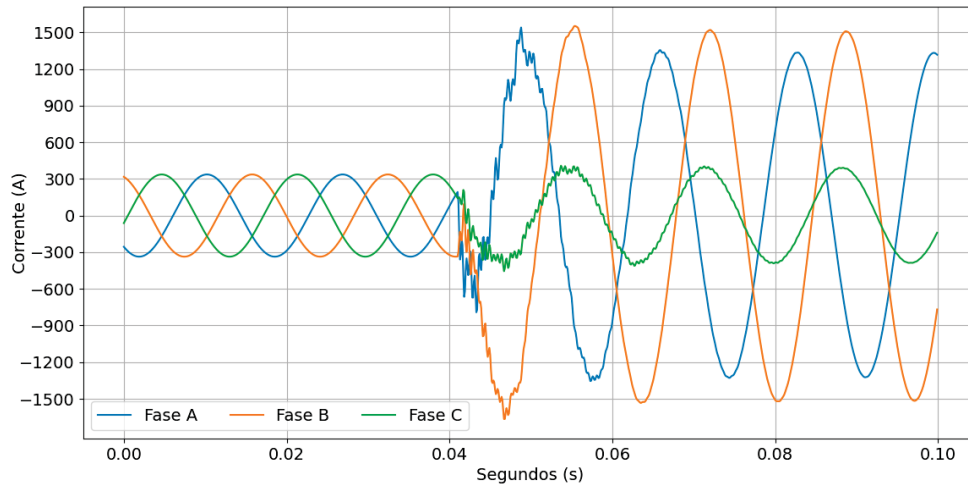
Na Tabela 6.8 estão reportados os desempenhos do algoritmo incluindo o processo de normalização, mas ainda sem os atributos baseados na energia e sem o atributo *groundDetection*. Por consequência, a etapa de normalização é essencial para permitir que o modelo possa generalizar sem distinção entre diferentes níveis de tensão das linhas de transmissão sob análise.

Tabela 6.8: Média e desvio padrão da acurácia para a classificação do tipo de falha com a inclusão da etapa de normalização

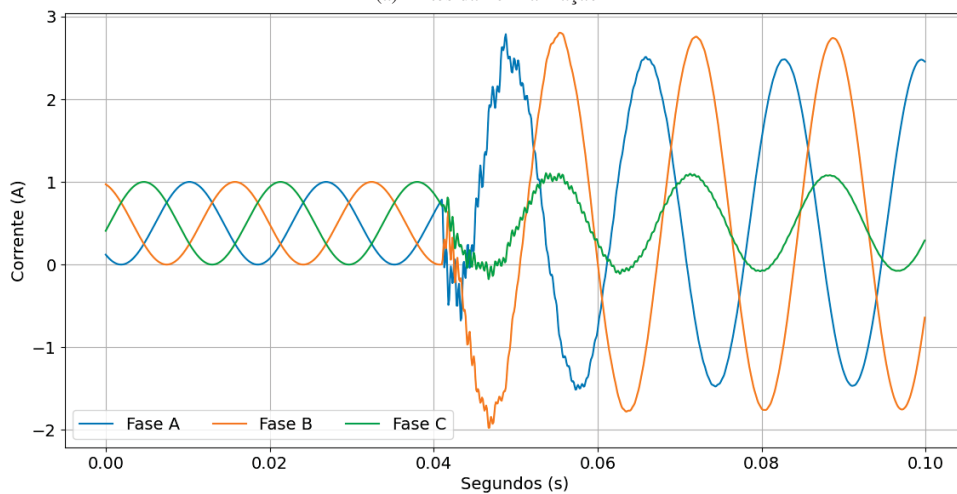
Taxa de amostragem	LT1	LT2
<b>256 Hz</b>	99,9212% $\pm$ 0,0070%	87,4556% $\pm$ 0,5107%
<b>1 kHz</b>	99,8724% $\pm$ 0,0160%	96,4012% $\pm$ 0,1898%
<b>5 kHz</b>	99,9486% $\pm$ 0,0080%	92,3822% $\pm$ 0,2796%
<b>10 kHz</b>	99,9410% $\pm$ 0,0090%	92,8051% $\pm$ 0,3359%

✓ Normalização  
✓ Atributos máximos e mínimos  
✗ Atributo *groundDetection*  
✗ Atributos baseados na energia

Conforme representada pela Figura 6.1, o escalonamento em nosso método é responsável por transformar o sinal da corrente em uma amplitude menor, tornando os valores do sinal pré-falta (nível normal de operação) próximos ao intervalo  $[0, 1]$ . Ao mesmo tempo, esse procedimento preserva o comportamento pós-falta que excede as amplitudes do nível normal de operação, mas em uma escala menor. Para este propósito, utilizamos os valores de máximos e mínimos de um ciclo pré-falta para normalizar o primeiro ciclo pós-falta proveniente da mesma oscilografia (vide Equação 5.2). Apesar da retratação de dados normalizados pré e pós-falha na Figura 6.1, o método proposto utiliza apenas o sinal pós-falta para extrair as características.



(a) Antes da normalização



(b) Depois da normalização

Figura 6.1: Representação do processo de normalização considerando um evento de falha ABG. Na imagem (a) está retratado um trecho de uma simulação original do conjunto LT1, representando o estado normal de operação (*i.e.*, sinal pré-falta) com correntes simétricas e equilibradas até o início da falha, a qual ocorre por volta do tempo 0,041 s. Conforme pode ser observado, a amplitude da corrente da forma de onda pré-falta permanece entre aproximadamente -300 A e 300 A. Portanto, o processo da normalização visa converter a oscilografia para uma escala menor, conforme apresentada pela imagem (b), porém conservando o comportamento pós-falta que excede a amplitude do nível do estado normal de operação, a qual após a normalização encontra-se no intervalo  $[0, 1]$ . Conseqüentemente, oscilografias com amplitudes distintas da primeira imagem (a) resultarão em amplitudes equivalentes a segunda imagem (b), permitindo que o método lide com qualquer oscilografia de maneira equivalente.

É crucial salientar que o processo de normalização é utilizado apenas para transformar o ciclo pós-falta para uma escala (amplitude) inferior, utilizando para isso os dados pré-falta da mesma oscilografia. Portanto, não estamos interessados em dimensionar o ciclo pós-falta ou o vetor de atributos em um intervalo fixo de valores, como  $[0, 1]$ . Esse processo é essencial para que nosso método lide com linhas de transmissão com diferentes particularidades, as quais podem acarretar em diferentes níveis de corrente (amplitudes), de modo equivalente.

Em complemento, podemos utilizar qualquer ciclo pré-falta para esse procedimento (extração de valores máximos e mínimos para a etapa de normalização), uma vez que o SEP opera predominantemente com cargas simétricas e balanceadas. Logo, o desempenho do nosso método não deve ser afetado pela existência de possíveis variações entre as amplitudes da corrente dos ciclos pré-falta, visto que essas variações serão mínimas. Em caso de valores discrepantes (*outliers*), eles representarão a incidência de uma falta na linha de transmissão sob análise, os quais irão compor os dados pós-falta.

Repare que utilizamos valores de máximos e mínimos para dois propósitos distintos em nosso trabalho: (i) extraídos do ciclo pré-falta e utilizados como parâmetros para a Equação 5.2 para normalizar o ciclo pós-falta; (ii) extraídos do ciclo pós-falta normalizado, individualmente para cada fase, para comporem o vetor de características.

Vale ressaltar que o processo de normalização em nosso método difere daquele realizado por outros trabalhos relacionados como Ray et al. (2013) e Ray e Mishra (2016), os quais empregam essa técnica como etapa de pré-processamento diretamente no vetor de características. Silva et al. (2007), por sua vez, normalizaram as amostras de tensão e corrente no intervalo de falta pelo respectivos valores máximos encontrados no registro. No entanto, os autores não especificam se o valor máximo utilizado refere-se a cada fase individualmente ou entre elas em conjunto, além dos demais problemas relatados na Seção 3.2.

É possível identificar a melhora expressiva no desempenho para a classificação dos exemplos da LT2 com a inclusão do processo de normalização por meio da comparação entre as Tabelas 6.7 e 6.8. No entanto, o algoritmo classificou erroneamente alguns eventos bifásicos LL como LLG e vice-versa, por exemplo AB como ABG, ou seja, a abordagem não conseguiu distinguir precisamente quanto à atuação ou não da terra na falha. Tal aspecto foi identificado por meio das análises das matrizes de confusão, conforme exemplificado na Tabela 6.9. Tal comportamento prevaleceu nos demais experimentos para todas as frequências, especialmente para a taxa de 256 Hz que teve o pior desempenho geral para LT2 comparada as demais taxas.

Tabela 6.9: Matriz de confusão do experimento nº 4 para a taxa de amostragem de 1 kHz

		<b>Predito</b>										
		<b>AG</b>	<b>BG</b>	<b>CG</b>	<b>AB</b>	<b>AC</b>	<b>BC</b>	<b>ABG</b>	<b>ACG</b>	<b>BCG</b>	<b>ABC</b>	
<b>Real</b>	<b>AG</b>	4080	0	0	0	0	0	8	1	0	11	4100
	<b>BG</b>	0	4084	0	0	0	0	0	0	3	13	4100
	<b>CG</b>	0	0	4093	0	0	0	0	1	0	6	4100
	<b>AB</b>	0	0	0	3898	0	0	202	0	0	0	4100
	<b>AC</b>	0	0	0	0	3844	0	0	256	0	0	4100
	<b>BC</b>	0	0	0	0	0	3828	0	0	272	0	4100
	<b>ABG</b>	0	0	0	138	0	0	3962	0	0	0	4100
	<b>ACG</b>	0	0	0	0	169	0	0	3927	0	4	4100
	<b>BCG</b>	0	0	0	0	0	386	0	0	3711	3	4100
	<b>ABC</b>	0	0	0	0	0	0	0	0	0	4100	4100

Essa circunstância representou a necessidade de inclusão de uma característica que auxiliasse o algoritmo a distinguir especialmente entre as faltas LL e LLG, e vice-versa. Para resolver esse problema, nós adicionamos o atributo *groundDetection* e, conseqüentemente, o modelo atingiu os desempenhos reportados na Tabela 6.10. Cabe ressaltar que todas essas análises preliminares do *ablation study* foram realizadas exclusivamente a partir dos conjuntos de treinamento e validação.

Tabela 6.10: Média e desvio padrão da acurácia para a classificação do tipo de falha com a etapa de normalização e com a inclusão do atributo *groundDetection*

Taxa de amostragem	LT1	LT2
<b>256 Hz</b>	99,9960% ± 0,0046%	99,5512% ± 0,0797%
<b>1 kHz</b>	99,9998% ± 0,0008%	99,8866% ± 0,0285%
<b>5 kHz</b>	99,9995% ± 0,0015%	99,9368% ± 0,0208%
<b>10 kHz</b>	99,9998% ± 0,0008%	99,9612% ± 0,0142%

✓ Normalização  
 ✓ Atributos máximos e mínimos  
 ✓ Atributo *groundDetection*  
 ✗ Atributos baseados na energia

A Equação 5.1, utilizada para a obtenção do atributo *groundDetection* conforme relatado na Subseção 5.1.1, baseia-se no componente da corrente de sequência-zero, o qual é um indicador da presença da terra na falha (vide Subseção 2.1.1). Este componente apresenta valores substancialmente altos durante a presença da terra na falta. Akke e Thorp (1998) originalmente definiram o valor do limiar (*threshold*) como 0,05 a partir da observação dos valores dos índices obtidos a partir de diferentes condições de falha. Entretanto, os autores utilizaram os dados gerados sinteticamente (simulados) para a definição deste limiar. Por conseguinte, atualizamos o valor desse limiar para 0,1 fundamentados pela análise de um subconjunto de falhas reais coletadas da rede da COPEL, conforme mencionado anteriormente.

Todos os experimentos reportados até então nesta subseção não consideraram os atributos baseados na energia (*energiaPropA*, *energiaPropB* e *energiaPropC*). De fato, o método manifestou comportamento acurado e preciso tanto ao ser avaliado com dados provenientes da mesma linha de transmissão para a qual foi treinado (LT1) quanto ao ser avaliado com dados de uma linha desconhecida ao algoritmo (LT2), mesmo sem utilizar os atributos de energia. Em outros termos, os resultados apresentados até então não demonstram de maneira efetiva a real necessidade da inclusão de tais características ao método, uma vez que os desempenhos já exprimem taxas de acerto superiores à 99%.

Nas Tabelas 6.11 a 6.14 estão retratados os desempenhos do método proposto para o conjunto TL2 com a inclusão dos atributos *energiaPropA*, *energiaPropB* e *energiaPropC*, separadamente para cada frequência. Os resultados para os conjuntos de validação e teste da LT1 foram anteriormente apresentados nas Tabelas 6.2 a 6.6. Ao compararmos estes valores com aqueles previamente apresentados na Tabela 6.10, constata-se pouca influência no desempenho dos cenários analisados, ou seja, sem a clara demonstração da necessidade destas novas características. Todavia, a importância destes atributos estará devidamente justificada na Subseção 6.1.4.

Tabela 6.11: Desempenho do método proposto para a taxa de amostragem de 256 Hz (conjunto LT2)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	99,9817% ± 0,0086%	99,9976% ± 0,0051%	99,9896% ± 0,0050%
BG	99,8720% ± 0,0723%	100,0000% ± 0,0000%	99,9359% ± 0,0362%
CG	99,5500% ± 0,1755%	99,9804% ± 0,0473%	99,7647% ± 0,0956%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,9890% ± 0,0107%	99,9945% ± 0,0053%
ACG	99,9768% ± 0,0499%	99,9927% ± 0,0063%	99,9847% ± 0,0262%
BCG	100,0000% ± 0,0000%	99,4243% ± 0,1681%	99,7113% ± 0,0846%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9380% ± 0,0187%	99,9384% ± 0,0185%	99,9381% ± 0,0186%

Tabela 6.12: Desempenho do método proposto para a taxa de amostragem de 1 kHz (conjunto LT2)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	98,8024% ± 0,2571%	100,0000% ± 0,0000%	99,3975% ± 0,1301%
BG	99,9354% ± 0,0377%	100,0000% ± 0,0000%	99,9677% ± 0,0189%
CG	99,9695% ± 0,0295%	100,0000% ± 0,0000%	99,9848% ± 0,0147%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,7604% ± 0,0657%	99,8800% ± 0,0330%
ACG	100,0000% ± 0,0000%	98,9592% ± 0,2785%	99,4767% ± 0,1407%
BCG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,8707% ± 0,0295%	99,8720% ± 0,0289%	99,8707% ± 0,0295%

Tabela 6.13: Desempenho do método proposto para a taxa de amostragem de 5 kHz (conjunto LT2)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	99,8939% ± 0,1060%	100,0000% ± 0,0000%	99,9469% ± 0,0531%
BG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
CG	99,9890% ± 0,0121%	100,0000% ± 0,0000%	99,9945% ± 0,0061%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,9427% ± 0,0504%	99,9714% ± 0,0052%
ACG	100,0000% ± 0,0000%	99,9403% ± 0,0731%	99,9701% ± 0,0366%
BCG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9883% ± 0,0108%	99,9883% ± 0,0108%	99,9883% ± 0,0109%



Tabela 6.14: Desempenho do método proposto para a taxa de amostragem de 10 kHz (conjunto LT2)

Tipo de falta	Acurácia	Precisão	F1-Score
AG	99,9354% ± 0,0702%	100,0000% ± 0,0000%	99,9677% ± 0,0351%
BG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
CG	99,9963% ± 0,0082%	100,0000% ± 0,0000%	99,9982% ± 0,0041%
AB	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
AC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
BC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABG	100,0000% ± 0,0000%	99,9452% ± 0,0748%	99,9726% ± 0,0375%
ACG	100,0000% ± 0,0000%	99,9866% ± 0,0211%	99,9933% ± 0,0105%
BCG	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
ABC	100,0000% ± 0,0000%	100,0000% ± 0,0000%	100,0000% ± 0,0000%
Total	99,9932% ± 0,0071%	99,9932% ± 0,0071%	99,9932% ± 0,0071%

Apesar de treinarmos o algoritmo para uma linha de transmissão específica (500 kV, 414 km, 60 Hz) e testá-la para uma outra linha distinta (230 kV, 106,646 km, 60 Hz), o modelo apresentou alta taxa de acerto com médias totais superiores à 99% para todas as taxas avaliadas. Esses resultados, os quais respondem a nossa QP1, comprovam a capacidade de generalização de nosso método para classificar corretamente faltas para LTs distintas daquela usada para o treinamento. Vale ressaltar que os casos de falha trifásicos também foram corretamente classificados com acurácia de 100% para todas as taxas de amostragem, similarmente ao reportado na Subseção 6.1.2.

A capacidade de generalização é essencial, visto que a LT sob supervisão pode apresentar eventuais alterações nas amplitudes dos sinais da corrente e da tensão durante sua operação, as quais não necessariamente são provenientes de falhas. Isso ocorre devido a possíveis fatores como variações contínuas de carga e desequilíbrio do fluxo de carga (Saha et al., 2010; Vasantharathna, 2016). Mesmo considerando uma mesma linha de transmissão, as amplitudes das correntes das oscilografias diferem entre os seus terminais. Além disso, as subestações podem supervisionar diversas LTs simultaneamente, cada qual com seu nível de tensão, reforçando ainda mais a importância da generalização de modo a evitar a necessidade do treinamento do algoritmo com dados para cada linha que será supervisionada, ou, ainda, requerer modelos treinados individualmente para cada linha.

Em complemento, na Figura 6.2 está apresentada uma sumarização dos resultados para todos os cenários avaliados até este ponto do trabalho, considerando todos os tipos de falta para ambos os terminais em conjunto para a taxa de amostragem de 10 kHz. Vale mencionar que o comportamento dos resultados é semelhante para as demais taxas de amostragem, conforme pode ser constatado pela análise das tabelas previamente apresentadas. Por meio desta análise gráfica fica evidente a importância destes procedimentos (normalização e atributo *groundDetection*) para a capacidade de generalização do método.



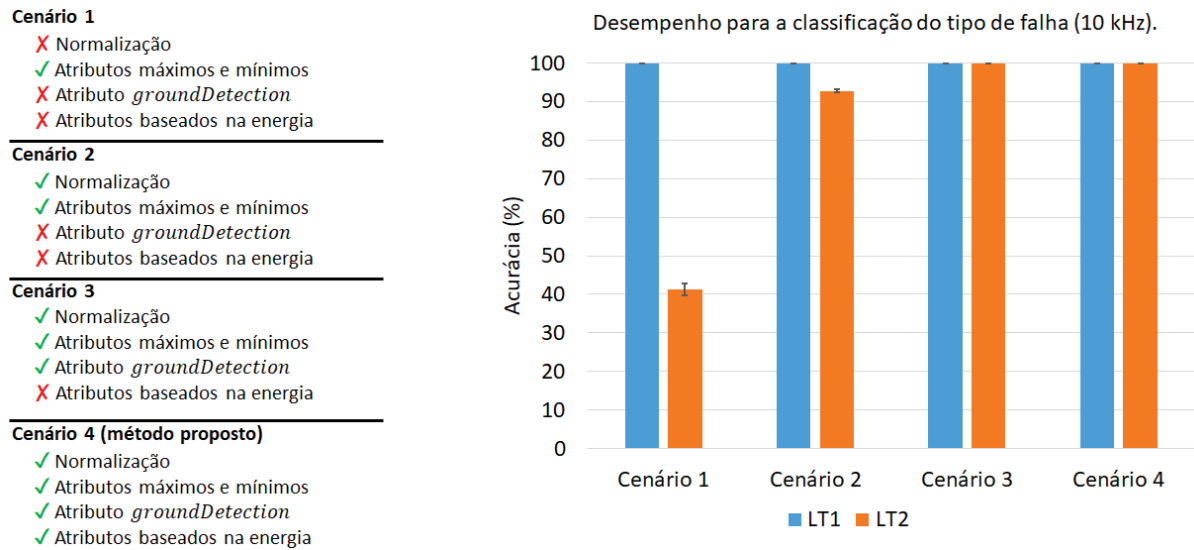


Figura 6.2: Sumarização dos resultados para os diversos cenários avaliativos realizados neste trabalho (*ablation study*) para a taxa de amostragem de 10 kHz

#### 6.1.4 Desempenho com Eventos Reais de Falta

Na Tabela 6.15 estão apresentados os resultados para a classificação de eventos reais de falta coletados da rede da COPEL (vide Tabela 6.1) para dois cenários distintos:

- **Cenário 1:** não utilizamos os atributos baseados na energia (*energiaPropA*, *energiaPropB* e *energiaPropC*);
- **Cenário 2:** utilizamos todos os atributos do método proposto.

A etapa de normalização foi efetuada para ambos os cenários, conforme justificada na subseção anterior. Visto que alguns dos eventos de falha possuem dados disponíveis apenas para uma das extremidades da linha, identificamos os terminais cujos dados não estão disponíveis por um traço (–). Para cada caso, o algoritmo foi treinado e avaliado dez vezes, seguindo os mesmos procedimentos adotados nos experimentos anteriores.

Ao compararmos ambos os cenários, fica evidente a importância da inclusão dos atributos baseados na energia (*energiaPropA*, *energiaPropB* e *energiaPropC*). Sem o uso destas características (cenário 1), o algoritmo classifica erroneamente alguns eventos LG como LLG. Exemplo disso ocorre para o evento #7, cujo tipo esperado era CG, mas foi identificado como ACG ou BCG ao longo dos experimentos para ambas as extremidades da linha. A Figura 6.3 retrata graficamente o sinal da corrente do evento #7 para ambos os terminais. Ao analisarmos o comportamento pós-falta, percebe-se uma alteração acentuada para a fase C, a qual foi efetivamente afetada pela falha. Entretanto, mesmo que as fases A e B não tenham sido envolvidas pela falha, também tiveram mudanças em suas amplitudes em comparação aos respectivos períodos pré-falta. Cabe salientar que, apesar deste comportamento também estar presente em alguns eventos simulados, os aumentos nas respectivas amplitudes não eram tão intensos a ponto de refletir no desempenho do algoritmo. Como consequência, o algoritmo possivelmente identificou esses comportamentos como faltosos, atribuindo assim alternativamente de maneira incorreta as fases A e B junto a fase C na falha com atuação da terra (G).

Tabela 6.15: Acurácia para a classificação do tipo de falta para eventos reais

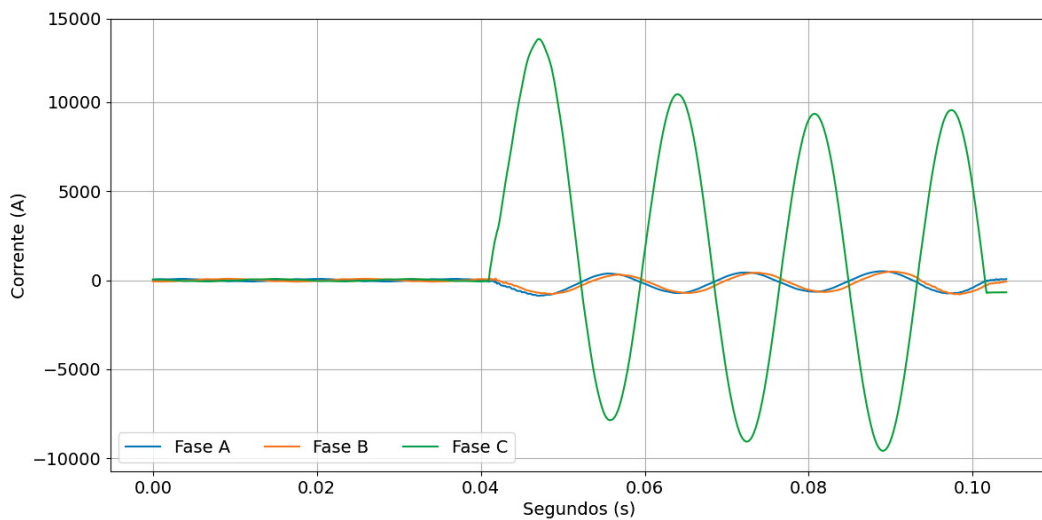
ID*	Tipo da falha	Acurácia			
		Cenário 1 <sup>1</sup>		Cenário 2 <sup>2</sup>	
		Terminal 1	Terminal 2	Terminal 1	Terminal 2
#1	AG	100%	100%	100%	100%
#2	BG	100%	100%	100%	100%
#3	CG	100%	100%	100%	100%
#4	AG	100%	100%	100%	100%
#5	AG	100%	100%	100%	100%
#6	AG	100%	–	100%	–
#7	CG	0%	0%	100%	100%
#8	AG	100%	100%	100%	100%
#9	CG	100%	100%	100%	100%
#10	BCG	100%	–	100%	–
#11	CG	–	100%	–	100%
#12	CG	100%	100%	100%	100%
#13	AG	100%	100%	100%	100%
#14	BG	0%	0%	100%	100%
#15	CG	–	100%	–	100%
#16	CG	100%	20%	100%	100%
#17	ABC	–	100%	–	100%
#18	AG	100%	80%	100%	100%
#19	CG	100%	100%	100%	100%
#20	AG	–	50%	–	0%
#21	ABC	–	0%	–	0%
#22	ACG	100%	100%	100%	100%
#23	CG	100%	80%	100%	100%
#24	CG	100%	100%	100%	100%
#25	CG	100%	100%	100%	100%
#26	BG	100%	100%	100%	100%
#27	AG	90%	100%	100%	100%
#28	AG	100%	100%	100%	100%
#29	AG	50%	100%	80%	100%
#30	AG	100%	100%	70%	100%
#31	BC	100%	100%	100%	100%
#32	BG	–	100%	–	100%
#33	BG	100%	70%	100%	100%
#34	BG	100%	100%	100%	100%
#35	AG	50%	100%	100%	100%
<b>Acurácia Geral</b>		86,9355%		95,9677%	

\*Confira a Tabela 6.1 para verificar os detalhes sobre cada evento

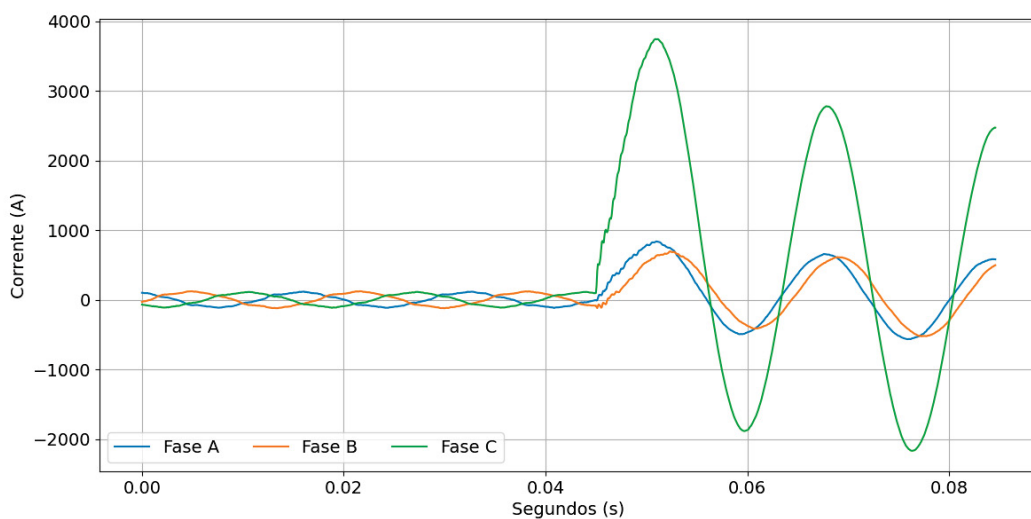
<sup>1</sup>Cenário 1: método sem os atributos baseados na energia

<sup>2</sup>Cenário 2: método com todos os atributos listados na Subseção 5.1.1

Esse comportamento demonstra a limitação de dados simulados, utilizados neste trabalho para o treinamento do algoritmo, frente a dados reais. Em outros termos, dados sintéticos não conseguem representar por completo toda a complexidade em que as linhas de transmissão estão suscetíveis e que, conseqüentemente, resultam numa grande diversidade de assinaturas de faltas devido a uma combinação de diferentes fatores, tanto ambientais quanto operacionais. Uma possível solução seria utilizar dados emulados no lugar de dados simulados, uma vez que a emulação demonstra um comportamento muito mais fidedigno e preciso do que a simulação. Todavia, a emulação de falhas demanda de equipamentos especiais, como o equipamento SMART GRID DL SGWD-W<sup>1</sup> da De Lorenzo.



(a) Terminal 1



(b) Terminal 2

Figura 6.3: Sinal trifásico da corrente para o evento de falha #7

<sup>1</sup><https://delorenzogloball.com/product/smart-grid-dl-sgwd-w-1?Lang=En>

Diante do exposto, identificamos a necessidade de novos atributos que pudessem contribuir para que o algoritmo distinguísse mais precisamente quais fases realmente estão sob a influência ou não de uma falha, mesmo que apresentem aumento na amplitude da corrente comparado ao sinal pré-falha. A partir da inclusão dos atributos baseados na energia, apenas quatro dos 62 eventos não tiveram taxa de acerto de 100% (cenário 2), e tão somente dois foram classificados incorretamente para todos os dez experimentos. No quesito desempenho, a acurácia do método de 87% no cenário 1 passa a atingir cerca de 96% de acurácia para a classificação com o cenário 2, uma melhoria notável e significativa no desempenho.

Neste momento, com a ocorrência de erros de predição por parte do método proposto (cenário 2), demonstra-se a necessidade e a importância da saída probabilística do método com o *ranking* das três classes (tipos de falta) com maior probabilidade de representar os dados de entrada. Nas Tabelas 6.16 a 6.19 estão apresentadas as predições de cada experimento para os eventos de falha que não atingiram 100% de acerto pelo método proposto (#20, #21, #29 e #30, respectivamente). Nestas tabelas estão apresentadas as três classes com maiores probabilidades estabelecidas pelo algoritmo ERT, juntamente com os respectivos valores entre parênteses.

Apesar do algoritmo classificar erroneamente o tipo de falha para todos os experimentos para o evento #20 (Tabela 6.16), o tipo esperado está como o segundo com maior probabilidade para todos os experimentos. De modo similar ocorreu para os eventos #29 (Tabela 6.18) e #30 (Tabela 6.19), os quais quando foram incorretamente preditos, possuíam o tipo correto na segunda posição de maior probabilidade. Essas análises ressaltam e amparam nosso interesse de incluir ao método a apresentação não apenas da predição final, mas também um *ranking* com os principais tipos de falta determinados pelo algoritmo com as maiores chances de representar os dados de entrada, juntamente com os respectivos valores de probabilidade.

Tabela 6.16: Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #20 do terminal 2. Os valores entre parênteses representam as probabilidades

<b>Experimento</b>	<b>Rank 1 (predito)</b>	<b>Rank 2</b>	<b>Rank 3</b>
<b>#1</b>	ACG (60%)	AG (40%)	–
<b>#2</b>	ACG (56%)	AG (40%)	ABG/BCG (2%)
<b>#3</b>	ACG (66%)	AG (28%)	ABG (8%)
<b>#4</b>	ACG (62%)	AG (30%)	ABG (8%)
<b>#5</b>	ACG (68%)	AG (26%)	ABG (6%)
<b>#6</b>	ACG (68%)	AG (32%)	–
<b>#7</b>	ACG (64%)	AG (34%)	ABG (2%)
<b>#8</b>	ACG (60%)	AG (36%)	ABG (4%)
<b>#9</b>	ACG (52%)	AG (40%)	ABG (6%)
<b>#10</b>	ACG (64%)	AG (30%)	AB/ABC/ABG (2%)

Tabela 6.17: Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #21 do terminal 2. Os valores entre parênteses representam as probabilidades

<b>Experimento</b>	<b>Rank 1 (predito)</b>	<b>Rank 2</b>	<b>Rank 3</b>
<b>#1</b>	ACG (48%)	ABC/ABG (16%)	BCG (8%)
<b>#2</b>	ACG (42%)	ABG (26%)	BCG (18%)
<b>#3</b>	ABG (40%)	ACG (26%)	AG (16%)
<b>#4</b>	ACG (42%)	BCG (18%)	ABC (16%)
<b>#5</b>	ABG (34%)	BCG (28%)	ACG (26%)
<b>#6</b>	ACG (54%)	ABG (22%)	ABC (8%)
<b>#7</b>	ABG (40%)	ACG (34%)	ABC/BCG (10%)
<b>#8</b>	ACG (38%)	ABG (26%)	BCG (14%)
<b>#9</b>	ACG (46%)	ABG (24%)	BCG (20%)
<b>#10</b>	ACG (30%)	ABG (22%)	BCG/BG (18%)

Tabela 6.18: Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #29 do terminal 1. Os valores entre parênteses representam as probabilidades

<b>Experimento</b>	<b>Rank 1 (predito)</b>	<b>Rank 2</b>	<b>Rank 3</b>
<b>#1</b>	AG (54%)	ABG (46%)	–
<b>#2</b>	AG (60%)	ABG (40%)	–
<b>#3</b>	AG (70%)	ABG (30%)	–
<b>#4</b>	AG (52%)	ABG (46%)	AB (2%)
<b>#5</b>	AG (68%)	ABG (32%)	–
<b>#6</b>	AG (60%)	ABG (40%)	–
<b>#7</b>	AG (62%)	ABG (38%)	–
<b>#8</b>	ABG (54%)	AG (44%)	BG (2%)
<b>#9</b>	AG (52%)	ABG (48%)	–
<b>#10</b>	ABG (54%)	AG (46%)	–

Tabela 6.19: Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #30 do terminal 1. Os valores entre parênteses representam as probabilidades

<b>Experimento</b>	<b>Rank 1 (predito)</b>	<b>Rank 2</b>	<b>Rank 3</b>
<b>#1</b>	ABG (60%)	AG (40%)	–
<b>#2</b>	ABG (54%)	AG (46%)	–
<b>#3</b>	AG (52%)	ABG (48%)	–
<b>#4</b>	AG (60%)	ABG (38%)	AB (2%)
<b>#5</b>	AG (56%)	ABG (42%)	ACG (2%)
<b>#6</b>	AG (60%)	ABG (40%)	–
<b>#7</b>	AG (56%)	ABG (42%)	AB (2%)
<b>#8</b>	AG (60%)	ABG (40%)	–
<b>#9</b>	AG (62%)	ABG (38%)	–
<b>#10</b>	ABG (54%)	AG (46%)	–

Por outro lado, o evento #21 foi incorretamente classificado de maneira generalizada, pouco aparecendo no *ranking* das maiores probabilidades (Tabela 6.17). Esse comportamento ocorreu devido ao reconhecimento errôneo da incidência da terra na falha, ou seja, atribuindo ao atributo *groundDetection* o valor um (1). De modo a justificar essa hipótese, refizemos as dez repetições dos experimentos fixando o valor do atributo *groundDetection* como zero (0), ou seja, indicando que não houve atuação da terra. Os resultados para esta avaliação estão apresentados na Tabela 6.20, a qual demonstra claramente a influência do atributo *groundDetection* na predição do algoritmo. Com o valor corretamente delegado à esta característica, percebe-se a alta probabilidade atribuída pelo algoritmo para a classe (tipo de falta) correta: ABC.

Tabela 6.20: Ranking dos tipos de falha com maiores probabilidades determinadas pelo ERT para cada experimento para o evento de falha ID #21 do terminal 2 com o valor do atributo *groundDetection* estabelecido com o valor zero (0) – indicativo que não houve atuação da terra. Os valores entre parênteses representam as probabilidades

Experimento	Rank 1 (predito)	Rank 2	Rank 3
#1	ABC (96%)	AC (4%)	–
#2	ABC (92%)	AC (8%)	–
#3	ABC (98%)	AC (2%)	–
#4	ABC (96%)	AC (4%)	–
#5	ABC (92%)	AC (8%)	–
#6	ABC (98%)	AC (2%)	–
#7	ABC (98%)	AC (2%)	–
#8	ABC (100%)	–	–
#9	ABC (98%)	AC (2%)	–
#10	ABC (92%)	AC (8%)	–

Este aspecto representa a necessidade de melhorias na estratégia que determina o valor do atributo *groundDetection*. Neste cenário, possíveis soluções podem estar relacionadas ao desenvolvimento de técnicas mais eficientes para determinar a presença ou não da terra na falta, uma vez que a técnica utilizada neste trabalho (Equação 5.1) foi estabelecida originalmente a partir de dados simulados (Akke e Thorp, 1998). A tentativa de um novo ajuste no valor do limiar não é apropriado, podendo acarretar em um sobreajuste. É digno de nota que o valor do limiar foi definido neste trabalho considerando exemplos reais não tabelados pela COPEL, de modo que não compuseram o conjunto de eventos utilizados para testar o método (Tabela 6.1).

Conforme evidenciado nesta subseção e pela Tabela 6.1, as subestações podem supervisionar diversas linhas de transmissão simultaneamente. Neste aspecto, os métodos disponíveis no estado da arte necessitam predominantemente de treinamento individualizado com dados para cada uma destas linhas, além de eventualmente requisitarem de modelos individualizados para cada uma, demonstrando a incapacidade de classificar falhas para LTs desconhecidas, ou seja, linhas que não foram utilizadas para o treinamento dos algoritmos. Por outro lado, nosso método demonstrou a capacidade de classificar o tipo de falta mesmo para linhas de transmissão com configurações diferentes daquela usada para treinar o algoritmo. Logo, uma vez treinado, nosso modelo pode ser utilizado em subestações sem a necessidade de novos treinamentos com dados para cada linha sob supervisão.

Uma propriedade crucial que requer destaque, e que até então não havia sido constatada nas análises anteriores, consiste na capacidade do método em classificar o tipo da falha para uma oscilografia com taxa de amostragem diferente daquela usada para treinar o algoritmo. Enquanto os dados usados para treinar o algoritmo compreenderam a frequência de 10 kHz, todos os eventos analisados estavam em 15,36 kHz, sem a necessidade de realizar uma reamostragem nos dados do sinal antes da extração dos atributos. Portanto, os resultados reportados manifestam que, uma vez que o algoritmo esteja treinado, o método pode ser utilizado para a classificação de faltas independentemente do terminal analisado, da taxa de amostragem configurada no equipamento e das características da LT (*e.g.*, comprimento e tensão da linha).

Assim sendo, os resultados do método proposto para a classificação de tipos de falhas são satisfatórios. Apesar do algoritmo ter sido treinado unicamente com dados sintéticos, nosso método foi capaz de classificar corretamente o tipo de falta para casos reais, especialmente quando apoiado pela apresentação de um *ranking* de possíveis tipos de falta. Logo, esses resultados corroboram a QP2, visto que mesmo com o método sendo treinado apenas com dados simulados/sintéticos, ele foi eficaz na tarefa de classificação de eventos reais de falta. A diversidade de experimentos apresentados neste trabalho demonstra a versatilidade e a efetividade do nosso método.

#### 6.1.5 Análise da Importância dos Atributos

Os resultados alcançados nas subseções anteriores foram possíveis graças a engenharia de atributos realizada neste trabalho. Demonstramos a importância dessa engenharia a partir de experimentos abrangentes decorridos nas subseções anteriores, justificando a contribuição de cada característica para o método. Objetivando o complemento dessas análises sobre os atributos, apresentamos nesta subseção a avaliação para quantificarmos a importância de cada atributo para o algoritmo ERT utilizado neste trabalho para a classificação de faltas em linhas de transmissão. Para tal finalidade, empregamos a análise de importância de Gini (Ishwaran, 2015).

Na Tabela 6.21 e na Figura 6.4 estão apresentadas as importâncias dos atributos para o algoritmo ERT, em que quanto maior o valor, mais relevante é o atributo. As pontuações dispostas na tabela representam a média e o desvio padrão para cada característica ao longo de dez repetições. As relevâncias foram computadas a partir da importância de Gini com a faixa de valores normalizada (a soma dos valores equivale a 1).

Tabela 6.21: Importância dos atributos para o algoritmo ERT usado no método proposto para classificação de falhas

<b>Atributo</b>	<b>256 Hz</b>	<b>1 kHz</b>	<b>5 kHz</b>	<b>10 kHz</b>
<i>energiaPropA</i>	0,1235 ± 0,0040	0,1503 ± 0,0050	0,1616 ± 0,0028	0,1614 ± 0,0082
<i>energiaPropB</i>	0,1215 ± 0,0026	0,1494 ± 0,0029	0,1625 ± 0,0064	0,1679 ± 0,0072
<i>energiaPropC</i>	0,1201 ± 0,0034	0,1511 ± 0,0067	0,1627 ± 0,0084	0,1653 ± 0,0030
<i>maxA</i>	0,0544 ± 0,0017	0,0586 ± 0,0045	0,0578 ± 0,0048	0,0556 ± 0,0060
<i>maxB</i>	0,0578 ± 0,0028	0,0580 ± 0,0064	0,0546 ± 0,0055	0,0526 ± 0,0057
<i>maxC</i>	0,0669 ± 0,0025	0,0576 ± 0,0041	0,0540 ± 0,0037	0,0563 ± 0,0056
<i>minA</i>	0,1018 ± 0,0025	0,0671 ± 0,0059	0,0569 ± 0,0021	0,0543 ± 0,0035
<i>minB</i>	0,1034 ± 0,0037	0,0667 ± 0,0050	0,0542 ± 0,0047	0,0569 ± 0,0050
<i>minC</i>	0,0966 ± 0,0030	0,0673 ± 0,0063	0,0580 ± 0,0049	0,0562 ± 0,0057
<i>groundDetection</i>	0,1541 ± 0,0055	0,1740 ± 0,0105	0,1777 ± 0,0066	0,1735 ± 0,0050



A característica *groundDetection* manifestou a maior contribuição para as predições. Este aspecto pode ser explicado pela análise conduzida ao longo da Seção 6.1.3, a qual revelou que o atributo *groundDetection* é crucial ao algoritmo ERT para distinguir entre os tipos de falhas considerando o envolvimento ou não da terra na falta, *e.g.*, discernindo entre AB e ABG.

Essa comprovação reforça a análise conduzida na subseção anterior, onde debatemos sobre os erros do método para o evento de falha #21, os quais ocorreram justamente pela atribuição incorreta do valor ao atributo em questão. Portanto, o aperfeiçoamento na técnica para determinar o valor do atributo *groundDetection* é fundamental, uma vez que possui uma influência notória ao algoritmo ERT na realização das predições.

Em seguida na ordem de relevância estão os atributos baseados na energia, cujos valores tendem a ter maior peso para as maiores taxas de amostragem, alcançando valores superiores a 0,16 para as duas maiores frequências (5 kHz e 10 kHz). Um comportamento interessante de ser observado, porém já esperado, corresponde a similaridade dos valores das pontuações entre as características *energiaPropA*, *energiaPropB* e *energiaPropC* dentro de uma mesma taxa de amostragem. Isso pode ser explicado pelo motivo de que esses atributos baseiam-se na mesma informação (energia), cada uma representando sua respectiva fase, mas principalmente é justificado pelo fato da base de dados ser balanceada entre os tipos de falta, evitando que as características de uma determinada fase prevaleçam na relevância sobre as demais. Tal comportamento também ocorre para os atributos de máximos e mínimos.

Em contrapartida, o restante dos atributos de máximos e mínimos demonstraram, em geral, pontuações similares, salvo para a taxa de amostragem de 256 Hz que atestou maior importância para os atributos de mínimos em comparação aos máximos. Este comportamento era aguardado, visto que cada par de atributos (*e.g.*, *maxA* e *minA*) visava contribuir de maneira igualitária para identificar variações nas amplitudes individualmente para cada fase, permitindo que o algoritmo reconhecesse se a falha afetou uma fase em particular, ou seja, distinguir sobre o envolvimento ou não da fase na falha.

Do mesmo modo, os atributos de energia possuem papéis semelhantes as características de máximos e mínimos. As fases que estão envolvidas em uma falta demonstram valores de energia substancialmente altos, e, conseqüentemente, revelam uma maior contribuição para a energia total. Logo, a principal função desses atributos é distinguir em relação a presença ou não de cada fase em uma falha em particular, *e.g.*, distinguir entre AG e ABG. O uso dessas características baseadas na energia diverge de outros trabalhos na literatura que também utilizam valores de energia, os quais aplicam os valores de energia diretamente como atributos de seus métodos. Por outro lado, nosso método calcula a proporção da contribuição de energia. Esse uso distinto que propomos visa potencializar e reforçar ao algoritmo sobre a presença de uma fase na falha em questão.

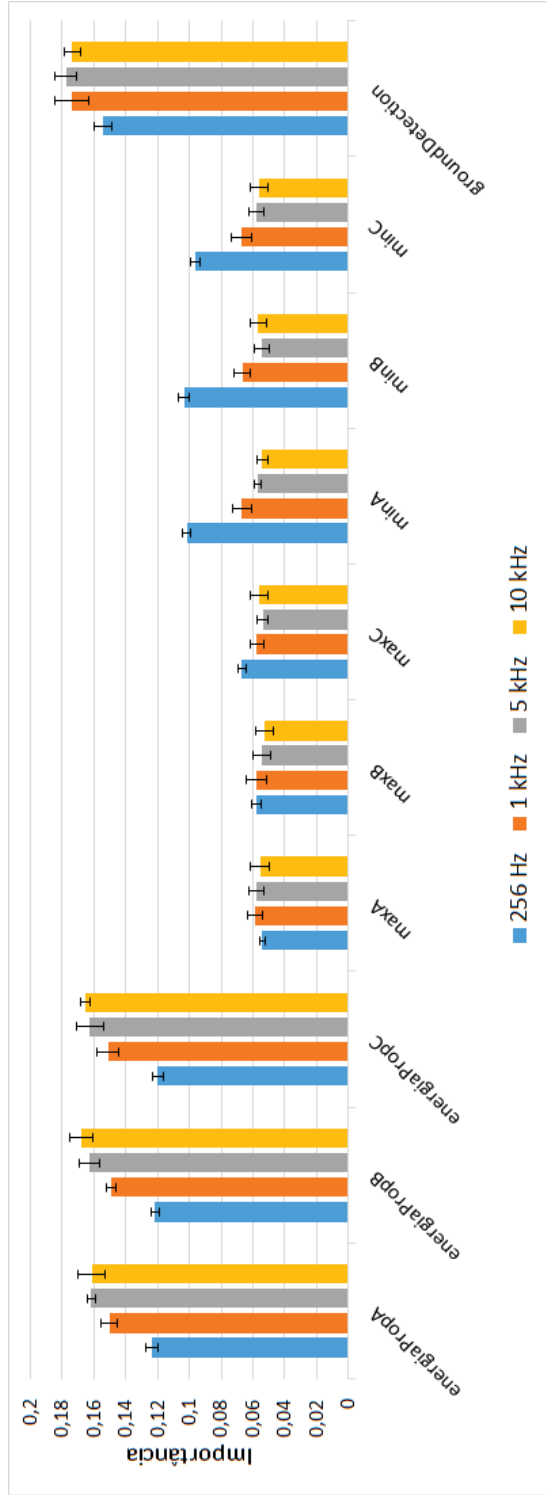


Figura 6.4: Importância dos atributos para o algoritmo ERT

Fonte: autoria própria

### 6.1.6 Comparação com o Estado da Arte

Inicialmente nesta subseção, esclarecemos o motivo pela escolha do algoritmo ERT como classificador ao compararmos seu desempenho frente a outros algoritmos de modo a corroborar nossa hipótese (Seção 1.3). Nas Tabelas 6.22 e 6.23 estão retratados os desempenhos dos algoritmos de AM avaliados para os conjuntos LT1 e LT2, respectivamente. Cabe mencionar que a arquitetura usada na MLP baseia-se na estrutura proposta por Ferreira et al. (2020), enquanto os demais algoritmos seguem as configurações padrão das bibliotecas (vide Seção 5.3), salvo os algoritmos *ensemble* cujos números de estimadores foram estabelecidos em 50.

Tabela 6.22: Média e desvio padrão da acurácia para a classificação do tipo de falta para os algoritmos avaliados (conjunto de validação LT1)

Algoritmo	256 Hz	1 kHz	5 kHz	10 kHz
<b>CART</b>	99,8710 ± 0,0145	99,9342 ± 0,0117	99,9560 ± 0,0100	99,9556 ± 0,0108
<b>SVM</b>	92,7387 ± 0,0943	95,3323 ± 0,0765	98,1844 ± 0,0467	98,2289 ± 0,0663
<b>KNN (k=3)</b>	99,1793 ± 0,0325	99,0165 ± 0,0423	99,3639 ± 0,0349	99,4255 ± 0,0232
<b>MLP</b>	99,9235 ± 0,1344	99,9575 ± 0,0316	99,9742 ± 0,0219	99,9451 ± 0,0627
<b>LightGBM</b>	99,9827 ± 0,0042	99,9892 ± 0,0045	99,9948 ± 0,0015	99,9954 ± 0,0023
<b>XGBoost</b>	99,9765 ± 0,0063	99,9911 ± 0,0041	99,9938 ± 0,0036	99,9968 ± 0,0025
<b>RF</b>	99,9923 ± 0,0024	99,9965 ± 0,0023	99,9970 ± 0,0023	99,9983 ± 0,0020
<b>ERT</b>	99,9980 ± 0,0024	99,9993 ± 0,0010	99,9989 ± 0,0012	99,9992 ± 0,0013

Tabela 6.23: Média e desvio padrão da acurácia para a classificação do tipo de falta para os algoritmos avaliados (conjunto LT2)

Algoritmo	256 Hz	1 kHz	5 kHz	10 kHz
<b>CART</b>	96,1917 ± 1,0248	99,3355 ± 0,1330	99,5805 ± 0,1286	99,6022 ± 0,1353
<b>SVM</b>	98,0288 ± 0,0275	98,0104 ± 0,0463	98,6900 ± 0,0571	99,0361 ± 0,0298
<b>KNN (k=3)</b>	91,3960 ± 0,2957	91,1077 ± 0,3384	93,1662 ± 0,4355	92,9883 ± 0,4160
<b>MLP</b>	99,6117 ± 0,1946	99,4246 ± 0,2436	99,6209 ± 0,5452	99,8421 ± 0,3187
<b>LightGBM</b>	94,3415 ± 0,4272	99,2090 ± 0,1158	99,1896 ± 0,1601	99,5506 ± 0,1350
<b>XGBoost</b>	98,7548 ± 0,2584	99,3878 ± 0,0980	99,7382 ± 0,0519	99,8138 ± 0,0224
<b>RF</b>	99,5093 ± 0,1314	99,5779 ± 0,0332	99,8920 ± 0,0167	99,9252 ± 0,0237
<b>ERT</b>	99,9380 ± 0,0187	99,8707 ± 0,0295	99,9883 ± 0,0108	99,9932 ± 0,0071

Ao compararmos os desempenhos entre os algoritmos pelo teste de Friedman, separadamente por taxa de amostragem e individualmente para LT1 e LT2, nota-se diferença estatisticamente significativa para todos os cenários examinados ( $p$ -valor < 0,0001 tanto para LT1 quanto para LT2). A principal razão pela diferença estatística para a LT1 está associada ao desempenho inferior do SVM. As maiores discrepâncias estão nos resultados para a LT2 (Tabela 6.23) em que avaliamos a competência para a generalização, contexto o qual evidencia-se o predomínio do algoritmo ERT sobre os demais, consolidada a partir do pós-teste de Dunn, principalmente para a menor frequência (256 Hz).

A superioridade geral do ERT frente aos demais algoritmos, em especial sobre os monolíticos, pode estar atrelada ao fato desta técnica ser composta por múltiplos especialistas (Subseção 2.2.3). Logo, a combinação de um grupo de estimadores com erros não correlacionados pode identificar padrões distintos no conjunto de características, o que pode resultar em um modelo mais preciso para um problema específico, como foi o caso da classificação de falhas.

Também comparamos o desempenho de nosso método diante de dois trabalhos relacionados previamente descritos na Seção 3.1 (Ferreira et al., 2020; Coban e Tezcan, 2021). A escolha por ambos os trabalhos se deve a dois fatores principais: (i) ambos os trabalhos, assim como o nosso, não dependem de processamento de sinais, como as *wavelets*; (ii) natureza das características utilizadas pelos respectivos métodos. Enquanto Ferreira et al. (2020) fornece como entrada atributos extraídos dos sinais, Coban e Tezcan (2021) utilizam os dados brutos diretamente como entrada. Nas Tabelas 6.24 e 6.25 estão apresentados os resultados comparativos para os conjuntos de teste LT1 e LT2, respectivamente, considerando ambos os terminais em conjunto. Vale mencionar que replicamos ambos os métodos utilizando o mesmo protocolo experimental e os mesmos conjuntos de dados (LT1 e LT2) empregados para avaliar nosso método.

Tabela 6.24: Comparação dos desempenhos pela acurácia para a classificação de falta para o conjunto de teste LT1

Taxa de amostragem	Método proposto	Ferreira et al. (2020)	Coban e Tezcan (2021)
<b>256 Hz</b>	99,9985%	93,9921%	99,9040%
	$\pm 0,0010\%$	$\pm 0,8810\%$	$\pm 0,0099\%$
<b>1 kHz</b>	99,9995%	94,4897%	99,9556%
	$\pm 0,0006\%$	$\pm 1,5213\%$	$\pm 0,0068\%$
<b>5 kHz</b>	99,9994%	95,5710%	99,9377%
	$\pm 0,0008\%$	$\pm 0,7403\%$	$\pm 0,0025\%$
<b>10 kHz</b>	99,9995%	96,6317%	99,9357%
	$\pm 0,0008\%$	$\pm 0,3018\%$	$\pm 0,0021\%$

Tabela 6.25: Comparação dos desempenhos pela acurácia para a classificação de falta para o conjunto de teste LT2

Taxa de amostragem	Método proposto	Ferreira et al. (2020)	Coban e Tezcan (2021)
<b>256 Hz</b>	99,9380%	79,3809%	67,4715%
	$\pm 0,0187\%$	$\pm 2,9962\%$	$\pm 0,7365\%$
<b>1 kHz</b>	99,8707 %	78,7740%	63,3211%
	$\pm 0,0295\%$	$\pm 2,5683\%$	$\pm 0,8601\%$
<b>5 kHz</b>	99,9883%	83,3813%	61,9565%
	$\pm 0,0108\%$	$\pm 2,2383\%$	$\pm 1,2086\%$
<b>10 kHz</b>	99,9932%	83,5744%	61,7146%
	$\pm 0,0071\%$	$\pm 2,4191\%$	$\pm 1,1788\%$

O método de Coban e Tezcan (2021) demonstrou alta taxa de acerto mesmo para as menores frequências de amostragem, quando treinado e testado com dados provenientes da mesma linha de transmissão (acurácia média superior a 99,9040%). Contudo, essa abordagem manifestou uma perda considerável de desempenho quando avaliada com dados de uma linha de transmissão distinta daquela usada para treinar o algoritmo (acurácia média inferior a 67,4715%). Por sua vez, o método elaborado por Ferreira et al. (2020) apresentou dependência em relação a taxa de amostragem para a primeira análise (Tabela 6.24), bem como resultados piores para a capacidade de generalização (Tabela 6.25), com acurácia média inferior a 83,5744%.

Em contrapartida, nosso método demonstrou desempenho superior comparado a ambos os trabalhos relacionados para todas as análises, especialmente ao considerarmos a competência da abordagem para classificar falhas para linhas de transmissão diferentes, com acurácia média superior a 99%. Essa melhor atuação foi atestada pelo teste estatístico de Wilcoxon pareado com intervalo de confiança de 95%. Os testes de hipótese revelaram diferentes estatisticamente significativa entre os desempenhos.

Esse aspecto representa que nossa proposta pode corretamente determinar o tipo de falta mesmo para linhas de transmissão desconhecidas, *i.e.*, aquelas que não foram utilizadas para treinar o algoritmo. Essa análise é crucial em razão das linhas de transmissão sob supervisão poderem apresentar disparidades nas magnitudes/níveis de tensão e corrente durante suas operações devido a diversos fatores, como a variação na potência dos geradores em uma situação menos extrema em comparação ao contexto considerado neste trabalho.

Realizamos ainda experimentos adicionais para esclarecer o motivo de nosso método funcionar melhor que os trabalhos de Coban e Tezcan (2021) e Ferreira et al. (2020) para a generalização. Neste cenário, o melhor desempenho pode ser explicado em parte pelo processo de normalização executado pelo nosso método, descrito detalhadamente no Subseção 5.1.1 e justificado pela Figura 6.1. Assim sendo, para avaliarmos o impacto da normalização para ambos os trabalhos, nós incluímos este processo em seus respectivos métodos, de modo semelhante ao que já é realizado para a nossa abordagem. Logo, normalizamos os sinais dos conjuntos de treinamento (LT1) e teste (LT2) previamente a extração de atributos, incluindo a forma de onda da tensão dado que Ferreira et al. (2020) também empregam este sinal junto ao da corrente. Para este caso em específico, os valores de máximo e mínimo usados na Equação 5.2 corresponderam ao ciclo pré-falta da tensão ao invés da corrente.

Na Tabela 6.26 estão apresentados os desempenhos com a inclusão da etapa de normalização para ambos os trabalhos, indicando uma melhora expressiva para o método de Coban e Tezcan (2021) para todas as taxas de amostragem em comparação a Tabela 6.25, enquanto Ferreira et al. (2020) demonstrou melhores resultados gerais apenas para as taxas de amostragem menores (256 Hz e 1 kHz), mas perca de desempenho para as taxas maiores (5 kHz e 10 kHz). Nosso método desempenhou melhor com diferença estatisticamente significativa para todas as taxas de amostragem ( $p$ -valor  $< 0,0001$  para todas as comparações por meio do teste de Wilcoxon com intervalo de confiança de 95%).

Esses resultados demonstram que apenas o processo de normalização é insuficiente para permitir que um método generalize a classificação do tipo de falta para linhas de transmissão desconhecidas, ou seja, linhas que não foram utilizadas para treinar o algoritmo. Essa análise reforça a importância de nossa engenharia de atributos, considerando cada etapa descrita na Subseção 5.1.1, juntamente com o uso do algoritmo ERT utilizado em nosso método.

Tabela 6.26: Comparação dos desempenhos para classificação de falta incluindo o processo de normalização para a capacidade de generalização (conjunto LT2)

Taxa de amostragem	Método proposto	Ferreira et al. (2020)	Coban e Tezcan (2021)
<b>256 Hz</b>	99,9380%	82,0427%	93,0045%
	$\pm 0,0187\%$	$\pm 0,6620\%$	$\pm 0,1715\%$
<b>1 kHz</b>	99,8707%	81,1492%	90,1602%
	$\pm 0,0295\%$	$\pm 0,5696\%$	$\pm 0,7773\%$
<b>5 kHz</b>	99,9883%	80,2492%	88,2118%
	$\pm 0,0108\%$	$\pm 1,5059\%$	$\pm 0,9141\%$
<b>10 kHz</b>	99,9932%	80,0675%	87,4419%
	$\pm 0,0071\%$	$\pm 1,0013\%$	$\pm 0,4137\%$

Vale lembrar que a configuração da rede MLP avaliada nas Tabelas 6.22 e 6.23 corresponde a rede proposta por Ferreira et al. (2020), cujos atributos usados naqueles experimentos refletem as características que propomos. Logo, se compararmos esses resultados com os apresentados nas Tabelas 6.24 e 6.25, os quais retratam os atributos originalmente utilizados pelos autores, nota-se o melhor desempenho de seu método ao empregar nossos atributos, especialmente para a capacidade de generalização. Essa constatação reforça ainda mais a importância da engenharia de atributos que fizemos.

## 6.2 LOCALIZAÇÃO DE FALTA

Esta seção divide-se em seis partes com experimentos abrangentes de avaliação de nosso método, nas quais responderemos as seguintes questões de pesquisa:

- **QP3:** a normalização incorporada leva a uma melhor definição da RC e, consequentemente, a melhores desempenhos?
- **QP4:** o módulo de filtro, agindo como uma etapa de pós-processamento, permite a remoção de predições imprecisas resultando na melhoria do sistema preditor?
- **QP5:** a técnica SDR que propomos permite atingir melhores resultados para a localização de falta comparada tanto a modelos *ensemble* convencionais quanto a modelos monolíticos tradicionais?
- **QP6:** de modo semelhante a QP2, o método proposto é capaz de generalizar seu funcionamento para LTs desconhecidas (*i.e.*, linhas não usadas para o treinamento do modelo) considerando a tarefa de localização?

Na Subseção 6.2.1 descrevemos o protocolo experimental usado para avaliar nosso método. Na Subseção 6.2.2 apresentamos o desempenho para o cenário que o algoritmo foi treinado e testado com dados da mesma LT, enquanto na Subseção 6.2.3 relatamos a importância dos atributos para o método localizar as falhas. Na Subseção 6.2.4 analisamos o desempenho de diversos algoritmos de AM (CART, SVM, KNN, MLP, LightGBM, XGBoost, *Bagging* de CART, RF e ERT) e comparamos o desempenho do nosso método frente a métodos dispostos no estado da arte. Já na Subseção 6.2.5 analisamos a capacidade do método localizar falhas para uma LT distinta daquela empregada para o seu treinamento. Para concluir esta seção, discutimos na Subseção 6.2.6 a complexidade da estrutura SDR proposta.



### 6.2.1 Protocolo Experimental

Em nosso estudo, treinamos o algoritmo ERT com  $N = 700$  estimadores e usamos as suas árvores para comporem nosso conjunto (*pool*) de regressores  $\mathcal{F}$ . Empregamos o ERT com 700 árvores pelo motivo de experimentos preliminares manifestarem que o conjunto com menor número de árvores ( $< 700$ ) resulta em piores resultados, enquanto um maior número de modelos ( $> 700$ ) não representa uma melhoria significativa no desempenho (Ensina et al., 2024).

O algoritmo foi treinado e examinado com exemplos embaralhados do conjunto de dados. Assim, as instâncias da FADbF para a LT de 500 kV foram divididas aleatoriamente em conjuntos de treinamento, validação e teste. O conjunto de treinamento  $\mathcal{T}$  compreende 50% de todos os exemplos (84.000 instâncias), enquanto os conjuntos de validação  $\mathcal{V}$  e teste  $\mathcal{X}$  contêm 25% de todos os eventos cada (42.000 exemplos). Esse processo de partição foi realizado dez vezes, resultando em diferentes conjuntos com instâncias distintas entre cada repetição. Logo, os resultados reportados na Seção 6.2 representam a média e o desvio padrão dessas dez repetições.

A escolha pelo treinamento em termos percentuais em vez de diretamente em quilômetros justifica-se pelo fato de as subestações poderem supervisionar diversas linhas de transmissão com diferentes comprimentos simultaneamente, o que poderia afetar as previsões do modelo. Portanto, a utilização de percentuais visa permitir que nosso método possa lidar com outras LTs independentemente de seus comprimentos, tornando o método mais flexível e tentando evitar a necessidade de novos treinamentos individualmente para cada linha sob supervisão.

Considerando o módulo “Normalização Incorporada” na segunda fase do SDR proposto (Subsubseção 5.2.2.2), aplicamos a técnica MinMax para dimensionar os atributos no intervalo  $[0, 1]$ . Ainda, estabelecemos  $K = 30$  para o módulo “Definição da RC” uma vez que apresentou melhores resultados que outros tamanhos de vizinhança em experimentos preliminares (Ensina et al., 2024). Analisamos quatro métricas de distância para avaliar o impacto da variação deste hiperparâmetro usado no algoritmo KNN para determinar a RC: Euclidiana, Cosseno, Geodésica e Cityblock. Esta avaliação é essencial para determinar se podemos obter melhores resultados variando a métrica de distância.

No módulo “Seleção Dinâmica”, a medida de competência utilizada foi o erro quadrático máximo proposto por Moura et al. (2019b) uma vez que esta métrica apresentou melhores resultados em experimentos preliminares do que as outras medidas listadas em seu estudo. Além disso, também avaliamos outras estratégias para selecionar os modelos mais competentes além da abordagem padrão mencionada em Subseção 2.2.4 (critérios de seleção). Portanto, avaliamos a mediana, a média aritmética, a média harmônica e a média geométrica juntamente com a técnica tradicional.

De modo similar ao descrito na Subseção 6.1.1, os dados da base também foram reamostrados para outras três taxas de amostragem, sendo considerados quatro cenários de avaliação: 256 Hz, 1 kHz, 5 kHz e 10 kHz. Inclusive, também usamos os dados da linha de 230 kV da FADbF para avaliarmos a capacidade de generalização do método, cujo treinamento ocorreu unicamente com os dados da LT de 500 kV, incluindo os dados de validação usados na estrutura SDR para a determinação da RC.

Todos os experimentos do método proposto foram avaliados pelas métricas do erro médio absoluto (*Mean Absolute Error* – MAE) e do erro quadrático médio (*Mean Squared Error* – MSE), definidas pelas Equações 6.4 e 6.5, respectivamente. A escolha por estas métricas justifica-se por elas serem as medidas usadas pela literatura para analisar o desempenho dos métodos de localização de faltas, principalmente o MAE.



$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6.4)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6.5)$$

sendo  $n$  o número de instâncias,  $y_i$  a localização real e  $\hat{y}_i$  a localização predita da falta para a  $i$ -ésima amostra.

### 6.2.2 Desempenho do Método Proposto

O desempenho do método proposto está apresentado na Tabela 6.27 para as quatro taxas de amostragem avaliadas, considerando o contexto em que o modelo foi treinado e testado com dados de uma mesma linha de transmissão (LT1 de 500 kV). Contempla-se pelos resultados uma melhor performance do método para as maiores frequências de amostragem, tanto para a métrica MSE quanto para MAE, corroborado pelo teste *one-way* ANOVA (p-valor < 0,0001). Em termos percentuais, por exemplo, o erro médio para a taxa de 10 kHz equivale a cerca de 0,1605% ± 0,0009% da extensão da linha de 500 kV de 414 km.

Tabela 6.27: Média e desvio padrão dos erros em km para a localização de falha (LT1)

Taxa de amostragem	MSE	MAE
<b>256 Hz</b>	2,2661 ± 0,0791	0,7599 ± 0,0079
<b>1 kHz</b>	2,0381 ± 0,0560	0,8249 ± 0,0061
<b>5 kHz</b>	1,3536 ± 0,0670	0,6612 ± 0,0032
<b>10 kHz</b>	1,3242 ± 0,0647	0,6646 ± 0,0040

Na Figura 6.5 está ilustrada a dispersão dos erros ao longo da linha de transmissão. Para esta análise, ordenamos o vetor correspondente as localizações reais do conjunto de teste, que varia de 4,14 km a 414 km, simultaneamente com o vetor das predições em ordem crescente, respeitando as posições dos valores entre ambos os vetores. O resultado desta ordenação é a curva contínua traçada na Figura 6.5. Podemos identificar que a distribuição dos erros ocorre, de modo geral, uniformemente ao longo da linha, com aumento dos erros sutilmente conforme a localização da falha se distancia do terminal usado como referência para a coleta dos dados.

Na Tabela 6.28 estão representados mais detalhes sobre a distribuição dos erros a partir de diferentes intervalos para a frequência de 10 kHz. Erros maiores que 10 km representam apenas cerca de 35 dos 42.000 eventos de falha dos conjuntos de teste, correspondendo por volta de 0,0833% das instâncias. Por outro lado, a maioria dos erros são menores ou iguais a 0,5 km e representam 57,2786% de todos os exemplos de teste (média de 24.057 das 42.000 ocorrências).

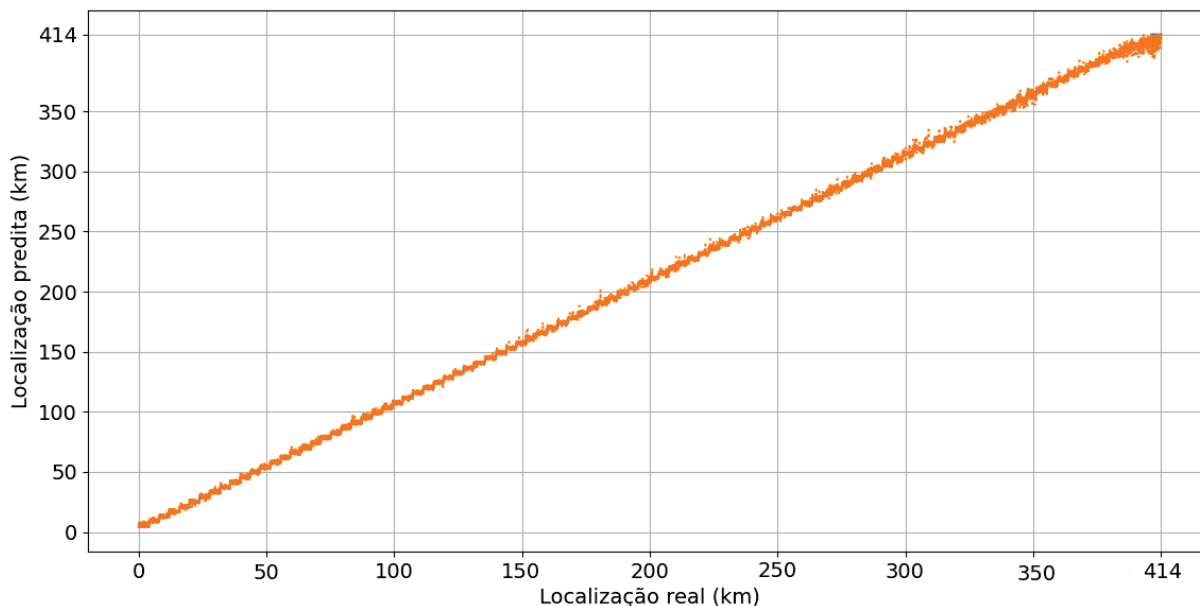


Figura 6.5: Representação dos erros de predição para a frequência de 10 kHz

Tabela 6.28: Distribuição dos erros para diferentes intervalos, considerando a frequência de 10 kHz. Os valores representam a média e o desvio padrão do número de ocorrências por intervalo ao longo das dez repetições

<b>Intervalo de erro</b>	<b>Números de ocorrências</b>
erro $\leq$ 0,5 km	24.057 $\pm$ 90
0,5 km < erro $\leq$ 1 km	10.085 $\pm$ 76
1 km < erro $\leq$ 5 km	7.523 $\pm$ 107
5 km < erro $\leq$ 10 km	300 $\pm$ 22
erro > 10 km	35 $\pm$ 8

Tais resultados puderam ser alcançados graças não somente pela diversidade e pela relevância dos atributos extraídos dos sinais de tensão e corrente, mas principalmente pela estrutura SDR proposta em nosso projeto. Em especial, nossa estrutura busca explorar a pluralidade dos modelos gerados durante o treinamento, selecionando aqueles ditos mais competentes individualmente para cada padrão de teste. Para isso, um dos processos fundamentais é a análise de seus hiperparâmetros, bem como dois módulos propostos nesta nova estrutura SDR: “Normalização Incorporada” e “Filtro”.

Exibimos a seguir, a partir de experimentos abrangentes, a importância tanto dos módulos anteriormente citados quanto dos hiperparâmetros do nosso método SDR que nos permitiram atingir os desempenhos acima reportados. Inicialmente, apresentamos na Subsubseção 6.2.2.1 os experimentos referentes ao processo de normalização, fundamentando a decisão pela normalização incorporada no esquema SDR proposto. Na Subsubseção 6.2.2.2 avaliamos o impacto no desempenho quando variamos a função de distância utilizada para determinar a RC. Na Subsubseção 6.2.2.3 analisamos os resultados referentes ao emprego de diferentes critérios para selecionar os regressores mais competentes do conjunto. Por fim, analisamos na Subsubseção 6.2.2.4 a relevância da aplicação do filtro como uma espécie de pós-processamento do nosso método. Cabe mencionar que utilizamos os mesmos modelos ERT pré-treinados ao longo das investigações, além de que tais experimentos compreenderam unicamente a taxa de amostragem de 10 kHz.

### 6.2.2.1 Avaliação do Processo de Normalização

Na Tabela 6.29 estão reportados os resultados das comparações entre três cenários de escalonamento:

1. **Sem normalização:** consiste no uso dos atributos sem qualquer pré-processamento (*i.e.*, sem o escalonamento dos dados);
2. **Normalização padrão:** consiste no escalonamento do conjunto de validação antes do início de todo o processo, utilizando seus parâmetros para normalizar cada padrão de teste durante a execução da SDR;
3. **Normalização incorporada:** proposto neste trabalho, este cenário foi descrito em detalhes na Subsubseção 5.2.2.2. Basicamente, este módulo normaliza o padrão de teste e o conjunto de validação concomitantemente e individualmente para cada novo exemplo para definir a RC. Antes de executarmos o processo de normalização, adicionamos o padrão de teste junto ao conjunto de validação para então conduzirmos a normalização, cujo padrão de teste escalonado é removido do conjunto de validação logo após a execução da normalização. Cabe destacar que seu uso refere-se única e exclusivamente para determinarmos a RC, enquanto os dados originais, sem qualquer tratamento, são usados como atributos para o modelo.

As comparações estatísticas pelo teste t pareado (normalização incorporada  $\times$  sem normalização e normalização incorporada  $\times$  normalização padrão) corroboraram o melhor funcionamento da estratégia proposta diante dos outros dois cenários. Apesar de não haver diferença estatística para a medida MSE (p-valor = 0,0577 e p-valor = 0,1025, respectivamente), podemos afirmar melhor desempenho de nossa estratégia pela métrica MAE (p-valor < 0,0001 em ambas as comparações), principal métrica utilizada na literatura para a localização de faltas.

Tabela 6.29: Média e desvio padrão dos resultados em km comparando diferentes cenários de normalização

Cenário de escalonamento	MSE	MAE
Sem normalização	1,3448 $\pm$ 0,0605	0,6718 $\pm$ 0,0041
Normalização padrão	1,3371 $\pm$ 0,0552	0,6750 $\pm$ 0,0033
Normalização incorporada	1,3242 $\pm$ 0,0647	0,6646 $\pm$ 0,0040

A performance inferior da normalização padrão em relação a incorporada pode estar relacionada a discrepância dos valores dos atributos do padrão de teste, ou seja, valores mínimos inferiores e valores máximos superiores nas características do teste amostral em comparação com aqueles dentro do conjunto de validação. Por outro lado, este problema não afeta a nossa normalização incorporada, uma vez que o método adiciona cada padrão de teste dentro do conjunto de validação original antes da normalização. Portanto, esta análise é crucial para endossar a importância da normalização do conjunto de validação junto ao padrão de teste.

O escalonamento incorporado dimensiona o padrão de teste e o conjunto de validação juntos e individualmente para cada novo exemplo para então definir a RC, levando a melhores resultados da SDR para estimar a localização da falta. O motivo deste processo *online* está relacionado ao fato de que o algoritmo KNN (baseado em distância) utilizado para determinar a RC pode ser afetado pela escala dos dados, o que pode influenciar os exemplos retornados para compor esta região. Ou seja, características com escalas discrepantes tendem a dominar as demais e podem levar a uma RC imperfeita e, conseqüentemente, a uma perda de desempenho.

Assim sendo, estes resultados respondem a QP3, confirmando que a normalização incorporada permite uma melhor definição da RC e, conseqüentemente, em melhores desempenhos.

Além do mais, a decisão pela normalização embutida fundamenta-se ainda pelo fato de que nem todo algoritmo preditor (seja ele classificador ou regressor) se beneficiará pela normalização das características como etapa de pré-processamento (Amorim et al., 2023). Mesmo que tal pré-processamento seja realizado, nada impede que o módulo de normalização da estrutura SDR também seja utilizada, uma vez que, no máximo, não acarretará em ganho no desempenho, mantendo a performance semelhante caso esse módulo não tivesse sido usado. Optamos por não normalizar os atributos como etapa de pré-processamento devido a natureza do problema de falhas, dada a sua complexidade e imprevisibilidade. Cabe esclarecer que, para a tarefa de localização, utilizamos o conceito de normalização em dois momentos distintos: (1) alteração das amplitudes de dois ciclos pós-falta a partir de um ciclo pré-falta previamente a extração de características, semelhantemente ao desempenhado para a classificação; e (2) definição da RC dentro do *framework* SDR.

### 6.2.2.2 Avaliação de Métricas de Distância

Na Tabela 6.30 estão dispostos os desempenhos do nosso método com o emprego de diferentes medidas de distância. O teste *one-way* ANOVA manifestou diferença estatisticamente significativa entre os desempenhos ( $p$ -valor = 0,0001 para ambas as métricas), enquanto o pós-teste de Tukey nos permitiu identificar quais métricas se destacaram sobre as demais. É possível afirmar que o método desempenhou melhor ao empregar as distâncias Cosseno e Cityblock frente ao uso das distâncias Euclidiana e Geodésica.

Tabela 6.30: Média e desvio padrão dos resultados em km comparando as diferentes métricas de distância

Medida de distância	MSE	MAE
Euclidiana	1,3827 ± 0,0641	0,6745 ± 0,0044
Cosseno	1,3237 ± 0,0525	0,6653 ± 0,0047
Geodésica	1,3776 ± 0,0585	0,6745 ± 0,0042
Cityblock	1,3242 ± 0,0647	0,6646 ± 0,0040

Apesar de não ser possível declarar qual das duas medidas foi a melhor (*i.e.*, Cosseno ou Cityblock) devido aos desempenhos não apresentarem diferença estatisticamente significativa tanto para MAE quanto para MSE, selecionamos o Cityblock como métrica de distância para compor nosso método devido aos seus melhores desempenhos em experimentos anteriores (Ensina et al., 2024). Inclusive, Swetapadma e Yadav (2018) observaram descobertas semelhantes sobre esta métrica em seu estudo ao também analisarem diferentes funções de distância para seu método baseado no algoritmo KNN. Seus resultados também determinaram o Cityblock como a métrica com melhor desempenho para o problema de localização de faltas. No entanto, é crucial enfatizar os diferentes usos do algoritmo KNN: enquanto Swetapadma e Yadav (2018) emprega tal algoritmo como regressor para prever a localização da falta, nosso método utiliza-o apenas para definir a RC.

Esta análise é interessante em razão de alcançarmos melhores resultados apenas alterando o modo em que a distância é computada, evidenciando a sua importância. Em complemento, essa investigação visa explorar o comportamento de cada medida na base de dados, uma vez que determinadas métricas podem se sobressair perante as demais por eventualmente serem beneficiadas pela distribuição diversificada e particular dos dados na base. Neste cenário, outras medidas como Canberra e Mahalanobis podem ser usadas alternativamente (Deza e Deza, 2016).

### 6.2.2.3 Avaliação de Critérios de Seleção

Na Tabela 6.31 estão reportados os resultados considerando diferentes critérios que estabelecem o limite de erro usado como referência para a seleção dos regressores mais competentes do conjunto a partir de seu desempenho na RC. Os resultados evidenciam que a abordagem padrão usualmente utilizada na literatura ( $\frac{E_{max}-E_{min}}{2}$ ) é superada por todos os outros critérios, confirmados pelo *one-way* ANOVA (p-valor = 0,0006 para a métrica MSE e p-valor < 0,0001 para MAE) e pós-teste de Tukey.

Tabela 6.31: Média e desvio padrão dos resultados em km comparando os diferentes critérios de seleção

<b>Critério de seleção</b>	<b>MSE</b>	<b>MAE</b>
Critério padrão	1,4004 ± 0,0676	0,6704 ± 0,0048
Mediana	1,3320 ± 0,0604	0,6669 ± 0,0045
Média aritmética	1,3553 ± 0,0609	0,6675 ± 0,0047
Média harmônica	1,3483 ± 0,0766	0,6698 ± 0,0048
Média geométrica	1,3242 ± 0,0647	0,6646 ± 0,0040

Embora não haja diferença estatisticamente significativa entre média aritmética, média geométrica, média harmônica e mediana para a medida MSE pelo pós-teste de Tukey, a média geométrica teve melhor desempenho frente os demais critérios para a métrica MAE. Ainda que a literatura da SDR evidencie maior atenção à avaliação do hiperparâmetro da medida de competência (Moura et al., 2019b, 2020), pudemos demonstrar que não devemos negligenciar os demais hiperparâmetros envolvidos em um método SDR. Neste cenário, nossos resultados comprovaram, por exemplo, que definir o critério de seleção é igualmente crucial, repercutindo diretamente na melhoria do desempenho.

### 6.2.2.4 Avaliação da Aplicação do Filtro

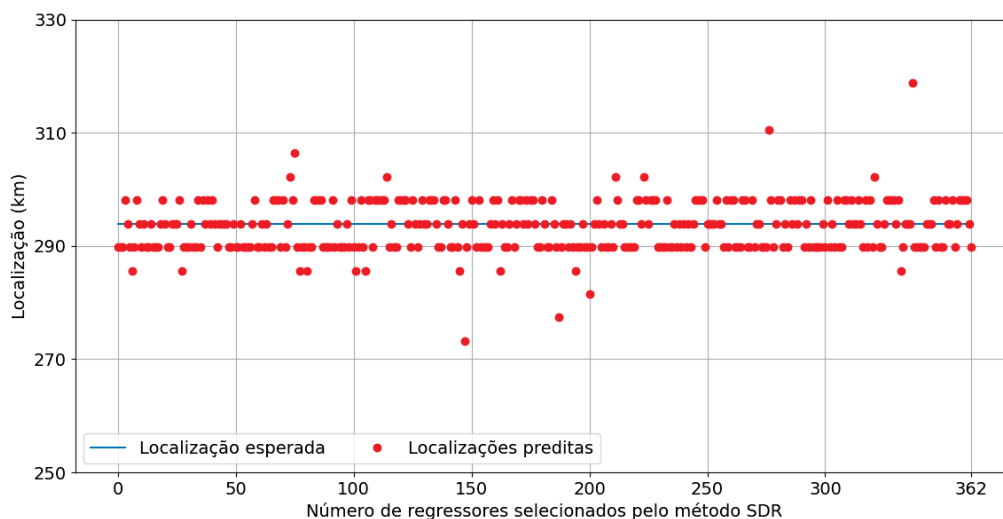
Na Tabela 6.32 confrontamos o desempenho de nosso método diante ao não uso do módulo de filtro, que objetiva remover as predições que sejam discrepantes provenientes dos regressores selecionados. Nota-se que o método desempenhou significativamente melhor comparada a não aplicação deste módulo para a predição do local da falha, corroborada pelo teste t (p-valor < 0,0001 tanto para MSE quanto para MAE). Portanto, o uso deste módulo inédito permite a remoção de predições imprecisas, reverberando em melhores desempenhos, o que responde a nossa QP4.

Tabela 6.32: Média e desvio padrão dos resultados em km referentes ao emprego do módulo de filtro

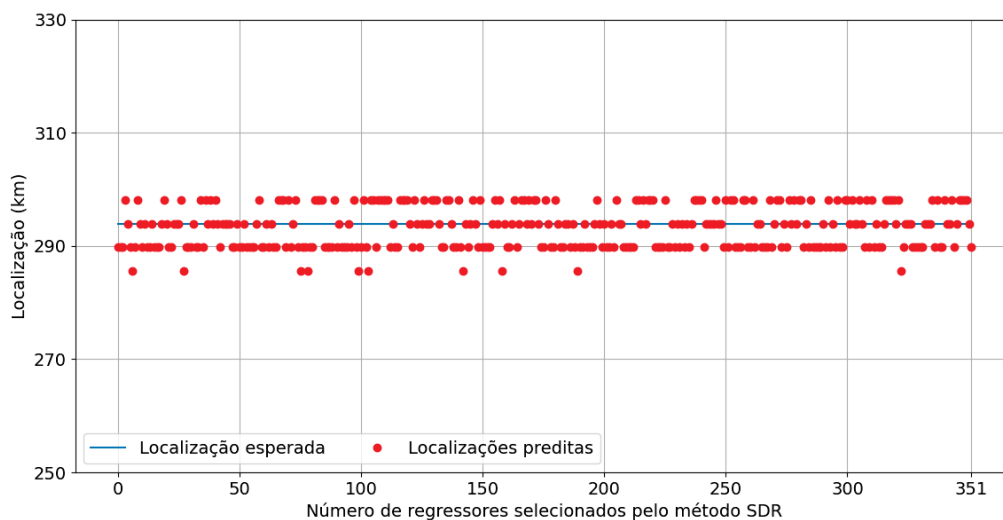
<b>Filtro</b>	<b>MSE</b>	<b>MAE</b>
Sem filtro	1,5794 ± 0,0838	0,6942 ± 0,0049
Com filtro	1,3242 ± 0,0647	0,6646 ± 0,0040

Na Figura 6.6 está ilustrado um exemplo que demonstra a operação deste pós-processamento. Observa-se pela Figura 6.6(a) que algumas predições estão destoantes da grande maioria (*i.e.*, *outliers*), assim como distantes do valor esperado para o atributo-alvo. Todavia, os regressores responsáveis por tais predições são removidos ao aplicarmos o módulo de filtro (Figura 6.6(b)), tornando a predição final do método mais assertiva. Nosso trabalho é o primeiro do estado da arte da SDR a propor uma etapa de pós-processamento como o módulo de “Filtro”, enquanto na classificação recentemente alguns trabalhos têm considerado essa etapa de pós-processamento (Cordeiro et al., 2023b,a).

Considerando os resultados apresentados para os quatro componentes avaliados neste estudo (ou seja, módulo de normalização embutido, função de distância, critério de seleção e módulo de filtro), podemos estabelecer que a aplicação do filtro levou a um maior ganho no desempenho, seguido pelo critério de seleção, pela função de distância e pela normalização. Essas investigações são interessantes pois demonstram a importância de buscar a melhor combinação de hiperparâmetros geralmente negligenciados ou inexplorados pelo estado da arte da SDR.



(a) Sem filtro



(b) Com filtro

Figura 6.6: Exemplo da dispersão das previsões (a) sem o uso e (b) com a aplicação do módulo de filtro

Fonte: autoria própria



### 6.2.3 Análise da Importância dos Atributos

Nesta subseção, avaliamos a importância dos atributos presentes no conjunto de dados FADbF, o qual compreende tanto os atributos normalmente utilizados em trabalhos relacionados para análise de falhas (Tabela 4.1), quanto os novos incluídos neste projeto.

O bom desempenho de nosso método pode ser justificado de certa maneira pela análise realizada em relação à importância das características para o algoritmo ERT, cujas árvores compuseram o conjunto de regressores para a técnica SDR. Na Tabela 6.33 estão apresentados os dez principais atributos ordenados por suas importâncias para o algoritmo ERT, em que quanto maior o valor, mais relevante é o atributo. As pontuações apresentadas nesta tabela representam a média e o desvio padrão de cada atributo considerando os dez modelos pré-treinados, os mesmos utilizados para as análises das subseções anteriores. A relevância de uma característica é calculada com base na importância de Gini com uma faixa normalizada de pontuações (Ishwaran, 2015).

Tabela 6.33: Listagem dos dez principais atributos ordenados por importância para o algoritmo ERT, considerando a média e o desvio padrão da medida de Gini

<b>Ranking</b>	<b>Atributo</b>	<b>Gini score</b>
#1	Desvio padrão (sinal da corrente)	0,0431 ± 0,0009
#2	Variância (sinal da corrente)	0,0298 ± 0,0003
#3	Média geométrica (sinal da corrente)	0,0251 ± 0,0016
#4	Variância (sinal da tensão)	0,0220 ± 0,0006
#5	Covariância (sinal da corrente)	0,0200 ± 0,0005
#6	Desvio padrão (sinal da tensão)	0,0196 ± 0,0007
#7	Distância (sinal da tensão)	0,0173 ± 0,0011
#8	Entropia de Shannon (sinal da tensão)	0,0163 ± 0,0005
#9	AUC (sinal da corrente)	0,0147 ± 0,0005
#10	Entropia de Shannon (sinal da corrente)	0,0139 ± 0,0000

Nesta análise, agrupamos os atributos inicialmente separados por fase (A, B ou C) em um único atributo na Tabela 6.33, mas separadamente por tipo de sinal. Por exemplo, considerando o atributo desvio padrão extraído para o sinal da corrente, originalmente existem três características baseadas nele, uma para cada fase (A, B ou C). Assim, somamos as pontuações de Gini para esses três atributos em um único grupo para permitir uma análise por característica, individualmente para cada tipo de sinal. Consulte o Apêndice B para obter detalhes completos sobre as pontuações de Gini, incluindo os valores individualmente para cada característica.

Esta análise evidencia que seis das dez principais características são originais do conjunto de dados FADbF: variância (sinal da corrente); média geométrica (sinal da corrente); variância (sinal da tensão); covariância (sinal da corrente); distância (sinal da tensão); e AUC (sinal da corrente). Embora o desvio padrão e a variância tenham certa relação, ainda assim manifestam perspectivas estatísticas diferentes. Como resultado, a variância neste contexto é considerada uma característica nova. Observe que optamos por não agrupar os atributos de ambos os tipos de sinais em um único atributo, por exemplo, desvio padrão para sinais de corrente e tensão em um mesmo agrupamento, e examiná-los separadamente devido ao fato de que as formas de onda de corrente e tensão manifestam propriedades distintas. Um exemplo ocorre para o próprio atributo de desvio padrão, determinado como o atributo mais importante a partir do sinal da corrente, mas apenas na sexta posição para a forma de onda da tensão.



Neste trabalho não realizamos nenhuma seleção de atributos, o que eventualmente poderia repercutir em ganho de desempenho. Poderíamos, por exemplo, utilizar apenas o desvio padrão ou a variância, uma vez que ambos são utilizados para medir a dispersão dos dados em torno da média aritmética. No entanto, este tipo de pré-processamento será considerado em trabalhos futuros.

#### 6.2.4 Comparação com o Estado da Arte

Nesta seção, comparamos o desempenho da nossa abordagem diante de cinco trabalhos relacionados descritos no Capítulo 3. Em seguida, discutimos sobre algumas propriedades do conjunto usado como *pool* de regressores, juntamente com a comparação com algoritmos *ensemble* tradicionais como RF, ERT e Gradient Boosting.

Na Tabela 6.34 está apresentada a comparação de desempenho entre alguns trabalhos do estado da arte revisados na Capítulo 3 e o método proposto. Reproduzimos esses métodos usando o mesmo conjunto de dados (FADbF) usado para avaliar nossa abordagem para uma comparação justa, com a mesma divisão em conjuntos de treinamento e teste conforme apresentamos na Subseção 6.2.1. Como resultado, os procedimentos de pré-processamento e os atributos utilizados como entrada para os algoritmos podem diferir de suas propostas originais. Por exemplo, técnicas de processamento de sinais como DWT, WPT e DFT não foram realizadas para os métodos que as utilizam, uma vez que esses procedimentos devem ser realizados nos sinais de corrente e tensão antes da extração de características, que não compõe o conjunto de dados FADbF pelos motivos discutido no final da Capítulo 3. Por outro lado, realizamos etapas como o uso de múltiplos módulos para os métodos que o exigem (Swetapadma e Yadav, 2017, 2018; Bhatnagar et al., 2023). Considerando que nosso objetivo é avaliar o conjunto de dados proposto, utilizamos todo o conjunto FADbF como entrada para todos os métodos, enquanto cada esquema de regressão foi reproduzido com a mesma configuração descrita nos respectivos trabalhos.

Tabela 6.34: Média e desvio padrão dos resultados em km para os métodos relacionados e o método proposto

Referência	MSE	MAE
Swetapadma e Yadav (2017)	33,9113 ± 4,7636	3,0266 ± 0,0359
Swetapadma e Yadav (2018)	183,0063 ± 5,7464	6,6843 ± 0,0715
Ferreira et al. (2020)	3959,0408 ± 189,9829	48,6720 ± 1,4212
Bhatnagar et al. (2023)	5,1419 ± 0,3567	1,3690 ± 0,0143
Método proposto	1,3242 ± 0,0647	0,6646 ± 0,0040

Nosso método demonstrou superioridade sobre todos os outros métodos, corroborado estatisticamente pelo teste de Friedman ( $p$ -valor < 0,0001 para ambas as métricas MSE e MAE) e pós-teste de Dunn. Essas altas taxas de erro dos métodos relacionados podem ser explicadas em parte pelo fato de que esses métodos foram originalmente desenvolvidos para linhas de transmissão diferentes daquela usada para o conjunto de dados FADbF, bem como pela utilização de diferentes taxas de amostragem para coleta de dados. Além disso, as propriedades dos sistemas de energia e as suas complexidades também diferem. Consequentemente, todos esses aspectos resultam em diferentes assinaturas de falta (por exemplo, amplitudes de sinal distintas mesmo para linhas de transmissão com o mesmo comprimento) entre os conjuntos de dados (Raza et al., 2020; Ensina et al., 2022b).

Como indicado por Ferreira et al. (2016), o desempenho do sistema de diagnóstico de falhas é garantido apenas para a linha de transmissão e condições de falha para as quais o sistema de diagnóstico de falhas foi desenvolvido. Esta afirmação está refletida nos resultados de Bhatnagar et al. (2023), que manifestou o segundo melhor desempenho para o conjunto de dados FADbF. Os autores usaram o sistema IEEE 14-barras para desenvolver e avaliar sua abordagem (Capítulo 3), o qual apresenta semelhanças com o sistema IEEE 9-barras usado para simular os eventos de falta para o conjunto de dados FADbF (Seção 4.1). Consequentemente, o melhor desempenho do seu método em comparação aos demais pode estar relacionado a tal aspecto. Além disso, o desempenho de seu método para o conjunto de dados FADbF é semelhante ao desempenho relatado para o sistema de energia considerado em seu estudo original, que representou MAE de cerca de 1 km contra MAE de 1,3690 km em nosso conjunto de dados.

Na sequência, buscamos fundamentar a escolha pelo uso das árvores do algoritmo ERT como *pool* de regressores em nossa SDR. A escolha de utilizar a árvore de decisão como modelo base para compor o conjunto de preditores é justificada em parte pelos resultados relatados na Tabela 6.35. A implementação da árvore de decisão utilizada neste estudo foi o algoritmo CART (Breiman et al., 1984), conforme disponibilizada pela ferramenta scikit-learn (Seção 5.3). Esses experimentos iniciais compreenderam o conjunto de dados FADbF usando apenas os conjuntos de treinamento e validação.

Tabela 6.35: Média e desvio padrão dos resultados em km para cada algoritmo base

Algoritmo	MSE	MAE
CART	42,7314 ± 4,9654	3,6478 ± 0,0564
MLP (Ferreira et al., 2020)	3959,0408 ± 189,9829	48,6720 ± 1,4212
SVM	2448,5088 ± 17,9166	37,5053 ± 0,1096
KNN ( $k = 3$ )	121,2048 ± 5,5794	6,5848 ± 0,0468

O algoritmo CART teve melhor desempenho individual frente aos demais modelos, conforme evidenciado pelo teste *one-way* ANOVA seguido do pós-teste de Tukey, que manifestou diferenças estatísticas entre o CART e cada preditor. Além disso, estes algoritmos requerem um custo computacional superior às árvores de decisão. Consequentemente, esses aspectos explicam a escolha da árvore de decisão como preditor base.

Por sua vez, na Tabela 6.36 retratamos os desempenhos de diferentes técnicas *ensemble*, todas baseadas em árvores de decisão.

Tabela 6.36: Média e desvio padrão dos resultados em km para cada algoritmo *ensemble*

Abordagem	MSE	MAE
ERT	2,2340 ± 0,1385	0,8029 ± 0,0085
RF	8,7946 ± 1,7002	1,4307 ± 0,0327
<i>Bagging</i> de CART	8,8600 ± 1,6808	1,4337 ± 0,0317
LightGBM	22,2744 ± 0,8779	3,4753 ± 0,0188
XGBoost	28,3831 ± 0,8044	3,8732 ± 0,0523

Como pode ser observado, o ERT teve um desempenho melhor comparado aos demais algoritmos, justificando a sua utilização como referência para obter o *pool* de regressores mais promissor. O algoritmo ERT obtém os limiares aleatoriamente para cada atributo candidato, selecionando o melhor desses limites gerados como regra de divisão (Subseção 2.2.3). Como resultado, pode reduzir a variância do modelo e controlar o sobreajuste. Esses aspectos podem ter contribuído para um melhor desempenho desta abordagem em relação às demais técnicas.

A escolha do ERT também se justifica pelo desempenho do seu oráculo (linha “Oráculo” na Tabela 6.37), alcançando um desempenho perfeito, ou seja, pelo menos uma árvore do conjunto consegue atingir a exata localização da falha. O oráculo é uma estratégia abstrata que sempre seleciona o modelo que prevê o rótulo correto para uma determinada amostra de teste, se tal modelo existir (Kuncheva, 2002; Cruz et al., 2018a). Em outras palavras, o oráculo estabelece o melhor desempenho possível que o conjunto pode atingir, representando o esquema ideal de seleção de preditores.

Para determinar o desempenho do oráculo, foram realizadas as seguintes etapas: primeiro, calculamos o erro absoluto entre a posição real da falha e a posição prevista da falha individualmente para cada estimador do *pool*; a seguir, selecionamos apenas o regressor que apresentou o menor erro para o padrão de teste específico para prever seu rótulo. É fundamental esclarecer que tivemos que verificar as previsões de cada modelo diretamente com o valor alvo real do padrão de teste para então selecionar o regressor que apresentasse o menor erro, o que não seria possível em um cenário real porque o valor alvo real obviamente não estará disponível. É por isso que o oráculo é considerado uma estratégia abstrata.

Na Tabela 6.37 estão apresentados os resultados dos experimentos comparando nosso método e a SDR proposta por Moura et al. (2020), denominada *Dynamic Weighting Selection with Measure Selection* (DWS-MS). A escolha desse método para a comparação com a nossa abordagem é justificada pelo fato de que nossa estrutura SDR baseia-se no DWS-MS. É essencial mencionar que fornecemos exatamente o mesmo conjunto pré-treinado de regressores para ambos os métodos com tamanho  $N = 700$ .

Os hiperparâmetros do DWS-MS foram definidos de forma semelhante à nossa técnica neste trabalho, com a mesma medida de competência (ou seja, erro quadrático máximo),  $K = 30$ , e o mesmo conjunto treinado de modelos. Tentamos obter a melhor configuração possível deste método para o conjunto de dados FADbF, a fim de permitir uma comparação justa. Para os resultados do nosso método relatados na Tabela 6.37, utilizamos o erro quadrático máximo como medida de competência, Cityblock como medida de distância e a média geométrica como critério para determinação do limite, no qual selecionamos apenas os regressores que apresentam erro inferior a este limite. Esta configuração demonstra o melhor desempenho do nosso método para o conjunto de dados FADbF.

Tabela 6.37: Média e desvio padrão dos resultados em km para cada método SDR apenas para 10 kHz

<b>Abordagem</b>	<b>MSE</b>	<b>MAE</b>
DWS-MS (Moura et al., 2020)	1,6877 ± 0,1030	0,7066 ± 0,0050
Método proposto	1,3242 ± 0,0647	0,6646 ± 0,0040
Oráculo	0,0000 ± 0,0000	0,0000 ± 0,0000
Oráculo dos modelos selecionados	0,0013 ± 0,0011	0,0003 ± 0,0003

Nossa técnica demonstrou desempenho superior para todas as comparações, considerando o teste t pareado com intervalo de confiança de 95%. Observe que comparamos nosso método individualmente com o DWS-MS e todos os *ensembles* listados na Tabela 6.36 sem considerar os modelos oráculos (duas últimas linhas da Tabela 6.37). O teste de hipótese revelou diferenças estatisticamente significativas comparando nossa abordagem individualmente com cada uma das outras técnicas tanto para a métrica MSE e quanto para MAE (p-valor < 0,0001 para a maioria das comparações). Tais resultados, os quais demonstram a superioridade do nosso método, respondem a nossa QP5.

O melhor desempenho da nossa técnica pode ser explicado pelo fato de termos ido mais longe na variação dos hiperparâmetros do SDR, juntamente com a adição da normalização embutida e da etapa de filtro. Enquanto o DWS-MS visa obter a melhor medida de competência, nosso método busca a métrica de distância ideal e os critérios de seleção juntamente com a melhor medida de competência.

Embora o oráculo geral não resulte em nenhum erro (linha “Oráculo” na Tabela 6.37), com pelo menos um regressor no conjunto de modelos prevendo exatamente a localização da falha de cada padrão de teste, o principal desafio é selecionar este regressor. Esta é a nossa principal motivação na utilização da técnica de ensemble com seleção dinâmica frente ao uso de modelos individuais amplamente utilizados no estado da arte de localização de faltas.

Da mesma maneira, o desempenho do oráculo evidenciou que podemos investigar melhor as estratégias de seleção dinâmica. Para tanto, exploramos outras métricas de distância e outros critérios de seleção visando ganho de desempenho, conforme demonstrado anteriormente. No entanto, devemos enfatizar que alcançar o mesmo desempenho do oráculo é uma tarefa complexa, uma vez que selecionamos múltiplos regressores para definir o valor alvo, a menos que todos informem exatamente o mesmo (e correto) valor, o que é improvável.

Além do oráculo de todo o *pool*, avaliamos o oráculo apenas entre os modelos selecionados pelo nosso esquema SDR (linha “Oráculo dos modelos selecionados” na Tabela 6.37). Esta análise visa determinar o melhor desempenho possível apenas em relação aos modelos escolhidos dinamicamente pela nossa técnica SDR.

Considerando o conjunto original de modelos com 700 estimadores, o método SDR proposto selecionou cerca de  $382 \pm 44$  regressores. Inclusive, se observarmos o oráculo entre esses modelos selecionados (linha “Oráculo dos modelos selecionados” na Tabela 6.37), é possível identificar que nossa técnica selecionou preditores precisos. Além disso, estes resultados demonstram que novas estratégias podem ser adicionadas como uma seleção dinâmica de múltiplas etapas, utilizando diferentes medidas de competência para cada nova etapa, a fim de alcançar melhores resultados (Moura et al., 2021).

Uma perspectiva para futuras investigações está relacionada ao estudo de preditores de aprendizado profundo (*e.g.*, LSTM e CNN) como aprendizes base. Esta análise é interessante dado que métodos que baseiam-se em algoritmos DL e outros que utilizam algoritmos clássicos como árvores de decisão podem manifestar diferentes dispersões dos erros, uma vez que usualmente empregam diferentes representações de entrada juntamente a natureza diferente dos algoritmos. Neste cenário, podemos agregar modelos baseados em DL com algoritmos clássicos para compor o conjunto de regressores (conjunto heterogêneo), de maneira semelhante ao já utilizado na literatura para previsão (*forecasting*) de séries temporais (de Oliveira et al., 2022). Além disso, podemos utilizar estruturas mais simples de modelos de DL para compor o *pool* ao invés de exigir uma arquitetura de DL única e complexa, como proposto em alguns trabalhos da literatura (Fan et al., 2019; Wang et al., 2022; Bon e Dai, 2022).

Devido à propriedade chave da estratégia SDR ser a diversidade entre os modelos, podemos treinar algoritmos DL com dados brutos de sinais de corrente e tensão e treinar algoritmos clássicos (por exemplo, CART e SVM) com características extraídas manualmente, como aquelas do conjunto de dados FADbF (Seção 4.2). Assim, podemos não só diversificar os tipos de dados utilizados para treinar os algoritmos, mas também explorar as diferentes propriedades de cada algoritmo, garantindo a diversidade sob diversas perspectivas. Pensando nisso, acreditamos que podemos conseguir melhores resultados e também reduzir o tamanho do *pool* (atualmente com  $N = 700$ ). Por último, mas não menos importante, apesar da avaliação apenas com a técnica DWS, nosso *framework* SDR também pode ser adaptado para trabalhar com DS e DW.

### 6.2.5 Desempenho do Método Proposto para Generalização

Neste conjunto final de experimentos, analisamos se o método para localizar falhas também é apto a generalização, de modo a prever a posição da falha mesmo que tal LT não tenha sido usada para o seu treinamento. É digno de nota que o nosso trabalho é o primeiro que realiza este tipo de análise para a localização a partir do uso de algoritmos de AM. Na Tabela 6.38 estão apresentados os resultados para a capacidade de generalização de nosso método para a localização de falhas.

Tabela 6.38: Média e desvio padrão dos erros em km para a capacidade de generalização (conjunto LT2)

Taxa de amostragem	Sem normalização		Com normalização	
	MSE	MAE	MSE	MAE
<b>256 Hz</b>	2830,8651 ± 5,1016	44,8405 ± 0,0433	1305,1393 ± 14,5791	27,0867 ± 0,2237
<b>1 kHz</b>	2547,5801 ± 14,5597	41,6653 ± 0,1292	1218,1812 ± 8,4187	26,5280 ± 0,1142
<b>5 kHz</b>	2945,7179 ± 6,8717	45,6699 ± 0,0570	1021,6649 ± 12,4856	24,7193 ± 0,1800
<b>10 kHz</b>	2925,3060 ± 7,2921	45,5166 ± 0,0688	967,6502 ± 5,6314	23,6016 ± 0,0973

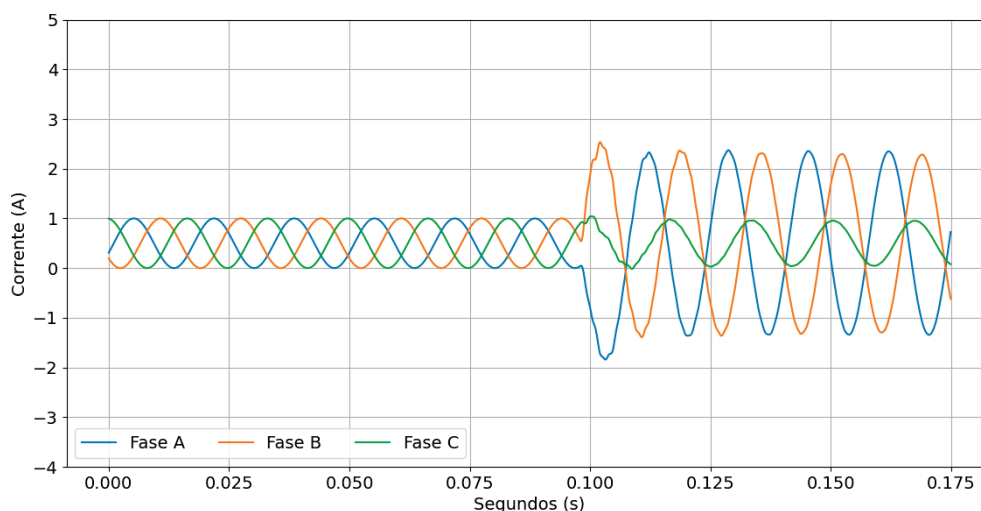
Apesar dos resultados manifestarem baixo desempenho, com MAE percentual por volta de 22,1308% para o melhor resultado na LT2 com 106,646 km de extensão na frequência de 10 kHz, ainda assim fica evidente a relevância da normalização dos ciclos pós-falha pelo ciclo pré-falha, etapa que antecede a extração de atributos. Por exemplo, ainda considerando a taxa de 10 kHz, sem a normalização o erro MAE percentual vai para 42,6501%.

Mesmo que a transformação dos sinais pela normalização tenha repercutido num ganho de desempenho comparado ao seu não uso, ainda há a necessidade de melhorias futuras no método para habilitar, de fato, a generalização (QP6). Apesar disso, decidimos por manter essa etapa no método devido ao fato de que, caso ocorra alguma alteração nas amplitudes dos sinais mesmo que para a mesma LT, apesar da tendência do aumento dos erros eles deverão ser inferiores do que caso essa normalização não tivesse sido realizada.

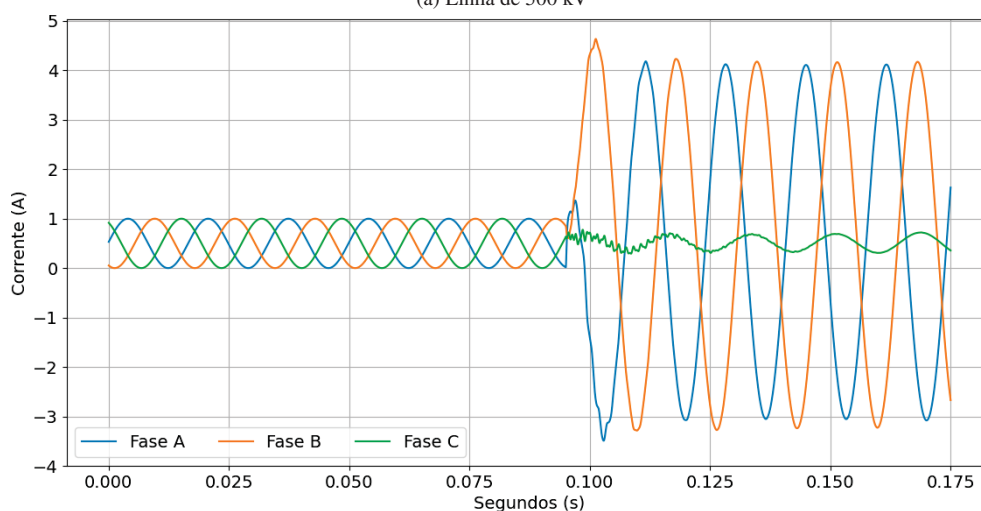
Uma possível explicação da inabilidade para a generalização pode estar atrelada a desigualdade entre as assinaturas das falhas para cada LT, mesmo que os parâmetros de falta sejam iguais. Exemplificamos na Figura 6.7 um exemplo de falha para ambas as linhas, cujos parâmetros são correspondentes: tipo ABG, resistência de 50  $\Omega$  e ocorrência da perturbação a 50% da extensão da linha. Repare que tais representações refletem os respectivos sinais da corrente após a etapa de normalização do sinal.

Embora os parâmetros para cada falha constituam valores equivalentes entre si (*i.e.*, mesmo valores para tipo, resistência e localização em termos percentuais do comprimento da linha), ainda assim as faltas reverberaram em assinaturas distintas para cada LT. Esse comportamento pode ser constatado pela comparação entre as amplitudes das formas de onda, de modo que as amplitudes do período pós-falha para a linha de 230 kV (Figura 6.7(b)) são consideravelmente maiores do que para a linha de 500 kV (Figura 6.7(a)). Isso se deve a localização em termos percentuais representar distâncias completamente distintas: 207 km para a LT1 de 414 km e 53,323 km para a LT2 de 106,646 km de extensão. Logo, as amplitudes pós-falha são superiores para a linha de 230 kV devido a falha estar mais próxima ao terminal em comparação a respectiva falha para a LT de 500 kV que está mais distante. Conforme descrevemos na Subseção 2.1.2, faltas próximas a um terminal provocam correntes elevadas, enquanto defeitos mais distantes provocam correntes de menor intensidade (Clark, 1979).





(a) Linha de 500 kV



(b) Linha de 230 kV

Figura 6.7: Representação de falhas para as LTs de (a) 500 kV e (b) 230 kV, cujos parâmetros de falta são iguais: tipo ABG, resistência de  $50 \Omega$  e ocorrência a 50% da extensão da linha

Fonte: autoria própria

Portanto, a realização da etapa de normalização nos sinais de corrente e tensão não foi o suficiente para tornar tais formas de onda proporcionais entre as linhas, o que, em tese, permitiria que a generalização fosse factível. É possível que o comportamento dessemelhante entre as assinaturas de falha no período pós-falha, mesmo que empregando os mesmos parâmetros, demonstre uma limitação do uso desses tipos de dados visando a generalização da tarefa de localização. Frente a isso, estudos futuros demandam atenção para que novos tratamentos sejam realizados nos dados e/ou novas representações sejam utilizadas para prover mais informação ao algoritmo, como os dados de ondas viajantes (*traveling waves*) (da Silva et al., 2005; Akmaz et al., 2018).

### 6.2.6 Complexidade

Apesar do bom desempenho e eficácia do método proposto para a localização de falhas, considerando o cenário que o modelo é treinado e avaliado com dados provenientes da mesma LT, podemos identificar duas desvantagens principais: o custo computacional e o número de modelos necessários no conjunto de regressores.

A primeira fraqueza está relacionada ao fato de que o método precisa realizar uma sequência de passos individualmente para cada novo exemplo a ser predito (Seção 5.2.2): (1) normalizar o conjunto de validação junto com o teste padrão pela técnica MinMax ( $O(n)$ , onde  $n$  é o número de exemplos no conjunto de validação, mais o padrão de teste); (2) definir a RC pelo algoritmo KNN ( $O(k \times n \times d)$ , onde  $k$  é o número de vizinhos,  $n$  é o número de exemplos no conjunto de validação e  $d$  é a dimensionalidade dos dados); (3) calcular o desempenho de cada regressor do *pool* para a RC com base na medida de competência ( $O(r)$ , onde  $r$  é o número de regressores); (4) determinar o limite de desempenho usando o critério de seleção ( $O(r)$ ); (5) selecionar os regressores que apresentarem erro inferior a este limite ( $O(r)$ ); (6) aplicar o filtro a partir das previsões dos regressores selecionados ( $O(m)$ , onde  $m$  é a quantidade de modelos selecionados no passo anterior); (7) calcular os pesos para cada modelo selecionado ( $O(m)$ , apesar do filtro remover alguns dos  $m$  regressores); e (8) então realizar a previsão final ( $O(m)$ ).

Conseqüentemente, há um aumento no tempo de processamento necessário para estimar o valor do atributo-alvo em comparação com algoritmos *ensemble* estáticos como RF e ERT, que basicamente apenas calculam a média aritmética de todas as previsões dos modelos para obter a decisão final. Porém, este tempo não é um problema, visto que nosso método foi desenvolvido para ser utilizado como uma aplicação *offline*, além do fato de que deverá ser usado para rotular um evento por vez, tornando esse custo irrelevante. Podemos estimar a complexidade do tempo de predição do nosso método somando as complexidades das etapas mencionadas anteriormente, as quais estão identificadas entre parênteses no último parágrafo. Em resumo, a complexidade aproximada de predição do método é  $O(n + (k \times n \times d) + r + m)$ .

Outro aspecto associado ao custo computacional é a necessidade de manter todo o conjunto de validação junto com o *pool* de modelos durante a operação do método SDR, semelhante a qualquer outra abordagem de seleção dinâmica. Essa propriedade é necessária devido ao método SDR requerer do conjunto de validação para utilizá-lo no procedimento de aprendizado lento realizado pelo algoritmo KNN, que é responsável por obter a RC para cada nova instância. Como resultado, há uma demanda por espaço de armazenamento extra. Neste estudo, por exemplo, cada arquivo do conjunto de validação corresponde a cerca de 96 MB com 42.000 instâncias e 126 atributos, o que não representa um tamanho significativo. É claro que o tamanho do arquivo deve variar dependendo da base de dados, mas isso não inviabilizará o método. Uma possível solução para este problema que será investigada posteriormente é a seleção de protótipos (Garcia et al., 2012). Esta técnica busca um conjunto de dados reduzido, servindo como uma visão condensada de todo o banco de dados original, podendo inclusive melhorar o desempenho (Kordos et al., 2019).

A última desvantagem refere-se ao elevado número de modelos no *pool* necessários para a tarefa de localização de faltas, permitindo que o método alcance os desempenhos apresentados ao longo desta seção. Este problema também repercute no custo computacional, pois exige que cada um desses regressores seja avaliado. Assim, esforços futuros serão direcionados para a melhoria da fase de geração (Subseção 5.2.2.1), seguindo orientações como as citadas no final da Subseção 6.2.4.



É importante ressaltar novamente que, mesmo com essas limitações, nosso método demonstrou ser preciso e eficaz. Esses problemas não afetam a sua viabilidade, uma vez que ele foi desenvolvido para ser utilizado como aplicação *offline*. Ferreira et al. (2016) enfatizam que o custo computacional não é crítico para aplicações *offline*, permitindo o uso de técnicas mais sofisticadas para localização de faltas.

### 6.3 CONSIDERAÇÕES FINAIS

Neste capítulo foram relatados e discutidos os resultados alcançados ao longo do projeto. Os desempenhos reportados para a classificação do tipo de falta demonstraram a versatilidade e a efetividade do método proposto, manifestando tanto a flexibilidade em relação a diferentes taxas de amostragem quanto a capacidade de generalização para linhas de transmissão com configurações distintas daquela usada para treinar o modelo. Por outro lado, apesar de nossa abordagem para a localização de falhas desempenhar bem perante diferentes frequências de amostragem, não foi constatada a eficácia para a generalização. De qualquer modo, os resultados de nossa estratégia SDR superam os métodos do estado da arte analisados.

## 7 CONCLUSÃO

As linhas de transmissão são elementos essenciais do sistema elétrico de potência para o transporte de energia. Devido as suas características físicas, essas linhas estão suscetíveis a diversas perturbações possíveis, como descargas atmosféricas, contato com vegetação e queimadas. Tais perturbações podem acarretar em faltas, as quais causam fluxos anormais nas grandezas elétricas dos componentes da linha de transmissão e, conseqüentemente, afetam negativamente a qualidade do fornecimento de energia.

Frente a isso, as atividades de classificação e localização de falhas demonstram-se de suma importância, fornecendo informações cruciais aos operadores do SEP sobre quais fases estão envolvidas na falha, além da respectiva posição na LT. A partir deste diagnóstico, os operadores podem realizar a tomada de decisão com maior confiança e segurança, em especial para o religamento da linha e/ou o envio de uma equipe até o local identificado. Em resposta, apresentamos neste trabalho um método para a realização de ambas as atividades, buscando duas propriedades importantes: independência da taxa de amostragem e capacidade de generalização para linhas com diferentes configurações. Em complemento, também desenvolvemos e disponibilizamos duas bases de dados à comunidade científica, uma com séries temporais dos sinais de tensão e corrente (Seção 4.1) e outra com atributos extraídos destes sinais (Seção 4.2).

Os resultados obtidos para a classificação constatarem bons desempenhos da abordagem para todas as taxas de amostragem analisadas (256 Hz, 1 kHz, 5 kHz e 10 kHz), assim como para a generalização da classificação, cenário em que o algoritmo foi treinado com dados de uma determinada LT e avaliado com dados de uma linha distinta. Inclusive, nosso método foi capaz de classificar corretamente casos reais de falhas coletados da rede da COPEL, mesmo que o algoritmo tenha sido treinado unicamente com dados sintéticos. A acurácia alcançada nesta análise corresponde por volta de 96%. Grande parte destes eventos eram provenientes de diferentes LTs, cada qual com suas próprias características, como extensão e amplitude dos sinais de tensão e corrente.

Já para a localização de falta, propomos uma nova estrutura SDR para desempenhar tal atividade. Em particular, esta nova estrutura introduz um módulo de normalização incorporada, assim como um módulo de filtro que permite tornar as predições mais precisas. Além disso, consideramos em nosso esquema a análise de diversos hiperparâmetros que muitas vezes são negligenciados pelos trabalhos correlatos. Os resultados reportados também indicaram bons desempenhos para todas as taxas de amostragem avaliadas, principalmente para as frequências maiores, demonstrando a relevância de cada módulo e de cada hiperparâmetro do nosso esquema SDR. Por outro lado, o método não foi capaz de realizar efetivamente a generalização da localização. Isso pode ser justificado pelas falhas apresentarem assinaturas diferentes entre as LTs, mesmo considerando os mesmos parâmetros de falha, como tipo e resistência.

De qualquer maneira, os desempenhos alcançados são satisfatórios e atendem ao objetivo proposto, demonstrando que nosso método é versátil, efetivo e acurado. Ainda, os resultados manifestaram concordância em relação a nossa hipótese, de modo que os modelos múltiplos preditivos superaram o desempenho de algoritmos monolíticos tanto para a classificação quanto para a localização. Conforme demonstrado pelo experimentos abrangentes, a engenharia de atributos foi essencial para habilitar o método, particularmente para a classificação, para ambas as propriedades desejadas.

Logo, podemos afirmar que o presente trabalho contribui em diferentes perspectivas: socialmente, economicamente e cientificamente. As contribuições social e economicamente estão atreladas ao fato da aplicação do método em si, uma vez que pode permitir na diminuição do tempo de indisponibilidade de energia elétrica, além da atenuação de multas para as concessionárias. Inclusive, o componente classificador de falha está em operação junto a um sistema computacional para análise de falhas na COPEL. Por sua vez, a contribuição científica para a área da computação está vinculada, principalmente, a nova estrutura SDR que propomos. Apesar desta estrutura ter sido desenvolvida no contexto da localização de falhas, ela também pode ser empregada por qualquer outra aplicação de regressão.

Todavia, também pudemos identificar algumas limitações em nosso trabalho que podem afetar seu pleno desempenho, além de determinados cenários que seu emprego pode ser comprometido. As limitações encontradas estão descritas na Seção 7.1, enquanto as perspectivas para trabalhos futuros estão dispostas na Seção 7.2.

## 7.1 LIMITAÇÕES

As principais limitações deste trabalho são as que seguem:

- O desempenho do componente classificador de falta é seriamente dependente do atributo binário *groundDetection*, o qual objetiva auxiliar o algoritmo a identificar a presença ou não da terra na falha. Conforme discutimos na Subseção 6.1.4, caso a incidência ou não da terra seja incorretamente identificada, a predição do modelo poderá ser afetada;
- O componente localizador de falta não demonstrou a capacidade para a generalização para prever a distância que a falha ocorreu na LT. Assim, há a necessidade de que um modelo seja treinado individualmente para cada linha sob supervisão;
- A estrutura SDR que propomos requer, até o presente momento, a avaliação empírica dos hiperparâmetros para a determinação da melhor configuração, assim como qualquer algoritmo de AM;
- Ainda em relação a nossa estrutura SDR, a sua complexidade é outro fator que podemos atribuir como limitação, conforme debatido na Subseção 6.2.6.

## 7.2 TRABALHOS FUTUROS

Entre os trabalhos futuros identificados estão:

- Aperfeiçoar a técnica que determina o valor do atributo *groundDetection*, uma vez que possui uma influência notória ao algoritmo ERT na realização das predições de classificação do tipo de falta;
- Extrair novos atributos dos sinais de corrente e tensão para comporem uma nova versão da FADbF, utilizando para esta finalidade outras ferramentas de extração de características de séries temporais, como o tsfresh (Christ et al., 2018);
- Avaliar algoritmos de seleção de atributos nas características da base FADbF de modo a atingir melhores desempenhos para a localização de falhas, considerando tanto algoritmos de *ranking* quanto de avaliação de subconjuntos. O uso de algoritmos de avaliação de subconjuntos permite explorar correlação entre atributos que não é

capturada pelo *ranking*. Além disso, a seleção de atributos poderá ser *user-driven* ou *data-driven*, permitindo uma análise mais ampla como trabalho futuro;

- Desenvolver novas estratégias para determinar a RC, uma vez que ela é fundamental para toda e qualquer técnica de SDR, inclusive a que propomos;
- Avaliar um *pool* heterogêneo de regressores em busca de tanto diminuir o número de modelos no *pool* quanto melhorar o desempenho do método para a localização de faltas. Neste quesito, podemos mesclar algoritmos tradicionais de AM (*e.g.*, CART e SVM) junto a algoritmos de DL, como LSTM e CNN;
- Avaliar outros algoritmos de DL, como o Mamba (Gu e Dao, 2023);
- Outra perspectiva para a localização de falhas é investigar o uso dos algoritmos Transformers (Vaswani et al., 2017). Os Transformers têm demonstrado grande capacidade de modelagem para dependências em dados sequenciais e, portanto, são atraentes para a modelagem de séries temporais (Wen et al., 2023), no nosso caso, sinais trifásicos de tensão e corrente;
- Investigar o emprego da seleção de protótipos (Garcia et al., 2012), a qual objetiva alcançar um conjunto de dados reduzido, servindo como uma visão condensada de toda a base de dados original. Como resultado está a redução da complexidade da estrutura SDR e, eventualmente, na melhoria de seu desempenho (Kordos et al., 2019);
- Avaliar a generalização do método (*domain generalization* - Wang et al. (2023a)) de localização para uma outra linha que contenha a mesma extensão, porém com nível de tensão e demais configurações distintas daquela empregada para o treinamento do modelo. Essa análise visa conferir se o método é capaz de generalizar a localização ao menos para outras linhas de possuam comprimento similar a linha usada para o seu treinamento;
- Explorar novas técnicas e caminhos para habilitar a generalização do método também para a localização. Neste cenário, é possível que haja a necessidade de utilizar outras fontes de dados além de apenas os sinais de tensão e corrente.

## REFERÊNCIAS

- Aggarwal, C. C. (2015). *Data mining: the textbook*. Springer, Cham.
- Akke, M. e Thorp, J. T. (1998). Some improvements in the three-phase differential equation algorithm for fast transmission line protection. *IEEE Transactions on Power Delivery*, 13(1):66–72.
- Akmaz, D., Mamiş, M. S., Arkan, M. e Tağluk, M. E. (2018). Transmission line fault location using traveling wave frequencies and extreme learning machine. *Electric Power Systems Research*, 155:1–7.
- Aleem, S. A., Shahid, N. e Naqvi, I. H. (2015). Methodologies in power systems fault detection and diagnosis. *Energy Systems*, 6(1):85–108.
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT Press, Cambridge, 4 edition.
- Amorim, L. B., Cavalcanti, G. D. e Cruz, R. M. (2023). The choice of scaling technique matters for classification performance. *Appl Soft Comput*, 133:109924.
- ANEEL (2020). Avaliação do resultado da qualidade dos serviços de transmissão de energia elétrica. Relatório Técnico Relatório de Avaliação do Resultado Regulatório nº 2/2020/SRT/ANEEL, Agência Nacional de Energia Elétrica – ANEEL, Brasília - DF. Documento juntado a Nota Técnica nº 27/2020-SRT/ANEEL, Processo nº 48500.000174/2020-10.
- Asadi Majd, A., Samet, H. e Ghanbari, T. (2017). k-nn based fault detection and classification methods for power transmission systems. *Protection and Control of Modern Power Systems*, 2(32):1–11.
- Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T. e Gamboa, H. (2020). Tsfel: Time series feature extraction library. *SoftwareX*, 11:100456.
- Barboza, E. V. L., de Almeida, P. R. L., de Souza Britto, A. e Cruz, R. M. O. (2023). Distance functions and normalization under stream scenarios. Em *2023 International Joint Conference on Neural Networks (IJCNN)*, páginas 1–8.
- Belagoune, S., Bali, N., Bakdi, A., Baadji, B. e Atif, K. (2021). Deep learning through lstm classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. *Measurement*, 177:109330.
- Bhadra, A. B., Hasan, K., Islam, S. S., Sarker, N., Tama, I. J. e Khan, S. M. (2021). Robust short-circuit fault analysis scheme for overhead transmission line. Em *2021 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, páginas 104–107, Dhaka, Bangladesh.
- Bhatnagar, M., Yadav, A. e Swetapadma, A. (2023). Random forest regression-based fault location scheme for transmission lines. Em Dash, R. N., Rathore, A. K., Khadkikar, V., Patel, R. e Debnath, M., editores, *Smart Technologies for Power and Green Energy*, páginas 201–210, Singapore. Springer Nature Singapore.

- Bichels, A. (2018). *Sistemas elétricos de potência: métodos de análise e solução*. EDUTFPR, Curitiba.
- Blume, S. W. (2017). *Electric power system basics for the nonelectrical professional*. Wiley-IEEE Press, Piscataway, 2 edition.
- Bon, N. N. e Dai, L. V. (2022). Fault identification, classification, and location on transmission lines using combined machine learning methods. *International Journal of Engineering and Technology Innovation*, 12(2):91–109.
- Boudraa, A.-O. e Salzenstein, F. (2018). Teager–kaiser energy methods for signal and image analysis: A review. *Digital Signal Processing*, 78:338–375.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2):123–140.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.
- Breiman, L., Friedman, J., Stone, C. J. e Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman & Hall/CRC, Boca Raton.
- Brillinger, D. (2001). Time series: General. Em Smelser, N. J. e Baltes, P. B., editores, *International Encyclopedia of the Social & Behavioral Sciences*, páginas 15724–15731. Pergamon.
- Britto, A. S., Sabourin, R. e Oliveira, L. E. S. (2014). Dynamic selection of classifiers—a comprehensive review. *Pattern Recognition*, 47(11):3665–3680.
- Chen, K., Hu, J. e He, J. (2018a). Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder. *IEEE Transactions on Smart Grid*, 9(3):1748–1758.
- Chen, K., Huang, C. e He, J. (2016). Fault detection, classification and location for transmission lines and distribution systems: a review on the methods. *High Voltage*, 1(1):25–33.
- Chen, T. e Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Em *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, página 785–794, San Francisco, USA.
- Chen, Y. Q., Fink, O. e Sansavini, G. (2018b). Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction. *IEEE Transactions on Industrial Electronics*, 65(1):561–569.
- Chollet, F. e Allaire, J. J. (2018). *Deep learning with R*. Manning Publications, Shelter Island.
- Christ, M., Braun, N., Neuffer, J. e Kempa-Liehr, A. W. (2018). Time series feature extraction on basis of scalable hypothesis tests (tsfresh – a python package). *Neurocomputing*, 307:72–77.
- Cidrás, J., Miñambres, J. F. e Alvarado, F. L. (2011). Análise de faltas e proteção de sistemas elétricos. Em Gómez-Expósito, A., Conejo, A. J. e Cañizares, C., editores, *Sistemas de energia elétrica: análise e operação*, páginas 255–297. LTC.
- Clark, H. K. (1979). *Proteção de sistemas elétricos de potência*, volume 7. Universidade Federal de Santa Maria, Santa Maria.



- Coban, M. e Tezcan, S. S. (2021). Detection and classification of short-circuit faults on a transmission line using current signal. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, 69(4).
- Cordeiro, P. R., Cavalcanti, G. D. e Cruz, R. M. (2023a). A post-selection algorithm for improving dynamic ensemble selection methods. Em *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, páginas 1142–1147.
- Cordeiro, P. R. G., Cavalcanti, G. D. C. e Cruz, R. M. O. (2023b). Dynamic ensemble algorithm post-selection using hardness-aware oracle. *IEEE Access*, 11:86056–86070.
- Cruz, R. M., Oliveira, D. V., Cavalcanti, G. D. e Sabourin, R. (2019). Fire-des++: Enhanced online pruning of base classifiers for dynamic ensemble selection. *Pattern Recognition*, 85:149–160.
- Cruz, R. M., Sabourin, R. e Cavalcanti, G. D. (2018a). Dynamic classifier selection: Recent advances and perspectives. *Information Fusion*, 41:195–216.
- Cruz, R. M. O., Cavalcanti, G. D. C. e Ing Ren, T. (2011). A method for dynamic ensemble selection based on a filter and an adaptive distance to improve the quality of the regions of competence. Em *International Joint Conference on Neural Networks*, páginas 1126–1133, San Jose, USA.
- Cruz, R. M. O., Sabourin, R. e Cavalcanti, G. D. C. (2018b). Prototype selection for dynamic classifier and ensemble selection. *Neural Computing and Applications*, 29(2):447–457.
- da Silva, M., Oleskovicz, M. e Coury, D. V. (2005). Uma nova ferramenta baseada na transformada wavelet para localização digital de faltas. *Revista Controle & Automação*, 16(3):345–358.
- da Silveira, E. G. (2007). *Localização de faltas em linhas de transmissão: desenvolvimento de novos algoritmos e implementação de sistema computacional para aplicações práticas*. Tese de doutorado, Pós-Graduação em Engenharia Elétrica - Universidade Federal de Minas Gerais, Belo Horizonte - MG.
- de Oliveira, J. F. L., Silva, E. G. e de Mattos Neto, P. S. G. (2022). A hybrid system based on dynamic selection for time series forecasting. *IEEE Trans Neural Netw & Learn Syst*, 33(8):3251–3263.
- de Vasconcelos, F. M. (2017). *Geração, transmissão e distribuição de energia elétrica*. Editora e Distribuidora Educacional S.A., Londrina.
- Deza, M. M. e Deza, E. (2016). *Encyclopedia of Distances*. Springer, Heidelberg, 4 edition.
- Domingos, P. (1997). Why does bagging work? a bayesian account and its implications. Em *Third International Conference on Knowledge Discovery and Data Mining*, página 155–158, Newport Beach, USA.
- Dong, X., Kong, W. e Cui, T. (2009). Fault classification and faulted-phase selection based on the initial current traveling wave. *IEEE Transactions on Power Delivery*, 24(2):552–559.
- Duda, R. O., Hart, P. E. e Stork, D. G. (2000). *Pattern classification*. John Wiley & Sons, New York, 2 edition.



- Ekici, S., Yildirim, S. e Poyraz, M. (2008). Energy and entropy-based feature extraction for locating fault on transmission lines by using neural network and wavelet packet decomposition. *Expert Systems with Applications*, 34(4):2937–2944.
- Ensina, L. A., Karvat, P. E. M., Almeida, E. C. e Oliveira, L. E. S. (2022a). Fault location in transmission lines based on lstm model. Em *X Symposium on Knowledge Discovery, Mining and Learning*, páginas 1–8, Campinas, Brasil.
- Ensina, L. A., Oliveira, L. E. S., Almeida, E. C., Santos, S. L. F. e Bernardino, L. S. (2022b). Fault classification in transmission lines with generalization competence. Em *IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society*, páginas 1–6, Brussels, Belgium.
- Ensina, L. A., Oliveira, L. E. S., Cruz, R. M. O. e Cavalcanti, G. D. C. (2023). Fault Analysis Database with Features (FADbF). <https://dx.doi.org/10.5281/zenodo.10275032>. Acessado em 06/12/2023.
- Ensina, L. A., Oliveira, L. E. S., Cruz, R. M. O. e Cavalcanti, G. D. C. (2024). Fault distance estimation for transmission lines with dynamic regressor selection. *Neural Computing and Applications*, 36(4):1741–1759.
- Ertel, W. (2017). *Introduction to Artificial Intelligence*. Springer, Cham, 2 edition.
- Esling, P. e Agon, C. (2012). Time-series data mining. *ACM Computing Surveys*, 45(1).
- Faceli, K., Lorena, A. C., Gama, J., de Almeida, T. A. e de Carvalho, A. C. P. L. F. (2021). *Inteligência artificial: uma abordagem de aprendizado de máquina*. LTC, Rio de Janeiro, 2 edition.
- Fan, R., Yin, T., Huang, R., Lian, J. e Wang, S. (2019). Transmission line fault location using deep learning techniques. Em *2019 North American Power Symposium (NAPS)*, páginas 1–5, Wichita, USA.
- Fang, J., Chen, K., Liu, C. e He, J. (2023). An explainable and robust method for fault classification and location on transmission lines. *IEEE Transactions on Industrial Informatics*, 19(10):10182–10191.
- Faulkenberry, L. M. e Coffey, W. (1996). *Electrical Power Distribution and Transmission*. Prentice Hall, Upper Saddle River.
- Fei, C., Qi, G. e Li, C. (2018). Fault location on high voltage transmission line by applying support vector regression with fault signal amplitudes. *Electric Power Systems Research*, 160:173–179.
- Ferreira, V. H., Zanghi, R., Fortes, M. Z., Júnior, S. G. e Alves da Silva, A. P. (2020). Probabilistic transmission line fault diagnosis using autonomous neural models. *Electric Power Systems Research*, 185:1–10.
- Ferreira, V. H., Zanghi, R., Fortes, M. Z., Sotelo, G. G., da Boa Morte Silva, R., de Souza, J. C. S., Guimarães, C. H. C. e Júnior, S. G. (2016). A survey on intelligent system application to fault diagnosis in electric power system transmission lines. *Electric Power Systems Research*, 136:135–153.

- Fleming, B. (1998). Negative-sequence impedance directional element. Em *10th Annual ProTest User Group Meeting*, páginas 1–11, Pasadena, USA.
- Fonseca, G. A., Ferreira, D. D., Costa, F. B. e Almeida, A. R. (2022). Fault classification in transmission lines using random forest and notch filter. *Journal of Control, Automation and Electrical Systems*, 33(2):598–609.
- Fortescue, C. L. (1918). Method of symmetrical co-ordinates applied to the solution of polyphase networks. *Transactions of the American Institute of Electrical Engineers*, 37(2):1027–1140.
- Fuchs, R. D. (1977). *Transmissão de energia elétrica: linhas aéreas; teoria das linhas em regime permanente*, volume 1. LTC, Rio de Janeiro.
- Furse, C. M., Kafal, M., Razzaghi, R. e Shin, Y.-J. (2021). Fault diagnosis for electrical systems and power networks: A review. *IEEE Sensors Journal*, 21(2):888–906.
- Gafoor, S. A. e Rao, P. V. R. (2006). Wavelet based fault detection, classification and location in transmission lines. Em *IEEE International Power and Energy Conference*, páginas 114–118, Putra Jaya, Malaysia.
- Garcia, S., Derrac, J., Cano, J. e Herrera, F. (2012). Prototype selection for nearest neighbor classification: Taxonomy and empirical study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3):417–435.
- Geurts, P., Ernst, D. e Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63:3–42.
- Gilbert, D. e Morrison, I. (1997). A statistical method for the detection of power system faults. *International Journal of Electrical Power & Energy Systems*, 19(4):269–275.
- Godse, R. e Bhat, S. (2020). Mathematical morphology-based feature-extraction technique for detection and classification of faults on power transmission line. *IEEE Access*, 8:38459–38471.
- Goh, H. H., Sim, S. V., Mohamed, M. A. H., Rahman, A. K. A., Ling, C. W., Chua, Q. S. e Goh, K. C. (2017). Fault location techniques in electrical power system: a review. *Indonesian Journal of Electrical Engineering and Computer Science*, 8(1):206–212.
- Grainger, J. J., Stevenson, W. D. e Chang, G. W. (2016). *Power System Analysis*. McGraw-Hill Education, New York.
- Gu, A. e Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces.
- Gururajapathy, S., Mokhlis, H. e Illias, H. (2017). Fault location and detection techniques in power distribution systems with distributed generation: a review. *Renewable and Sustainable Energy Reviews*, 74:949–958.
- Géron, A. (2021). *Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras & TensorFlow*. Alta Books, Rio de Janeiro, 2 edition.
- Haginomori, E., Koshiduka, T., Arai, J. e Ikeda, H. (2016). *Power System Transient Analysis: Theory and Practice using Simulation Programs (ATP-EMTP)*. Wiley, Chichester.

- Haq, E. U., Jianjun, H., Li, K., Ahmad, F., Banjerdpongchai, D. e Zhang, T. (2021). Improved performance of detection and classification of 3-phase transmission line faults based on discrete wavelet transform and double-channel extreme learning machine. *Electrical Engineering*, 103:953–963.
- Hedman, D. E. (1979). *Teoria das linhas de transmissão - I*, volume 2. Universidade Federal de Santa Maria, Santa Maria.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):832–844.
- Huang, G.-B., Zhu, Q.-Y. e Siew, C.-K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1):489–501.
- Høidalen, H. K., Prikler, L. e Peñaloza, F. (2019). *ATPDraw version 7.0 for Windows - Users' Manual*.
- Igual, L. e Seguí, S. (2017). *Introduction to Data Science: a python approach to concepts, techniques and applications*. Springer, Cham.
- Ishwaran, H. (2015). The effect of splitting on random forests. *Machine Learning*, 99(1):75–118.
- Javed, K., Gouriveau, R. e Zerhouni, N. (2014). Sw-elm: A summation wavelet extreme learning machine algorithm with a priori parameter initialization. *Neurocomputing*, 123:299–307.
- Kanagasabapathy, O. (2021). Fault location in transmission line through deep learning - a systematic review. Em Suma, V., Chen, J. I.-Z., Baig, Z. e Wang, H., editores, *Inventive Systems and Control*, páginas 223–238, Singapore.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. e Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. Em *31st International Conference on Neural Information Processing Systems*, página 3149–3157, Long Beach, USA.
- Kezunovic, M. e Rikalo, I. (1996). Detect and classify faults using neural nets. *IEEE Computer Applications in Power*, 9(4):42–47.
- Kindermann, G. (2018). *Curto-circuito*. Edição do autor, 6 edition.
- Kitchenham, B. e Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report, v2.3.
- Kononenko, I. e Kukar, M. (2007). *Machine learning and data mining: introduction to principles and algorithms*. Horwood Publishing, Chichester.
- Kordos, M., Álvar Arnaiz-González e García-Osorio, C. (2019). Evolutionary prototype selection for multi-output regression. *Neurocomputing*, 358:309–320.
- Krogh, A. e Vedelsby, J. (1994). Neural network ensembles, cross validation and active learning. Em *7th International Conference on Neural Information Processing Systems*, página 231–238, Cambridge, USA.
- Kubat, M. (2017). *An introduction to machine learning*. Springer, Cham, 2 edition.

- Kumar, A. e Bhagat, S. K. (2019). Study of stability analysis using lyapunov function for ieee-9 bus system. Em *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, páginas 1–6, Coimbatore, India.
- Kuncheva, L. (2002). A theoretical study on six classifier fusion strategies. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):281–286.
- Kuncheva, L. I. (2014). *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons, Hoboken, 2 edition.
- Labegalini, P. R., Labegalini, J. A., Fuchs, R. D. e de Almeida, M. T. (1992). *Projetos mecânicos das linhas aéreas de transmissão*. Edgard Blucher, São Paulo, 2 edition.
- LeCun, Y., Bengio, Y. e Hinton, G. (2015). Deep learning. *Nature*, 521:436–444.
- Liang, H., Han, X., Yu, H., Li, F., Liu, Z. e Zhang, K. (2022). Transmission line fault-cause identification method for large-scale new energy grid connection scenarios. *Global Energy Interconnection*, 5(4):362–374.
- Loh, W.-Y. (2011). Classification and regression trees. *WIREs Data Mining and Knowledge Discovery*, 1(1):14–23.
- Lopes, F. V., Ribeiro, E. P., Reis, R. L., Silva, K. M., Britto, A. M., Moraes, C. M., Agostinho, R. L. e Rodrigues, M. A. (2023). Three-parameter atp/atpdraw transmission line high impedance fault model. *Electric Power Systems Research*, 223:109577.
- Luo, X. e Kezunovic, M. (2005). Fault analysis based on integration of digital relay and dfr data. Em *IEEE Power Engineering Society General Meeting*, páginas 746–751, San Francisco, USA.
- Mallaki, M. e Dashti, R. (2012). Fault locating in transmission networks using transient voltage data. *Energy Procedia*, 14:173–180.
- Mallat, S. (2009). *A wavelet tour of signal processing: the sparse way*. Springer, Burlington, 3 edition.
- Mamede Filho, J. e Mamede, D. R. (2020). *Proteção de sistemas elétricos de potência*. LTC, Rio de Janeiro, 2 edition.
- Mendes-Moreira, J., Jorge, A. M., Freire de Sousa, J. e Soares, C. (2015). Improving the accuracy of long-term travel time prediction using heterogeneous ensembles. *Neurocomputing*, 150:428–439.
- Mendes-Moreira, J., Jorge, A. M., Soares, C. e de Sousa, J. F. (2009). Ensemble learning: A study on different variants of the dynamic selection approach. Em Perner, P., editor, *Machine Learning and Data Mining in Pattern Recognition*, páginas 191–205.
- Mendes-Moreira, J. a., Soares, C., Jorge, A. M. e Sousa, J. F. D. (2012). Ensemble approaches for regression: A survey. *ACM Computing Surveys*, 45(1).
- Mirzaei, M., Vahidi, B. e Hosseinian, S. H. (2018). Fault location on a series-compensated three-terminal transmission line using deep neural networks. *IET Science, Measurement & Technology*, 12(6):746–754.

- Mishra, D. e Ray, P. (2018). Fault detection, location and classification of a transmission line. *Neural Computing and Applications*, 30:1377–1424.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Moore, D. G. (2019). Pyinform. Acessado em 18/08/2023.
- Morais, J., Pires, Y., Cardoso, C. e Klautau, A. (2010). A framework for evaluating automatic classification of underlying causes of disturbances and its application to short-circuit faults. *IEEE Transactions on Power Delivery*, 25(4):2083–2094.
- Moreto, M. e Rolim, J. G. (2010). Análise automática de oscilografias em sistemas elétricos de potência. *Revista Controle & Automação*, 21(4):347–362.
- Moura, A. P., Moura, A. A. F. e Rocha, E. P. (2019a). *Engenharia de sistemas de potência: transmissão de energia elétrica em corrente alternada*. UFC, Fortaleza.
- Moura, T. J., Cavalcanti, G. D. e Oliveira, L. S. (2021). Mine: A framework for dynamic regressor selection. *Information Sciences*, 543:157–179.
- Moura, T. J. M. (2019). *MINE - a framework for dynamic regressor selection*. Tese de doutorado, Pós-Graduação em Ciência da Computação - Universidade Federal de Pernambuco, Recife - PE.
- Moura, T. J. M., Cavalcanti, G. D. C. e Oliveira, L. S. (2019b). Evaluating competence measures for dynamic regressor selection. Em *International Joint Conference on Neural Networks (IJCNN)*, páginas 1–8, Budapest, Hungary.
- Moura, T. J. M., Cavalcanti, G. D. C. e Oliveira, L. S. (2020). On the selection of the competence measure for dynamic regressor selection. Em *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, páginas 1630–1637, Toronto, Canada.
- Mukherjee, A., Kundu, P. K. e Das, A. (2021). Transmission line faults in power system and the different algorithms for identification, classification and localization: A brief review of methods. *Journal of The Institution of Engineers (India): Series B*, páginas 1–23.
- Panda, S., Mishra, D. P. e Dash, S. N. (2018). Comparison of anfis and ann techniques in fault classification and location in long transmission lines. Em *International Conference on Recent Innovations in Electrical, Electronics Communication Engineering (ICRIEECE)*, páginas 1112–1117, Bhubaneswar, India.
- Patel, B., Bera, P. e Saha, B. (2018). Wavelet packet entropy and rbfnn based fault detection, classification and localization on hvac transmission line. *Electric Power Components and Systems*, 46(1):15–26.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. e Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Pinto, M. O. (2014). *Energia elétrica: geração, transmissão e sistemas interligados*. LTC, Rio de Janeiro, 1 edition.



- Pradhan, A. K. (2010). Intelligent techniques for transmission line fault classification. Em Panigrahi, B. K., Abraham, A. e Das, S., editores, *Computational Intelligence in Power Engineering. Studies in Computational Intelligence*, volume 302, páginas 75–101. Springer.
- Prasad, A., Belwin Edward, J. e Ravi, K. (2018). A review on fault classification methodologies in power transmission systems: Part—i. *Journal of Electrical Systems and Information Technology*, 5(1):48–60.
- Python (2023). statistics — mathematical statistics functions. Acessado em 18/08/2023.
- Pérez-Arriaga, I. J., Rudnick, H. e Abbad, M. R. (2011). Sistemas de energia elétrica: uma visão geral. Em Gómez-Expósito, A., Conejo, A. J. e Cañizares, C., editores, *Sistemas de energia elétrica: análise e operação*, páginas 1–39. LTC.
- Qiao, Z. e Wang, B. (2021). Molten steel temperature prediction in ladle furnace using a dynamic ensemble for regression. *IEEE Access*, 9:18855–18866.
- Quinlan, J. (1987). Simplifying decision trees. *International Journal of Man-Machine Studies*, 27(3):221–234.
- Quinlan, J. R. (1979). Discovering rules by induction from large collections of examples. *Expert systems in the micro electronics age*.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1):81–106.
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- Rao, K. D. (2018). *Signals and Systems*. Birkhäuser, Cham, 1 edition.
- Ray, P. e Mishra, D. P. (2016). Support vector machine based fault classification and location of a long transmission line. *Engineering Science and Technology, an International Journal*, 19(3):1368–1380.
- Ray, P., Panigrahi, B. K. e Senroy, N. (2013). Hybrid methodology for fault distance estimation in series compensated transmission line. *IET Generation, Transmission & Distribution*, 7(5):431–439.
- Raza, A., Benrabah, A., Alquthami, T. e Akmal, M. (2020). A review of fault diagnosing methods in power transmission systems. *Applied Sciences*, 10(4):1312.
- Resende, L. e Silveira, E. (2016). Modeling of electrical protection functions of hydropower plants using the atp. *IEEE Latin America Transactions*, 14(3):1265–1270.
- Rich, E., Knight, K. e Nair, S. B. (2009). *Artificial Intelligence*. McGraw-Hill, New Delhi, 3 edition.
- Rodrigues, W. M., Walmsley, F. N., Cavalcanti, G. D. C. e Cruz, R. M. O. (2023). Security relevant methods of android’s api classification: A machine learning empirical evaluation. *IEEE Transactions on Computers*, 72(11):3273–3285.
- Rooney, N. e Patterson, D. (2007). A weighted combination of stacking and dynamic integration. *Pattern Recognition*, 40(4):1385–1388.

- Rooney, N., Patterson, D., Anand, S. e Tsymbal, A. (2004). Dynamic integration of regression models. Em Roli, F., Kittler, J. e Windeatt, T., editores, *Multiple Classifier Systems*, páginas 164–173, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Russell, S. e Norvig, P. (2020). *Artificial Intelligence: a modern approach*. Pearson, Upper Saddle River, 4 edition.
- Ruta, D. e Gabrys, B. (2005). Classifier selection for majority voting. *Information Fusion*, 6(1):63–81.
- Saha, M. M., Izykowski, J. e Rosolowski, E. (2010). *Fault location on power networks*. Springer, New York.
- Sammut, C. e Webb, G. I., editores (2017). *Random Subspace Method*, páginas 1055–1055. Springer.
- Sani, U. S., Malik, O. A. e Lai, D. T. C. (2022). Dynamic regressor/ensemble selection for a multi-frequency and multi-environment path loss prediction. *Information*, 13(11).
- Sato, F. e Freitas, W. (2015). *Análise de curto-circuito e princípios de proteção em sistemas de energia elétrica*. Elsevier, Rio de Janeiro.
- Schapire, R. E., Freund, Y., Bartlett, P. e Lee, W. S. (1998). Boosting the margin: a new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5):1651–1686.
- Schneider, J., Wenig, P. e Papenbrock, T. (2021). Distributed detection of sequential anomalies in univariate time series. *The VLDB Journal*, 30(4):579–602.
- Schuster, M. e Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681.
- Sergio, A. T., de Lima, T. P. e Ludermir, T. B. (2016). Dynamic selection of forecast combiners. *Neurocomputing*, 218:37–50.
- Shakiba, F. M., Azizi, S. M., Zhou, M. e Abusorrah, A. (2023). Application of machine learning methods in fault detection and classification of power transmission lines: a survey. *Artificial Intelligence Review*, 56(7):5799–5836.
- Shakiba, F. M., Shojaee, M., Azizi, S. M. e Zhou, M. (2022). Generalized fault diagnosis method of transmission lines using transfer learning technique. *Neurocomputing*, 500:556–566.
- Silva, E. G., De Mattos Neto, P. S. G. e Cavalcanti, G. D. C. (2021). A dynamic predictor selection method based on recent temporal windows for time series forecasting. *IEEE Access*, 9:108466–108479.
- Silva, K. M., Souza, B. A. e Brito, N. S. D. (2006). Fault detection and classification in transmission lines based on wavelet transform and ann. *IEEE Transactions on Power Delivery*, 21(4):2058–2063.
- Silva, K. M., Souza, B. A. d., Brito, N. S. D., Dantas, K. M. C., Costa, F. B. e Silva, S. S. B. (2007). Detecção e classificação de faltas a partir da análise de registros oscilográficos via redes neurais artificiais e transformada wavelet. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, 18(2):163–172.



- Singh, D. e Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97:105524.
- Singh, S. e Vishwakarma, D. N. (2015). Intelligent techniques for fault diagnosis in transmission lines — an overview. Em *2015 International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE)*, páginas 280–285, Noida, India.
- Smith, S. W. (2002). *Digital Signal Processing: A Practical Guide for Engineers and Scientists*. Newnes, Boston.
- Stapor, K., Ksieniewicz, P., García, S. e Woźniak, M. (2021). How to design the fair experimental classifier evaluation. *Applied Soft Computing*, 104:107219.
- Stevenson, W. D. (1986). *Elementos de análise de sistemas de potência*. McGraw-Hill, São Paulo, 2 edition.
- Swaminathan, R., Mishra, S., Routray, A. e Swain, S. C. (2021). A cnn-lstm-based fault classifier and locator for underground cables. *Neural Computing and Applications*, 33(22):15293–15304.
- Swetapadma, A. e Yadav, A. (2015). All shunt fault location including cross-country and evolving faults in transmission lines without fault type classification. *Electric Power Systems Research*, 123:1–12.
- Swetapadma, A. e Yadav, A. (2017). A novel decision tree regression-based fault distance estimation scheme for transmission lines. *IEEE Transactions on Power Delivery*, 32(1):234–245.
- Swetapadma, A. e Yadav, A. (2018). A novel single-ended fault location scheme for parallel transmission lines using k-nearest neighbor algorithm. *Computers & Electrical Engineering*, 69:41–53.
- Tamalu, N., Ensina, L. A., Almeida, E. C., Pena, E. H. M. e Oliveira, L. E. S. (2023). Fault detection in transmission lines: a denial constraint approach. Em *XXXVIII Simpósio Brasileiro de Bancos de Dados*, páginas 231–243, Belo Horizonte, Brasil.
- Tang, Y., Wang, H., Aggarwal, R. e Johns, A. (2000). Fault indicators in transmission and distribution systems. Em *International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, páginas 238–243, London, UK.
- Ueda, N. e Nakano, R. (1996). Generalization error of ensemble estimators. Em *International Conference on Neural Networks (ICNN'96)*, páginas 90–95, Washington, USA.
- Valabhoju, A., Yadav, A., Pazoki, M. e El-Sehiemy, R. A. (2021). Optimized ensemble of regression tree-based location of evolving faults in dual-circuit line. *Neural Computing and Applications*, 33(14):8795–8820.
- Vasantharathna, S. (2016). Electric power systems. Em Rashid, M. H., editor, *Electric Renewable Energy Systems*, páginas 403–456. Academic Press.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. e Polosukhin, I. (2017). Attention is all you need. Em *31st Conference on Neural Information Processing Systems (NeurIPS)*, Long Beach, USA.

- Vilalta, R. e Drissi, Y. (2002). A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95.
- Wang, B. e Mao, Z. (2020). A dynamic ensemble outlier detection model based on an adaptive k-nearest neighbor rule. *Information Fusion*, 63:30–40.
- Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., Chen, Y., Zeng, W. e Yu, P. S. (2023a). Generalizing to unseen domains: A survey on domain generalization. *IEEE Transactions on Knowledge and Data Engineering*, 35(8):8052–8072.
- Wang, X., Zhou, P., Peng, X., Wu, Z. e Yuan, H. (2022). Fault location of transmission line based on cnn-lstm double-ended combined model. *Energy Reports*, 8:781–791. ICPE 2021 - The 2nd International Conference on Power Engineering.
- Wang, Y., Cui, Q., Weng, Y., Li, D. e Li, W. (2023b). Learning picturized and time-series data for fault location with renewable energy sources. *International Journal of Electrical Power & Energy Systems*, 147:108853.
- Wang, Z. e Oates, T. (2015). Imaging time-series to improve classification and imputation. Em *24th International Conference on Artificial Intelligence*, página 3939–3945, Buenos Aires, Argentina.
- Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J. e Sun, L. (2023). Transformers in time series: A survey. Em *Thirty-Second International Joint Conference on Artificial Intelligence – IJCAI-23*, páginas 6778–6786, Macao, China.
- Witten, I. H. e Frank, E. (2017). *Data Mining: practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 4 edition.
- Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2):241–259.
- Yadav, A. e Dash, Y. (2014). An overview of transmission line protection by artificial neural network: Fault detection, fault classification, fault location, and fault direction discrimination. *Advances in Artificial Neural Systems*, 2014:1–21.
- Ye, L. e Campbell, A. (2006). Behavior investigations of superconducting fault current limiters in power systems. *IEEE Transactions on Applied Superconductivity*, 16(2):662–665.
- Ye, W., Jian, S., Ou, R., Huang, S., Gong, X., Peng, X. e Yuan, H. (2020). Fault classification of high voltage transmission line based on convolutional neural network. Em *10th International Conference on Information Science and Technology (ICIST)*, páginas 294–300, Bath, London, and Plymouth, UK.
- Zhang, F., Liu, Q., Liu, Y., Tong, N., Chen, S. e Zhang, C. (2020). Novel fault location method for power systems based on attention mechanism and double structure gru neural network. *IEEE Access*, 8:75237–75248.
- Zheng, A. e Casari, A. (2018). *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. O’Reilly.
- Zhou, Z.-H. (2012). *Ensemble Methods: foundations and algorithms*. CRC Press, Boca Raton, 1 edition.
- Zumbahlen, H. (2008). *Linear Circuit Design Handbook*. Newnes, Burlington.

## APÊNDICE A – LEVANTAMENTO BIBLIOGRÁFICO

O objetivo deste levantamento bibliográfico foi identificar os métodos propostos na literatura para as tarefas de detecção, classificação e localização de faltas em linhas de transmissão considerando dados de oscilografias, com foco para o uso de algoritmos de aprendizagem de máquina para tais tarefas. Para capturar um panorama replicável e abrangente sobre o tema, conduzimos um processo de levantamento que segue algumas características do método de revisão sistemática da literatura (Kitchenham e Charters, 2007).

Assim sendo, a questão de pesquisa definida foi: *"Quais métodos são empregados para a detecção, a classificação e a localização de faltas em linhas de transmissão utilizando dados de oscilografias por meio de algoritmos de aprendizagem de máquina?"*.

As buscas foram conduzidas para as bases de dados ACM, IEEE Xplore, SciELO e Scopus a partir da seguinte expressão de busca: (fault\*) AND (classification OR detection OR location OR diagnosis OR analysis OR distance) AND ("transmission line" OR "transmission lines" OR "transmission system" OR "transmission systems" OR "transmission network" OR "transmission networks").

Em complemento, as buscas compreenderam a expressão de busca para os campos de título, palavras-chave (*keywords*) e resumo. Os trabalhos encontrados nas bases de dados foram submetidos a um processo de filtragem, de modo a selecionar as obras mais relevantes, bem como remover principalmente aquelas sem relação com os objetivos da pesquisa. Dessa maneira, dois Critérios de Inclusão (CI) foram definidos:

**CI1** Uso de aprendizado de máquina na abordagem;

**CI2** Trabalhos a partir de 2018.

Apenas os artigos publicados a partir de 2018 foram selecionados pois na literatura já consta outras revisões e mapeamentos que relatam principalmente trabalhos anteriores a essa data (Yadav e Dash, 2014; Aleem et al., 2015; Singh e Vishwakarma, 2015; Chen et al., 2016; Ferreira et al., 2016; Mishra e Ray, 2018; Prasad et al., 2018; Raza et al., 2020; Mukherjee et al., 2021; Kanagasabapathy, 2021; Shakiba et al., 2023). Cabe destacar que os trabalhos citados pelas respectivas revisões e mapeamentos também foram considerados.

Já os Critérios de Exclusão (CE) foram:

**CE1** Não possui relação com a questão de pesquisa;

**CE2** Não é do idioma inglês, português ou espanhol;

**CE3** Não descreve o procedimento adotado para treinamento, validação e teste;

**CE4** Não apresenta a configuração da linha de transmissão avaliada;

**CE5** Não realiza a avaliação quantitativa do método;

**CE6** Trabalhos que sejam patentes, relatórios técnicos, livros, pôsteres e similares;

**CE7** Trabalho duplicado.

Ainda, os procedimentos adotados para a seleção dos artigos foram realizados em duas etapas:

- 1<sup>o</sup> **filtro**: leitura de título e resumo;
- 2<sup>o</sup> **filtro**: leitura do texto completo.

Assim sendo, os artigos foram primeiramente filtrados pela leitura de título e resumo (1<sup>o</sup> filtro), e, a partir dos artigos selecionados, foram novamente analisados pela leitura do texto completo (2<sup>o</sup> filtro). Os trabalhos resultantes foram utilizados para a extração das informações para compor os resultados deste mapeamento e respaldar nossas contribuições.

As principais informações extraídas dos trabalhos estão apresentadas a seguir:

- Atividades realizadas: detecção, classificação e/ou localização de falta;
- Configuração da(s) linha(s) de transmissão avaliada(s) (comprimento, tensão, frequência, etc);
- Sinais utilizados (corrente, tensão);
- Número de ciclos utilizados;
- Taxa(s) de amostragem avaliada(s);
- Técnica de processamento de sinais utilizada;
- Atributos extraídos e utilizados;
- Simulador utilizado para a geração das simulações;
- Base de dados de falhas utilizado/proposto, se ocorrer;
- Parâmetros de falta utilizados para gerar as simulações;
- Quantidade de dados gerados e destinados para treinamento, validação e teste;
- Método com aprendizagem de máquina utilizado para detecção, classificação e/ou localização de falta;
- Utilização de dados de falta reais e, se houver, indicar quais e como os dados foram empregados;
- Utilização de dados ambientais, se houver;
- Avaliação do método para diferentes taxas de amostragem;
- Avaliação do método para linha(s) de transmissão distinta(s) daquela(s) utilizada(s) para o treinamento (generalização);
- Métricas de desempenho utilizadas e os respectivos resultados.

Em complemento, outras informações adicionais também foram coletadas, como o país do primeiro autor, o ano de publicação e o periódico/congresso em que o trabalho foi publicado. É dito de nota que publicações citadas nas referências bibliográficas dos trabalhos selecionados e que não tenham sido identificadas durante as buscas também foram consideradas, desde que não violassem os critérios de seleção e exclusão.

## APÊNDICE B – ANÁLISE DOS ATRIBUTOS PARA A LOCALIZAÇÃO

Na Tabela B.1 estão apresentadas as características ordenadas individualmente por suas importâncias para o algoritmo ERT considerando a medida de Gini.

Tabela B.1: Características ordenadas por suas importâncias para o algoritmo ERT, considerando a média e o desvio padrão, entre parênteses, da medida de Gini calculada a partir dos dez modelos ERT pré-treinados

<b>Ranking</b>	<b>Atributo</b>	<b>Gini score</b>
#1	Desvio padrão – fase A (sinal da corrente)	4,4040E-02 (2,5690E-03)
#2	Desvio padrão – fase B (sinal da corrente)	4,3040E-02 (2,5440E-03)
#3	Desvio padrão – fase C (sinal da corrente)	4,2170E-02 (2,0330E-03)
#4	Variância – fase A (sinal da corrente)	3,0150E-02 (1,4900E-03)
#5	Variância – fase B (sinal da corrente)	2,9760E-02 (1,9630E-03)
#6	Variância – fase C (sinal da corrente)	2,9560E-02 (1,8470E-03)
#7	Média geométrica – fase C (sinal da corrente)	2,6990E-02 (2,5860E-03)
#8	Média geométrica – fase A (sinal da corrente)	2,4390E-02 (3,4570E-03)
#9	Média geométrica – fase B (sinal da corrente)	2,4000E-02 (2,1580E-03)
#10	Variância – fase C (sinal da tensão)	2,2730E-02 (1,3000E-03)
#11	Variância – fase B (sinal da tensão)	2,1720E-02 (1,0930E-03)
#12	Variância – fase A (sinal da tensão)	2,1620E-02 (1,3830E-03)
#13	Covariância – fases B-C (sinal da corrente)	2,0550E-02 (7,7560E-04)
#14	Desvio padrão – fase B (sinal da tensão)	2,0010E-02 (1,2140E-03)
#15	Desvio padrão – fase C (sinal da tensão)	2,0000E-02 (7,7940E-04)
#16	Covariância – fases A-C (sinal da corrente)	1,9870E-02 (7,9760E-04)
#17	Covariância – fases A-B (sinal da corrente)	1,9620E-02 (1,3880E-03)
#18	Desvio padrão – fase A (sinal da tensão)	1,8800E-02 (1,5140E-03)
#19	Distância – fase C (sinal da tensão)	1,8510E-02 (6,5940E-04)
#20	Distância – fase B (sinal da tensão)	1,7080E-02 (6,8470E-04)
#21	Entropia – fase B (sinal da tensão)	1,6910E-02 (1,0860E-03)
#22	Distância – fase A (sinal da tensão)	1,6360E-02 (7,3220E-04)
#23	Entropia – fase C (sinal da tensão)	1,6180E-02 (2,1140E-03)
#24	Entropia – fase A (sinal da tensão)	1,5840E-02 (2,1650E-03)
#25	AUC – fase C (sinal da corrente)	1,5190E-02 (1,5090E-03)
#26	AUC – fase A (sinal da corrente)	1,4670E-02 (1,4050E-03)
#27	AUC – fase B (sinal da corrente)	1,4120E-02 (1,7460E-03)
#28	Entropia – fase B (sinal da corrente)	1,3980E-02 (4,6120E-04)
#29	Entropia – fase C (sinal da corrente)	1,3930E-02 (5,7060E-04)
#30	Entropia – fase A (sinal da corrente)	1,3920E-02 (3,8500E-04)
#31	RMS – fase C (sinal da corrente)	1,3790E-02 (9,0440E-04)
#32	RMS – fase B (sinal da corrente)	1,2290E-02 (1,5450E-03)
#33	RMS – fase A (sinal da corrente)	1,2190E-02 (1,2220E-03)
#34	RMS – fase A (sinal da tensão)	1,1050E-02 (7,5330E-04)
#35	Correlação – fases A-C (sinal da tensão)	1,0970E-02 (7,6940E-04)
#36	Energia – fase A (sinal da tensão)	1,0880E-02 (8,7400E-04)
#37	Correlação – fases A-B (sinal da tensão)	1,0830E-02 (1,1480E-03)

Continua na próxima página



Tabela B.1: Continuação da página anterior

<b>Ranking</b>	<b>Atributo</b>	<b>Gini score</b>
#38	Energia – fase C (sinal da tensão)	1,0720E-02 (5,0680E-04)
#39	RMS – fase C (sinal da tensão)	1,0710E-02 (6,5260E-04)
#40	Energia – fase B (sinal da tensão)	1,0340E-02 (1,0820E-03)
#41	Correlação – fases B-C (sinal da tensão)	1,0300E-02 (6,4530E-04)
#42	RMS – fase B (sinal da tensão)	1,0190E-02 (1,0160E-03)
#43	Covariância – fases A-C (sinal da tensão)	9,1210E-03 (5,3660E-04)
#44	Distância pico a pico – fase B (sinal da corrente)	9,0000E-03 (8,8100E-04)
#45	Covariância – fases B-C (sinal da tensão)	8,9590E-03 (3,6130E-04)
#46	Covariância – fases A-B (sinal da tensão)	8,9010E-03 (3,6090E-04)
#47	Distância pico a pico – fase C (sinal da corrente)	8,7550E-03 (1,3090E-03)
#48	Distância pico a pico – fase A (sinal da corrente)	8,3270E-03 (1,1790E-03)
#49	Mínimo – fase B (sinal da corrente)	8,1260E-03 (6,1190E-04)
#50	Energia – fase C (sinal da corrente)	7,9990E-03 (1,0010E-03)
#51	Energia – fase A (sinal da corrente)	7,7880E-03 (7,5660E-04)
#52	Máximo – fase C (sinal da corrente)	6,9950E-03 (1,1780E-03)
#53	Energia – fase B (sinal da corrente)	6,9590E-03 (1,1480E-03)
#54	Mínimo – fase A (sinal da corrente)	6,7800E-03 (6,1190E-04)
#55	Mínimo – fase C (sinal da corrente)	6,2530E-03 (9,8160E-04)
#56	Máximo – fase A (sinal da corrente)	6,2400E-03 (7,0730E-04)
#57	Máximo – fase B (sinal da corrente)	6,2390E-03 (1,2000E-03)
#58	Correlação – fases A-B (sinal da corrente)	5,1500E-03 (2,9130E-04)
#59	Correlação – fases A-C (sinal da corrente)	5,0350E-03 (2,3570E-04)
#60	Correlação – fases B-C (sinal da corrente)	4,9310E-03 (3,0080E-04)
#61	Distância pico a pico – fase B (sinal da tensão)	3,8960E-03 (7,4150E-04)
#62	Média geométrica – fase A (sinal da tensão)	3,8690E-03 (2,1810E-04)
#63	Distância – fase C (sinal da corrente)	3,6130E-03 (3,1810E-04)
#64	Distância – fase A (sinal da corrente)	3,4940E-03 (3,9280E-04)
#65	Distância – fase B (sinal da corrente)	3,3930E-03 (1,5630E-04)
#66	Distância pico a pico – fase C (sinal da tensão)	3,3870E-03 (5,5380E-04)
#67	Média geométrica – fase B (sinal da tensão)	3,3220E-03 (2,6000E-04)
#68	Média geométrica – fase C (sinal da tensão)	3,2010E-03 (2,1700E-04)
#69	Distância pico a pico – fase A (sinal da tensão)	2,9380E-03 (4,9300E-04)
#70	Média harmônica – fase B (sinal da corrente)	2,9080E-03 (2,6110E-04)
#71	Média harmônica – fase A (sinal da tensão)	2,8940E-03 (1,7680E-04)
#72	Média harmônica – fase C (sinal da corrente)	2,5090E-03 (2,0680E-04)
#73	Frequência máxima – fase B (sinal da corrente)	2,4270E-03 (1,6050E-04)
#74	Frequência máxima – fase C (sinal da corrente)	2,3480E-03 (1,7490E-04)
#75	Média harmônica – fase A (sinal da corrente)	2,3270E-03 (3,0090E-04)
#76	Frequência máxima – fase A (sinal da corrente)	2,2590E-03 (8,2190E-05)
#77	Média harmônica – fase B (sinal da tensão)	2,2300E-03 (1,1260E-04)
#78	Média harmônica – fase C (sinal da tensão)	1,6430E-03 (1,3410E-04)
#79	Curtose – fase C (sinal da corrente)	1,2050E-03 (7,0240E-05)
#80	Curtose – fase A (sinal da corrente)	1,1780E-03 (4,4770E-05)
#81	Curtose – fase B (sinal da corrente)	1,1730E-03 (5,8360E-05)
#82	Frequência máxima – fase A (sinal da tensão)	8,2670E-04 (1,4690E-04)

Continua na próxima página



Tabela B.1: Continuação da página anterior

<b>Ranking</b>	<b>Atributo</b>	<b>Gini score</b>
#83	Frequência máxima – fase B (sinal da tensão)	7,6430E-04 (7,4350E-05)
#84	Frequência máxima – fase C (sinal da tensão)	7,4670E-04 (3,8560E-05)
#85	Mínimo – fase B (sinal da tensão)	7,2910E-04 (1,8550E-04)
#86	Curtose – fase B (sinal da tensão)	6,7700E-04 (2,5160E-05)
#87	Máximo – fase C (sinal da tensão)	6,1440E-04 (1,6490E-04)
#88	Máximo – fase B (sinal da tensão)	5,8100E-04 (1,4090E-04)
#89	Curtose – fase A (sinal da tensão)	5,7750E-04 (2,4460E-05)
#90	Curtose – fase C (sinal da tensão)	5,6960E-04 (1,6410E-05)
#91	Mínimo – fase A (sinal da tensão)	5,4000E-04 (8,6020E-05)
#92	Mínimo – fase C (sinal da tensão)	5,0790E-04 (1,6280E-04)
#93	Máximo – fase A (sinal da tensão)	4,8650E-04 (1,0410E-04)
#94	Média aritmética – fase A (sinal da corrente)	3,8070E-04 (8,5430E-05)
#95	Média aritmética – fase C (sinal da corrente)	3,5620E-04 (4,6050E-05)
#96	Média aritmética – fase B (sinal da corrente)	3,5130E-04 (2,8200E-05)
#97	Mediana – fase A (sinal da corrente)	2,3990E-04 (3,2350E-05)
#98	Assimetria – fase A (sinal da corrente)	2,3340E-04 (1,8600E-05)
#99	Assimetria – fase C (sinal da corrente)	2,2050E-04 (1,2300E-05)
#100	Mediana – fase C (sinal da corrente)	2,1580E-04 (3,2350E-05)
#101	Assimetria – fase B (sinal da corrente)	2,0890E-04 (1,0200E-05)
#102	Mediana – fase A (sinal da tensão)	2,0650E-04 (9,3340E-06)
#103	Mediana – fase B (sinal da corrente)	1,8840E-04 (2,5170E-05)
#104	Mediana – fase C (sinal da tensão)	1,8110E-04 (8,0190E-06)
#105	Mediana – fase B (sinal da tensão)	1,7570E-04 (4,1440E-06)
#106	Assimetria – fase C (sinal da tensão)	1,7030E-04 (8,4830E-06)
#107	Assimetria – fase A (sinal da tensão)	1,5840E-04 (6,4540E-06)
#108	Assimetria – fase B (sinal da tensão)	1,5630E-04 (9,5310E-06)
#109	AUC – fase C (sinal da tensão)	1,0480E-04 (1,1690E-05)
#110	AUC – fase B (sinal da tensão)	1,0120E-04 (1,8060E-05)
#111	AUC – fase A (sinal da tensão)	9,4550E-05 (9,0830E-06)
#112	Declive – fase B (sinal da tensão)	9,4210E-05 (4,0090E-06)
#113	Média aritmética – fase A (sinal da tensão)	9,3950E-05 (6,7030E-06)
#114	Média aritmética – fase C (sinal da tensão)	8,9050E-05 (6,6990E-06)
#115	Declive – fase A (sinal da tensão)	8,8490E-05 (4,7720E-06)
#116	Média aritmética – fase B (sinal da tensão)	8,8160E-05 (7,2000E-06)
#117	Declive – fase C (sinal da tensão)	8,7950E-05 (5,1530E-06)
#118	Declive – fase C (sinal da corrente)	8,4180E-05 (6,6080E-06)
#119	Declive – fase B (sinal da corrente)	8,1050E-05 (7,0570E-06)
#120	Declive – fase A (sinal da corrente)	7,7760E-05 (7,3300E-06)
#121	Largura de banda espectral – fase C (sinal da tensão)	4,5600E-05 (4,5580E-06)
#122	Largura de banda espectral – fase B (sinal da tensão)	4,3750E-05 (4,3950E-06)
#123	Largura de banda espectral – fase A (sinal da tensão)	3,8740E-05 (3,7420E-06)
#124	Largura de banda espectral – fase C (sinal da corrente)	4,3830E-06 (1,2380E-06)
#125	Largura de banda espectral – fase B (sinal da corrente)	3,4320E-06 (6,3130E-07)
#126	Largura de banda espectral – fase A (sinal da corrente)	3,1900E-06 (6,1300E-07)

## APÊNDICE C – PUBLICAÇÕES

As publicações realizadas no âmbito do projeto são as que seguem:

- Karvat, P. E. M., Almeida, E. C., Ensina, L. A., Oliveira, L. E. S., Santos, S. L. F. e Bernardino, L. S. (2022). Localização de Faltas em linhas de transmissão de energia elétrica utilizando as redes neurais recorrentes LSTM e GRU. Em *III Workshop de Trabalhos de Alunos de Graduação (WTAG)*, páginas 27-33, Búzios, Brasil. DOI: [https://doi.org/10.5753/sbbd\\_estendido.2022.21839](https://doi.org/10.5753/sbbd_estendido.2022.21839)
- Ensina, L. A., Karvat, P. E. M., Almeida, E. C. e Oliveira, L. E. S. (2022). Fault location in transmission lines based on LSTM model. Em *X Symposium on Knowledge Discovery, Mining and Learning*, páginas 1–8, Campinas, Brasil. DOI: <https://doi.org/10.5753/kdmile.2022.227805>
- Ensina, L. A., Oliveira, L. E. S., Almeida, E. C., Santos, S. L. F. e Bernardino, L. S. (2022). Fault classification in transmission lines with generalization competence. Em *IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society*, páginas 1–6, Brussels, Belgium. DOI: <https://doi.org/10.1109/IECON49645.2022.9968993>
- Ensina, L. A., Oliveira, L. E. S., Cruz, R. M. O. e Cavalcanti, G. D. C. (2024). Fault distance estimation for transmission lines with dynamic regressor selection. *Neural Computing and Applications*, 36(4):1741-1759. DOI: <https://doi.org/10.1007/s00521-023-09155-y>
- Tamalu, N., Ensina, L. A., de Almeida, E. C., Pena, E. H. M. e Oliveira, L. E. S. (2023). Fault detection in transmission lines: a denial constraint approach. Em *XXXVIII Simpósio Brasileiro de Bancos de Dados*, páginas 231–243, Belo Horizonte, Brasil. DOI: <https://doi.org/10.5753/sbbd.2023.231718>

Vale destacar ainda que este trabalho colaborou para o desenvolvimento de duas monografias de trabalho de conclusão de curso dos alunos Patrick Eduardo Marchor Karvat e Nicolas Tamalu, formandos do curso de Ciência da Computação da UFPR/Curitiba. Além disso, o método proposto para a classificação está atualmente em operação em um sistema computacional desenvolvido durante o projeto de pesquisa e desenvolvimento (P&D) em parceria entre UFPR, COPEL, Lactec e SIMEPAR (PD-06491-0407/2015). Em especial, o sistema objetiva fornecer informações sobre a análise de falhas aos operadores da COPEL, de modo a auxiliá-los para a tomada de decisão como o religamento de uma linha.

## APÊNDICE D – GLOSSÁRIO

**Abordagem** neste trabalho, utilizamos o termo abordagem para nos referirmos como sinônimo a palavra método.

**Barra** elemento pelo qual se conectam diversos componentes elétricos presentes na subestação. Ela facilita a interconexão dos circuitos, permitindo a transferência eficiente de energia entre diferentes equipamentos, linhas de transmissão e transformadores dentro da subestação (Blume, 2017).

**Barramento** ver “Barra”.

**Critério de seleção** regra utilizada como referência para determinar o desempenho mínimo aceitável para que um preditor do *pool* seja selecionado. O desempenho é mensurado a partir da medida de competência utilizada.

**Classificação de falta** objetiva identificar qual o tipo da falha, ou seja, qual(is) fase(s) está(ão) envolvida(s) na falha, com a incidência ou não da terra.

**Corrente** grandeza elétrica referente ao fluxo de elétrons em um condutor (Blume, 2017).

**Detecção de falta** trata-se da tarefa responsável por determinar o instante de início da falha em uma oscilografia.

**Emulação** refere-se à reprodução do comportamento de um sistema específico em outro sistema, muitas vezes usando hardware ou software para reproduzir as funções do sistema original. Procura reproduzir as características específicas de um sistema em outro visando imitar um sistema em particular.

**Ensemble** algoritmo de aprendizado de máquina composto por um conjunto de preditores, de modo que a decisão final ocorre pela combinação das previsões de cada modelo (Zhou, 2012).

**Escalonamento** ver “Normalização”.

**Falta** fluxo anormal das grandezas elétricas nos componentes do sistema de potência, como o aumento das amplitudes da corrente para uma ou mais fases (Yadav e Dash, 2014).

**Frequência** termo usado para descrever o número de ciclos de onda em um segundo (Blume, 2017).

**Generalização** neste trabalho, trata-se da capacidade de classificar/localizar falhas a partir de dados coletados de linhas de transmissão diferentes daquela usada para treinar o algoritmo.

**Isolação** refere-se à qualidade de isolar (qualitativo), atrelado ao material utilizado para isolar eletricamente, como borracha, plástico, etc.

**Isolamento** refere-se a uma característica quantitativa, como um cabo com isolamento para 750 V.

**Linha de transmissão** são as linhas que operam com as tensões mais elevadas do sistema, tendo com função principal o transporte de energia entre os centros de produção e os centros de consumo (Fuchs, 1977).

**Localização de falta** objetiva determinar o ponto da ocorrência da falha na LT, reportando a distância identificada da falta em relação a um dos terminais da linha.

**Manobra** refere-se a ações como energização, desligamento ou religamento de linhas de transmissão e demais equipamentos elétricos envolvidos no contexto das subestações (Mamede Filho e Mamede, 2020).

**Medida de competência** métrica empregada para avaliar o desempenho (competência) dos modelos do *pool* na região de competência (Moura et al., 2019b).

**Modelos múltiplos preditivos** ver “Ensemble”.

**Normalização** processo que objetiva ajustar as escalas dos atributos de modo que todos variem dentro da mesma faixa de valores (Zheng e Casari, 2018). Porém, o tratamento que realizamos nos sinais de tensão e corrente previamente a extração de atributos, a qual chamamos de normalização ao longo do texto, não torna todo o sinal dentro de uma faixa fixa de valores, vide Figura 6.1.

**Operador** profissional responsável por operar e monitorar as subestações elétricas.

**Oscilografia** representa o registro das oscilações de grandezas elétricas do sistema, como tensões e correntes (Moreto e Rolim, 2010).

**Pool** no contexto da seleção dinâmica, refere-se ao conjunto de preditores original, dos quais apenas um subconjunto será selecionado para a predição do atributo-alvo de um determinado padrão de teste (Cruz et al., 2018a).

**Região de competência** região específica do espaço de características, usualmente estabelecida pelo algoritmo KNN (Cruz et al., 2018a). Em outras palavras, a RC é a região de exemplos no conjunto de validação (ou de treinamento) que cercam o padrão de teste.

**Seleção dinâmica** realiza a seleção de um subconjunto de modelos básicos a partir do conjunto de preditores original (*pool*) individualmente para cada novo padrão de teste a ser predito. Essa estratégia pode ser adotada tanto para a classificação quanto para a regressão (Cruz et al., 2018a).

**Série temporal** refere-se a uma coleção de valores obtidos de medições sequências ao longo do tempo para um determinado fenômeno (Esling e Agon, 2012), por exemplo, sinais de corrente e tensão.

**Simulação** refere-se ao processo de modelar ou reproduzir o comportamento de um sistema real ou processo por meio de um modelo computacional.

**Sinal** sequência ordenada de números que descrevem as tendências e variações de uma grandeza física, como a tensão e a corrente, medida ao longo do tempo (Rao, 2018).

**Sistema elétrico de potência** grande sistema de energia que engloba a geração, a transmissão e a distribuição de energia elétrica (Blume, 2017).

**Subestação** instalação elétrica formada por um conjunto de equipamentos responsáveis pela transmissão e distribuição da energia elétrica, além de equipamentos de proteção e controle (Pérez-Arriaga et al., 2011).

**Tensão** grandeza elétrica referente a diferença de potencial elétrico entre dois pontos em um circuito. É a força que impulsiona os elétrons por meio de um condutor (Blume, 2017).

**Terminal** em sistemas de transmissão de energia elétrica, um terminal, também referenciado como extremidade, é o ponto onde a LT se conecta a uma subestação ou a outro equipamento elétrico (Sato e Freitas, 2015).