



UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
TELECOMUNICAÇÕES

Felipe Hugo Braga Bittar

**Identificação e Classificação de Manobras de
Chaves e Disjuntores usando Algoritmos de
Aprendizado de Máquinas em Subestação de
500kV**

NITERÓI

2022

UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
TELECOMUNICAÇÕES

Felipe Hugo Braga Bittar

Identificação e Classificação de Manobras de Chaves e Disjuntores usando Algoritmos de Aprendizado de Máquinas em Subestação de 500kV

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense como requisito parcial para a obtenção do título de Mestre em Engenharia Elétrica e de Telecomunicações. Área de concentração: Sistemas de Energia Elétrica.

Orientador:

Prof Dr. Vinicius Nunes Henrique Silva

Co-orientador:

Prof Dr. André da Costa Pinho

NITERÓI

2022

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

B624i Bittar, Felipe Hugo Braga
Identificação e Classificação de Manobras de Chaves e Disjuntores usando Algoritmos de Aprendizado de Máquinas em Subestação de 500kV / Felipe Hugo Braga Bittar ; Vinicius Nunes Henrique Silva, orientador ; André Da Costa Pinho, coorientador. Niterói, 2022.
71 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense, Niterói, 2022.

DOI: <http://dx.doi.org/10.22409/PPGEET.2022.m.14993911700>

1. Aprendizado de máquina. 2. Subestações elétricas. 3. Transformada de Wavelet. 4. Processamento digital de sinais. 5. Produção intelectual. I. Nunes Henrique Silva, Vinicius, orientador. II. Da Costa Pinho, André, coorientador. III. Universidade Federal Fluminense. Escola de Engenharia. IV. Título.

CDD -

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

uff Universidade Federal Fluminense
ESCOLA DE ENGENHARIA

COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E DE TELECOMUNICAÇÕES

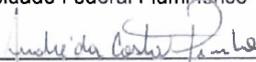
Ata da sessão de julgamento de Dissertação
de Mestrado em Engenharia Elétrica e de
Telecomunicações na área de concentração
Sistemas de Telecomunicações, do aluno
FELIPE HUGO BRAGA BITTAR.

Aos onze dias do mês de março de dois mil e vinte e dois, via videoconferência, reuniu-se a Comissão Examinadora, designada na forma regimental pela Coordenação do Curso, para a defesa da Dissertação de Mestrado apresentada pelo aluno Felipe Hugo Braga Bittar, sob o título “**Identificação e Classificação de Manobras de Chaves e Disjuntores usando Algoritmos de Aprendizado de Máquinas em Subestação de 500kV**”, como requisito para obtenção de grau de Mestre em Engenharia Elétrica e de Telecomunicações – área de concentração Sistemas de Telecomunicações. A Comissão foi presidida pelo professor doutor Vinicius Nunes Henrique Silva, tendo como membros da banca os professores doutores, André da Costa Pinho, Vitor Hugo Ferreira, Alexandre Bessa dos Santos e José Eduardo da Rocha Alves Junior. Aberta a sessão pública, foi concedido ao candidato o tempo de 45 (quarenta e cinco) minutos para a exposição do trabalho, sendo a defesa seguida pela arguição de cada examinador. A seguir, a Comissão reuniu-se para deliberar sobre o texto da dissertação e sua defesa oral, concluindo pela aprovação da mesma sem restrições, e nos termos do Regulamento Geral dos Cursos de Pós-Graduação desta Universidade, foi lavrada a presente ata, lida e julgada, conforme vai assinada pelos membros da Comissão.

Niterói, 11 de março de 2022.



Prof. Dr. Vinicius Nunes Henrique Silva - Orientador
Universidade Federal Fluminense - UFF



Prof. Dr. André da Costa Pinho - Coorientador
Universidade Federal Fluminense - UFF

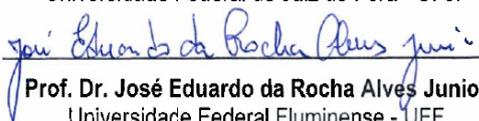


Assinado de forma digital por VITOR HUGO
FERREIRA em 11/03/2022 às 09:35:51 -0300
ID: 2022.03.10.0935:51-0300

Prof. Dr. Vitor Hugo Ferreira
Universidade Federal Fluminense - UFF



Prof. Dr. Alexandre Bessa dos Santos
Universidade Federal de Juiz de Fora - UFJF



Prof. Dr. José Eduardo da Rocha Alves Junior
Universidade Federal Fluminense - UFF

Dedico este trabalho à todos que os educadores que fomentam a ciência. Espero que um dia os ídolos de nossas crianças sejam figuras de professores e educadores como para mim foi o Carl Sagan, Julius Sumner Miller, Neil deGrasse Tyson entre outros

Agradecimentos

Primeiramente gostaria de registrar o meu agradecimento ao Laboratório de Diagnóstico em Equipamentos e Instalações Elétricas(LABDIG) do Cepel-Eletróbrás e a todos os seus membros que colaboraram e me ajudaram na escrita desse trabalho. Sou grato pela oportunidade e pela experiência vivida nesse período.

Agradeço a todos os professores que passaram na minha vida de estudante que me ajudaram a me formar como cidadão e profissional. Para elaboração deste trabalho, em especial, agradeço a participação dos colegas do Laboratório de comunicações ópticas (LACOP-UFF) e do FRIENDS Labs.

"We wish to find the truth, no matter where it lies. But to find the truth we need imagination and skepticism both. We will not be afraid to speculate, but we will be careful to distinguish speculation from fact."

"Desejamos encontrar a verdade, não importa onde ela esteja. Mas, para encontrar a verdade, precisamos de imaginação e ceticismo. Não teremos medo de especular, mas teremos o cuidado de distinguir a especulação do fato"

- Carl Sagan

Resumo

Este trabalho propõe a aplicação de técnicas de processamento digital de sinais (PDS) em conjunto com utilização de algoritmos de aprendizado de máquina para a análise de transitórios eletromagnéticos gerados por manobras de seccionamento no interior de uma subestação que compõe o Sistema Elétrico de Potência (SEP) brasileiro. Além da descrição física da origem do fenômeno transitório, é proposto, através destes eventos, o monitoramento da operação da subestação através da classificação do fenômeno transitório utilizando algoritmos de aprendizado de máquina. Para que o estudo de classificação de eventos transitórios seja factível, foram utilizadas técnicas de tratamento e manipulação de dados cujo o objetivo é extrair características relevantes no domínio do tempo e da frequência das formas de onda dos eventos transitórios amostrados. Desta forma, o trabalho tem sua contribuição no estudo e desenvolvimento de um sistema inteligente capaz de monitorar eventos transitórios gerados através de manobras de seccionamento, gerando relatórios de ocorrência de modo autônomo e criando um histórico de manobras que podem auxiliar em uma investigação e diagnóstico em casos de falhas ocasionadas pelos sinais transitórios.

Palavras-chave: VFTO, Machine Learning, Seleção de atributos, Expansão de banco de dados, Subestação, Transformada Discreta de Wavelet.

Abstract

This work consists of an application of digital signal processing (DSP) techniques in conjunction with machine learning. Regarding the origin of electromagnetic transients generated by disconnect switches and circuit breaker maneuver inside a brazilian substation, this work created a methodology focused to utilize the transient signal propagated from these switching maneuvers in order to monitoring the operation of the substation through maneuver classification by using machine learning methods. For the study of classification of transient events be feasible, some DSP techniques were used in order to extract the relevant characteristics both in time and frequency domain of the sampled transient waveforms. In this way, this work contributes for the study of smartgrids enabling the development of historical switching events by an intelligent system able to sample and classify the transients. The system developed is useful to learn about transients occurrences, investigation of generated transients by maneuvers, assist in an investigation and failure diagnosis caused by transient events.

Keywords: VFTO, Machine Learning, Feature Selection, Data Augmentation, Electrical substation, Wavelet transform.

Lista de Figuras

2.1	Subestação-500kV de interesse.	8
2.2	Sistema de medição simplificado.	9
2.3	Exemplo de manobra de fechamento.	10
3.1	Diagrama de blocos da metodologia.	11
3.2	Exemplo da técnica de <i>Data augmentation</i> por janelamento.	14
3.3	Assinatura espectral dos <i>spikes</i> gerados pela manobra da chave 9005.	14
3.4	Banco de sinais intra-classes de chaves desbalanceado.	15
3.5	Banco de sinais intra-classes de chaves balanceado.	15
3.6	Divisão do espectro após passagem pelo nível 1 da TDW.	18
3.7	TDW visualizada como um banco de filtros.	18
3.8	Amostra VFTO no tempo versus energia dos coeficientes da TDW.	19
3.9	Similaridade entre sinal VFTO e wavelet Daubechies	20
3.10	Fluxo do método <i>filter</i>	24
3.11	Fluxo do método <i>Wrapper</i>	24
3.12	Topologia de uma RNA.	26
3.13	Entidades do algoritmo MVS.	28
3.14	Árvores de decisão compondo um <i>Random Forest</i>	29
3.15	Cálculo do ganho de informação após divisão dos nós.	30
4.1	Diagrama em árvore do fluxo de classificações.	32
4.2	Matriz de confusão.	34
4.3	Distribuição Chave vs Disjuntores.	35
4.4	Distribuição da classe Chaves Seccionadores.	38

4.5	Custo computacional dos algoritmos com todas as 23 features <i>vs</i> 18 features.	40
4.6	Distribuição classe Aberto <i>versus</i> Fechado das chaves	40
4.7	Distribuição classe dos disjuntores.	42
4.8	Distribuição classe aberto <i>versus</i> fechado dos disjuntores.	44

Lista de Tabelas

3.1	Distribuição original do banco de dados	13
3.2	Distribuição do banco de dados após <i>data augmentation</i>	16
3.3	Entradas e seus acrônimos	21
3.4	Bandas dos coeficientes de wavelets	22
4.1	Resultados Tarefa 1 - Acurácia clássica.	34
4.2	Resultados Tarefa 1 - métrica F1.	36
4.3	Resultados Tarefa 1 - <i>Undersample</i>	37
4.4	Resultados Tarefa 1.1.1	39
4.5	Resultados Tarefa 1.1.2	41
4.6	Resultados Tarefa 1.2.1	43
4.7	Resultados Tarefa 1.2.2	45
4.8	desempenho com novos subconjuntos pelo método <i>wrapper</i>	47
4.9	Combinações com melhor desempenho em acurácia.	49
4.10	Comparação qualitativa de trabalhos da área.	50

Lista de Abreviaturas e Siglas

VFTO	: Very fast transient over voltage;
SEP	: Setor elétrico de potência ;
TE	: Transitórios eletromagnéticos;
LT	: Linha de transmissão;
RDP	: Registradores Digitais de Pertubação;
SIN	: Sistema interligado nacional;
SM	: Sistema de medição;
SE	: Subestação de energia;
ML	: Machine Learning;
FS	: Feature selection;
DA	: Data augmentation;
A_{cc}	: Data augmentation;
F1	: Média harmônica (<i>F1-score</i>);
TW	: Transformada de Wavelet;
Em1	: Energia média do coeficiente de detalhe nível 1;
Em2	: Energia média do coeficiente de detalhe nível 2;
Em3	: Energia média do coeficiente de detalhe nível 3;
Em4	: Energia média do coeficiente de detalhe nível 4;
Em5	: Energia média do coeficiente de detalhe nível 5;
Em6	: Energia média do coeficiente de detalhe nível 6;
Em7	: Energia média do coeficiente de detalhe nível 7;
Em8	: Energia média do coeficiente de detalhe nível 8;
Em9	: Energia média do coeficiente de detalhe nível 9;
Em10	: Energia média do coeficiente de detalhe nível 10;
EmA10	: Energia média do coeficiente de aproximação nível 10;

WEE1	:	Entropia do coeficiente de detalhe nível 1;
WEE2	:	Entropia do coeficiente de detalhe nível 2;
WEE3	:	Entropia do coeficiente de detalhe nível 3;
WEE4	:	Entropia do coeficiente de detalhe nível 4;
WEE5	:	Entropia do coeficiente de detalhe nível 5;
WEE6	:	Entropia do coeficiente de detalhe nível 6;
WEE7	:	Entropia do coeficiente de detalhe nível 7;
WEE8	:	Entropia do coeficiente de detalhe nível 8;
WEE9	:	Entropia do coeficiente de detalhe nível 9;
WEE10	:	Entropia do coeficiente de detalhe nível 10;
WEEA10	:	Entropia do coeficiente de aproximação nível 10;
RMS	:	Valor eficaz;

Sumário

1	Introdução	1
1.1	Motivação	3
1.2	Organização do trabalho	4
2	Geração e aquisição de sinais de transitórios	5
2.1	Elementos de uma subestação	5
2.2	Distúrbio transitório gerado por manobras	6
2.3	Topologia da subestação de interesse	7
2.4	Sistema de medição	9
3	Metodologia e técnicas de processamento digital de sinais para Machine Learning	11
3.1	Aumento de dados (<i>Data augmentation</i>)	12
3.2	Extração de características (<i>Feature Extraction</i>)	16
3.3	Seleção de atributos (<i>Feature Selection</i>)	22
3.3.1	Métodos de seleção de atributos utilizados	23
3.4	Aprendizado de máquina (<i>Machine learning</i>)	25
3.4.1	Redes neurais artificiais	25
3.4.2	Máquina de vetor suporte	27
3.4.3	<i>Random Forest</i>	28
4	Resultados	31
4.1	Análise de desempenho	31

4.1.1	Chave <i>versus</i> Disjuntor (Tarefa 1)	33
4.1.2	Identificação da Chave (Tarefa 1.1.1)	38
4.1.3	Identificação da manobra da chave (Tarefa 1.1.2)	40
4.1.4	Identificação do Disjuntor (Tarefa 1.2.1)	42
4.1.5	Identificação da manobra dos disjuntores (Tarefa 1.2.2)	44
4.2	Otimização de desempenho pelo método <i>wrapper</i>	46
4.3	Resumo de resultados	48
4.4	Trabalhos correlatos	50
5	Conclusão	51
	Referências	53

Capítulo 1

Introdução

Transitórios Eletromagnéticos (TE) são respostas elétricas oriundas de alterações súbitas das condições operacionais de um sistema de energia elétrica. No contexto de Sistemas Elétricos de Potência (SEP) os TE podem ser originados devido a uma descarga atmosférica ou por manobras de seccionamento [1]. Esses eventos são indesejáveis tanto para quem fornece energia (subestações geradoras ou distribuidoras) quanto para o consumidor final [1]-[2]. Entretanto, os transitórios eletromagnéticos são intrínsecos nas operações elétricas do SEP e eventualmente podem ocasionar um mal-funcionamento e até a total falha de um equipamento elétrico [1]

Os TE podem ser classificados como: impulsivos, causados por descargas atmosféricas ou de altas frequências, causados pela comutação do estado de chaves seccionadoras e/ou disjuntores [1]-[3]-[4]. Os transitórios de altas frequências tem sua geração muito rápida e de curta duração, em outras palavras, do início da perturbação até sua máxima amplitude o evento tem duração na ordem de nano segundos [4]. Esses transitórios são conhecidos como *very fast transient overvoltage* (VFTO) [5]. Durante um VFTO os equipamentos elétricos são submetidos a esforços devido a variações de corrente podendo levar a um mal-funcionamento ou a uma degradação prematura dos equipamentos, e não obstante a sua falha total [2]. Embora haja estudos de simulações de esforços causados por manobras, feito por softwares durante a fase de projeto, o evento do transitório mesmo que não ocasione um falha catastrófica, reduz a qualidade de operação e a vida útil dos ativos da rede [6]. Assim, para que seja possível entregar uma energia de qualidade, os ativos das subestações precisam passar por eventuais manutenções devido ao conseqüente desgaste causado pelos VFTOs. Logo, o monitoramento e a classificação destes eventos se tornam interessantes, pois estes dão amparo em diagnósticos de falhas e investigação da causa logo após eventos VFTOs [2].

Nesse contexto, com o intuito de mensurar VFTO incidentes em um ativo da rede, Guo-ming et al. [7] criaram um sensor capaz de adquirir sinais oscilográficos de um VFTO dentro de uma subestação isolada a gás pela blindagem da bucha de um transformador. O sensor relatado é constituído de um divisor capacitivo em série com um divisor resistivo para atenuação e leitura do sinal pelo osciloscópio.

S.A. Probert et al. [8], exploraram VFTOs causados por manobras de chaves, falhas na linha de transmissão (LT) e descargas atmosféricas através da energia dos coeficientes da transformada de wavelet utilizando a wavelet Daubechies. A geração dos sinais oscilográficos foram gerados via software de simulação PSCAD, e então, mostrou-se a possibilidade de classificar transitórios, uma vez que cada tipo de evento VFTO possuía distribuição de energia espectral distinta.

Apresentando interesse na classificação de VFTOs, P.L. Mao et al. [9] utilizaram redes neurais artificiais (RNA) para discriminar transitórios causados por faltas internas e correntes de magnetização em um transformador de potência. Os sinais utilizados foram gerados por software e após a classificação foi sugerido à atuação de sistemas de proteção específicos para cada evento.

Utilizando também as RNAs, Zhengyou He et al. [10], propuseram a classificação de seis categorias de distúrbios transitórios: manobra de disjuntor, energização de banco de capacitores, falha de curto-circuito em LT, arco primário, perturbação por descarga atmosférica adjacente a uma LT e falha por descarga atmosférica diretamente incidente na LT.

De maneira mais ampla e com implementação no interior de subestações de energia, B.Souza et al. [11] desenvolveram um sistema de monitoramento com classificação de distúrbios através da aquisição de sinais por Registradores Digitais de Perturbação (RDP) comerciais, onde os sinais oscilográficos gerados no interior de subestações da Companhia Hidrelétrica do São Francisco (CHESF) são classificados pelos tipos de faltas e/ou distúrbio analisando-se a energia dos coeficientes de wavelet da família Daubechies. Desta maneira, em um eventual desligamento por falta e/ou falha, a retomada da operação da subestação poderia ser realizada com mais rapidez e segurança, uma vez que se dispôs do máximo de informação sobre a ocorrência dos distúrbios, reduzindo-se multas e prejuízos pelos desligamentos.

Sob a perspectiva de monitorar os transitórios que incidem nos ativos mais valiosos da rede, esse trabalho propõe um sistema de monitoramento autônomo que seja capaz de amostrar VFTO nocivos aos equipamentos e sua classificação por um algoritmo de

aprendizado de máquina, conhecido também por *machine learning*(ML). Tal proposta baseia-se em uma abordagem economicamente viável de um sistema de monitoramento para pequenos empreendimentos, como as subestações locais e regionais que dificilmente são instrumentadas com RDP de qualidade, limitando a aquisição de VFTOs por possuírem baixa taxa de amostragem, uma vez que os tais eventos estão na faixa de 500kHz até dezenas de MHz [4]-[12]. Para a aquisição de VFTOs e criação de um banco de sinais, esse trabalho utilizou medidas feitas por um sistema de medição (SM) do tipo *high frequency current transformer* (HFCT) da marca Pearson, modelo 3525, com largura de banda de 5 - 15MHz, tempo de subida igual a 25ns e corrente máxima de pico de 5kA [13]. O SM foi elaborado e relatado por Vasques M. Rodrigues et al. [13], que o instalaram na cordoalha do terra de um transformador de corrente, coletando transitórios no trecho de interesse da subestação de estudo. Deste modo, o estudo do sistema de monitoramento de VFTOs foi realizado através do banco de sinais coletado pelo SM relatado. Por fim, a implementação deste sistema pode auxiliar a equipe de manutenção a analisar falhas ocasionadas após manobras de seccionamento de chaves e disjuntores.

1.1 Motivação

No caso de uma falha que ocasione a interrupção no fornecimento de energia, a empresa concessionária de energia fica sob o risco de multas pelo órgão fiscalizador, logo, uma investigação e diagnóstico da origem da falha para um religamento mais ágil mostra-se pertinente para redução de prejuízos financeiros. Nessa lógica, o registro histórico de eventos e das formas de onda de sinais transitórios podem servir de apoio para dar maior agilidade em um diagnóstico no caso de falhas [11]-[14]-[15]. Esse trabalho tem como objetivo criar uma metodologia de classificação de transitórios VFTOs ocasionados por chaves seccionadoras e disjuntores.

O registro de sinais transitórios e sua classificação pelas diversas manobras realizadas no interior de uma subestação visam auxiliar em um diagnóstico de falha, pois após estes serem classificados e discriminados, a equipe de manutenção dispõe de um histórico de ocorrências a respeito da rede elétrica em um momento de pré-falha, acelerando a investigação e a reconstituição de eventos para o total entendimento da falha em questão.

O desafio deste trabalho, cuja proposta é de uma metodologia de classificação automática de VFTOs, é a transformação de dados disponíveis da forma de onda do sinal transitório, em dados de entrada úteis para os algoritmos de aprendizado de máquina.

Assim, após a criação da metodologia, esta pode ser utilizada para o monitoramento de sinais transitórios do tipo VFTOs, como por exemplo, em projetos personalizados no interior de uma subestação de energia.

1.2 Organização do trabalho

Este trabalho foi organizado em três blocos: No capítulo 2 foram descritos os principais equipamentos instalados no interior de uma subestação de energia, suas respectivas funcionalidades, origem dos transitórios VFTOs devido a manobras de chaves seccionadoras/disjuntores, e descrição do sistema de medição usado para a coleta de sinais. O capítulo 3 descreve sobre a metodologia adotada até a extração dos resultados. Em seguida, o capítulo 4 expõe os resultados obtidos e as respectivas análises. Por fim, no capítulo 5 foram apontadas as conclusões sobre o estudo de viabilidade do sistema autônomo de monitoramento com algoritmos de ML embutido e trabalhos futuros.

Capítulo 2

Geração e aquisição de sinais de transitórios

Nesse capítulo são apresentados os principais equipamentos que compõem uma subestação de energia, em especial as chaves seccionadoras e os disjuntores que são os elementos de estudo desse trabalho. As chaves e os disjuntores são elementos que permitem fazer ligações ou isolamentos entre trechos da rede, porém, uma vez manobrados estes geram transitórios que podem gerar danos aos ativos da rede. Por se tratar de eventos que eventualmente trazem degradação e falha de equipamentos ligados a rede, esses transitórios necessitam de acompanhamento a respeito de seus níveis de sobre-tensão gerados e, além disso, as próprias formas de ondas podem ser utilizadas como informação para monitoramento e controle da rede de modo autônomo. Nesse sentido, o capítulo ainda traz a explanação da topologia da subestação utilizada como teste para coleta das formas de onda e maiores detalhes SM.

2.1 Elementos de uma subestação

Uma subestação pode ser descrita como um conjunto de equipamentos interligados entre si com o objetivo de controlar e direcionar o fluxo da energia, funcionando como ponto de entrega e distribuição entre diferentes pontos de um sistema, modificando os níveis de tensão para os respectivos usuários. As subestações, de maneira geral, são pontos de convergência, entrada e saída, de linhas de transmissão ou distribuição [16]-[17].

Alguns dos equipamentos mais comuns encontrados em uma subestação são: barramentos, linhas aéreas, transformadores de corrente (TC) e potencial (TP), relés, fusíveis, para-raios e malha de terra, transformadores de potência e, por fim, disjuntores e chaves

seccionadoras. Em particular, os dois últimos equipamentos citados são destinados a realizar manobras de seccionamento isolando um trecho do circuito elétrico ou modificando o fluxo de energia elétrica [16]-[17].

As chaves seccionadoras são elementos que interligam trechos da rede modificando de forma rápida a topologia de uma subestação. Em operação normal, essas devem manter a condução de sua corrente nominal, inclusive a corrente de curto-circuito até que um elemento de proteção atue [18]. Apesar de as chaves e os disjuntores fazerem operações de conexão ou desconexão entre pontos da rede, os disjuntores são equipamentos diferentes das chaves, pois os disjuntores são equipamentos de manobra destinados a isolar um trecho da rede de modo a interromper correntes de curto-circuito durante uma falha ou defeitos [16]-[18].

As manobras de seccionamento de chaves e disjuntores são eventos comuns durante a operação de uma subestação, no entanto, estas geram transitórios de sobretensões de altas frequências que podem danificar e degradar prematuramente equipamentos [19]-[20]. Os transitórios carregam informação a respeito da manobra realizada que lhe deu origem. Logo, um sistema de aquisição e monitoramento de sinais transitórios pode dar suporte a operação da subestação a respeito de sua incidência em ativos e geração de relatórios em casos de falhas e interrupções no fornecimento de energia.

2.2 Distúrbio transitório gerado por manobras

A origem do distúrbio transitório causado por manobra de seccionamento se dá na geração de arcos elétricos, conhecidos também como *strikes* ou *surge*, que ocorrem durante o movimento dos contatos metálicos das chaves e dos disjuntores [4]-[20]. Em uma única manobra podem ocorrer centenas de arcos elétricos. O número exato dependerá do projeto, características das chaves/disjuntores e do tempo total de manobra [18]. Em teoria, durante esse processo, o primeiro arco voltaico gerado é o de maior amplitude para uma manobra de fechamento e o menor para uma manobra de abertura [20].

Em particular, os arcos elétricos ocorrem quando a diferença de potencial elétrico entre os contatos da chave superam a tensão contrária gerada pelo dielétrico do meio que está entre os contatos da chave [20]. A ocorrência do arco elétrico é um fenômeno de natureza aleatória. Desta forma, à medida que os contatos da chave se aproximam em uma manobra de fechamento, a chance de ocorrer um arco aumenta, porém, nada garante que o mesmo possa ocorrer sempre no mesmo ponto de manobra, isto é, no mesmo distanciamento entre

os contatos[20]-[21].

Uma vez registrado o evento VFTO, que é inerente a operação de manobras por chaves seccionadoras e disjuntores no interior da SE, pode-se ter mais conhecimento a respeito da natureza espectral do evento, operação e registro de manobras, além de auxílio em diagnósticos de falhas [20]-[22]-[11].

2.3 Topologia da subestação de interesse

O banco de sinais VFTO utilizado para realização deste trabalho foi adquirido no interior de uma subestação de 500kV que compõe o Sistema Interligado Nacional (SIN). No total foram obtidos sinais de 8 chaves seccionadoras e 2 disjuntores. O diagrama unifilar da subestação de interesse é ilustrado na Figura 2.1 e as chaves e disjuntores manobrados são identificados com círculos azuis. Os sinais VFTOs gerados por essas manobras foram coletados utilizando o SM [13] ligado a um transformador de corrente.

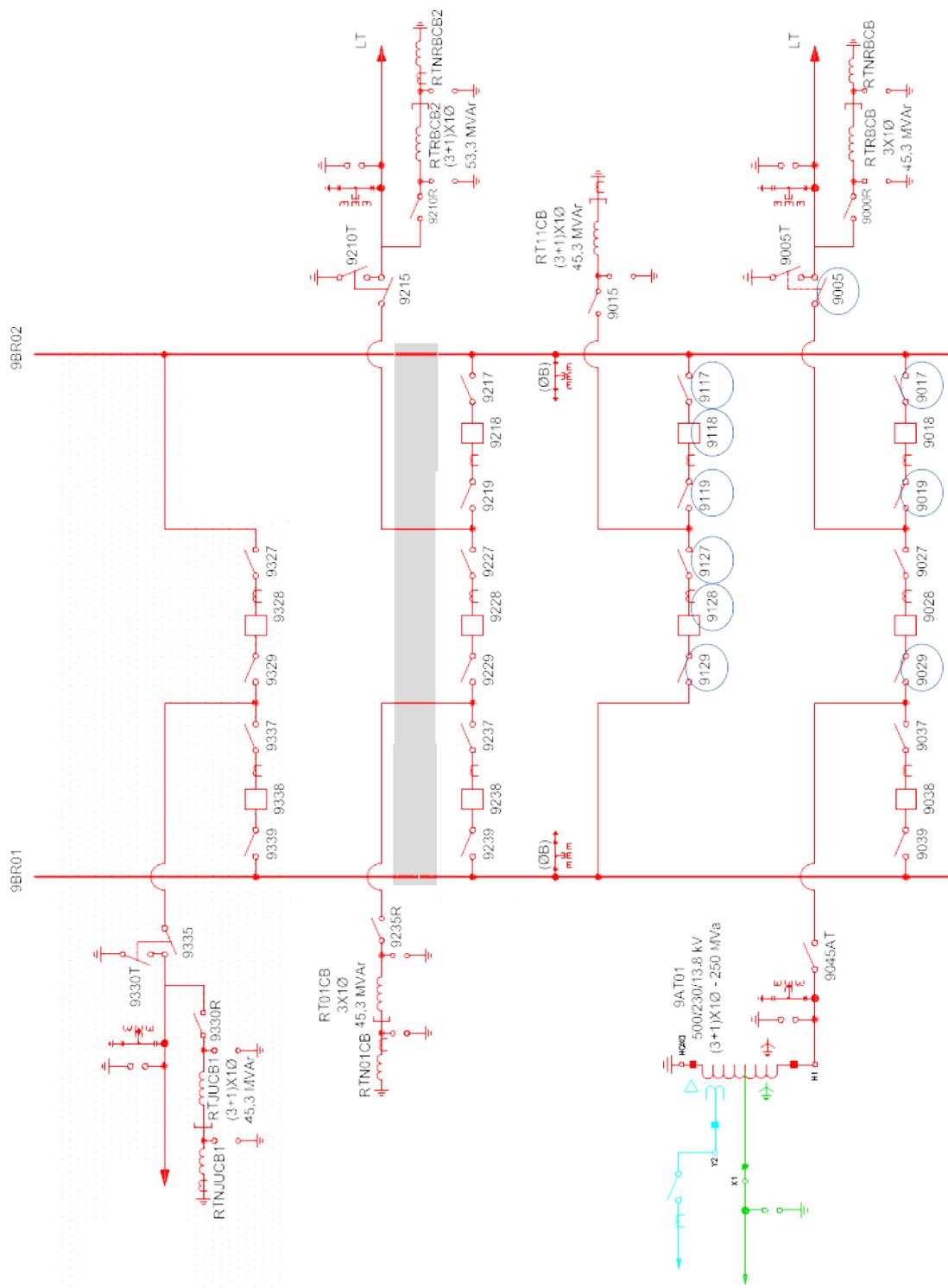


Figura 2.1: Subestação-500kV de interesse.

A base de dados é originalmente é composta por 143 amostras distintas e apresentam uma distribuição desbalanceada, pois há 124 amostras de chaves seccionadoras e apenas 19 de disjuntores.

2.4 Sistema de medição

A aquisição dos transitórios de sobrecorrentes foram realizadas fazendo-se uso de um transformador de corrente (TC) instalado no interior da SE e de um sensor transformador de corrente para altas frequências (High Frequency Current Transformer – HFCT modelo 3525) da marca Pearson, com tempo de subida de 25ns, corrente máxima de pico igual a 5 kA, largura de banda de 5 Hz – 15 MHz e com um atenuador em cascata. No diagrama em blocos do Sistema de medição (SM), na Figura 2.2, é possível ver, em vermelho, o HFCT conectado à cordoalha de aterramento do TC.

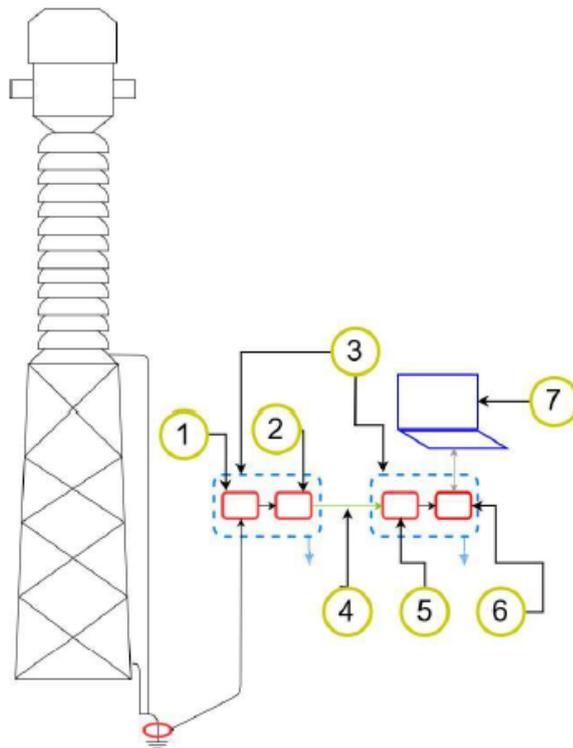


Figura 2.2: Sistema de medição simplificado.

Nesse sistema, o atenuador (1) é responsável por diminuir o nível de potência de um VFTO e adequar o sinal para o sistema sensor. Já o sistema ótico, composto por uma fibra ótica (4), transmissor (2) e receptor (5), é utilizado para transmitir os sinais provenientes do atenuador para o sistema de aquisição de dados com o mínimo de interferência eletromagnéticas que venham prejudicar as medidas. No sistema de aquisição, o osciloscópio (6) digitaliza os formatos de onda dos sinais VFTO que posteriormente são tratados pelo computador (7). Todo o sistema é devidamente blindado para evitar interferências eletromagnéticas externas provenientes dos outros equipamentos, dispositivos e sistemas que compõem uma subestação.

A Figura 2.3 mostra um exemplo de VFTO provocado pelo fechamento de uma chave seccionadora medida por este SM.

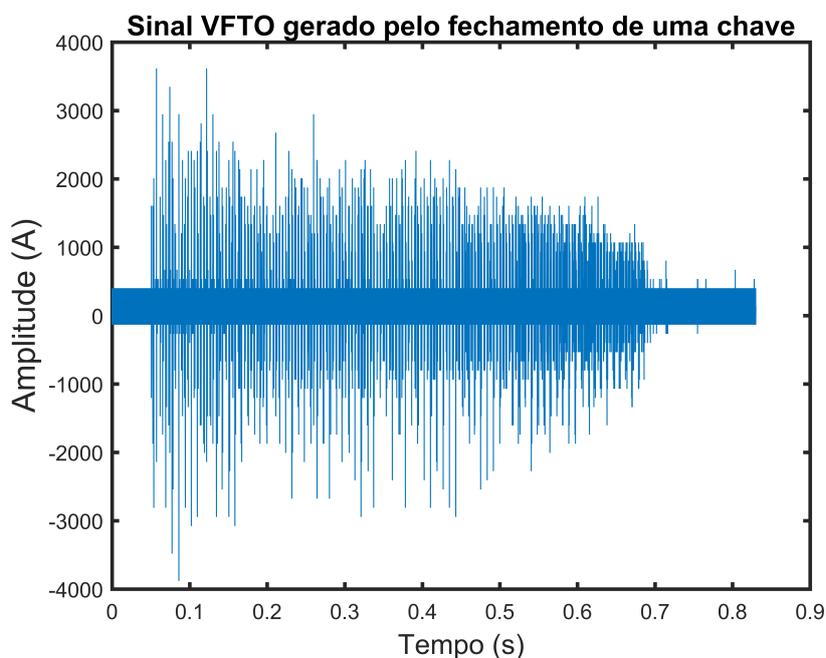


Figura 2.3: Exemplo de manobra de fechamento.

Este e outros sinais oscilográficos de chaves seccionadoras, assim como de disjuntores, serão classificados e identificados através de uma metodologia computacional que é apresentada no Capítulo 3.

Capítulo 3

Metodologia e técnicas de processamento digital de sinais para Machine Learning

A utilização de ferramentas de ML exige intensa atividade de PDS de modo a preparar as entradas para os algoritmos [23]. A metodologia adotada se baseia em uma expansão da base de dados(Data Augmentation) dos sinais VFTOs seguida de um PDS para extração de atributos(Feature extraction) que representam os sinais das chaves seccionadoras e dos disjuntores, além de otimizações de desempenho e redução do conjunto de atributos utilizados nos algoritmos de ML, conforme mostra a Figura do diagrama de blocos 3.1.

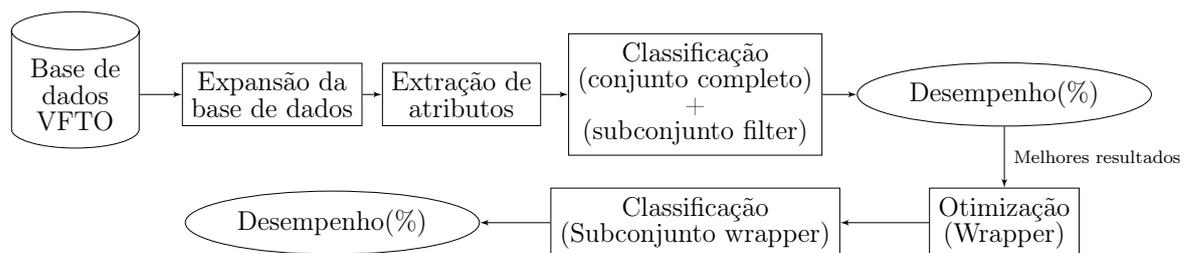


Figura 3.1: Diagrama de blocos da metodologia.

Após a primeira avaliação de desempenho da classificação pelos conjuntos completos, composto de todas as entradas, e do subconjunto do método de seleção de atributos do tipo filtragem, composto com das 80% entradas mais significativas, os melhores resultados foram encaminhados para uma otimização de desempenho e redução do número de entradas, e, assim, avaliadas de maneira definitiva chegando aos melhores desempenhos alcançados pela metodologia descrita.

3.1 Aumento de dados (*Data augmentation*)

O banco de sinais adquirido originalmente, cuja distribuição é exibida na Tabela 3.1, apresenta um desbalanceamento entre as classes dos equipamentos de chaves e disjuntores. A classe de chaves representa cerca de 87% das amostras e dos disjuntores 13%. Observa-se ainda que há também uma discrepância interna na classe das chaves, onde a chave 9019 é composta de 29 amostras e a 9005 é representada por apenas 2. Esse desbalanceamento tanto inter-classes quanto intra-classes pode prejudicar a fase de treinamento dos algoritmos de ML, pois a generalização dos dados observados pode ser influenciada a classificar com uma performance melhor apenas as classes com mais observações [24]-[25].

Visando obter um conjunto de dados balanceado, foram utilizadas manipulações nas amostras de sinais VFTO de modo a extrair mais amostras advindas das originais, e, portanto, expandindo-se o banco de sinais. A essa técnica dá-se o nome de *Data augmentation*, também conhecida por *oversampling* e todo processamento dos dados foram feitos no ambiente Matlab [24]-[25].

A etapa de *data augmentation* é executada de forma a equilibrar o número de amostras de classes distintas e presentes no banco de dados. Ao fim do processo, o banco de sinais terá mais amostras observáveis e uma distribuição mais uniforme entre as classes, assim, são apresentadas aproximadamente a mesma quantidade de amostras de cada classe ao algoritmo de aprendizado de máquina durante o processo de treinamento [23]-[25]. O principal objetivo do aumento de dados é o balanceamento da base de dados, além de ampliar a quantidade de amostras disponíveis no processo de treinamento e, conseqüentemente, a utilização de métricas de acurácia menos complexas para avaliação de desempenho dos algoritmos.

Total de amostras	Equipamento	ID	Total por ID	Manobra	Amostras por manobra
143	Chave	9017	24	Abrindo	12
				Fechando	12
		9019	30	Abrindo	14
				Fechando	16
		9029	9	Abrindo	3
				Fechando	6
		9117	25	Abrindo	13
				Fechando	12
		9119	13	Abrindo	5
				Fechando	8
		9127	17	Abrindo	8
				Fechando	9
	9129	4	Abrindo	1	
			Fechando	3	
	9005	2	Abrindo	2	
			Fechando	0	
Disjuntor	9118	8	Abrindo	2	
			Fechando	7	
	9128	11	Abrindo	7	
			Fechando	4	

Tabela 3.1: Distribuição original do banco de dados

Para dados oscilográficos, sinais no tempo ou séries temporais as técnicas de *data augmentation* podem ser aplicadas diretamente na forma de onda [25]. A ideia central destas técnicas é a síntese de amostras semelhantes às originais. Por se utilizarem amostras originais, as novas amostras devem necessariamente ser rotuladas com a mesma classe que a amostra original. Portanto, nesse trabalho, para cada sinal transitório no domínio do tempo, foi feito um janelamento de $200 \mu\text{s}$ centrado em cada um dos picos dos *strikes* de sobrecorrente e depois normalizados em amplitude, conforme ilustra a Figura 3.2. A escolha de duração do janelamento do sinal foi feita de modo a se obter resolução espectral mínima de 5kHz [26]. Portanto, a faixa espectral de interesse está contida nas médias e altas frequências, cujas bandas abrangem, respectivamente 5k-500kHz e 500k-5MHz [4].

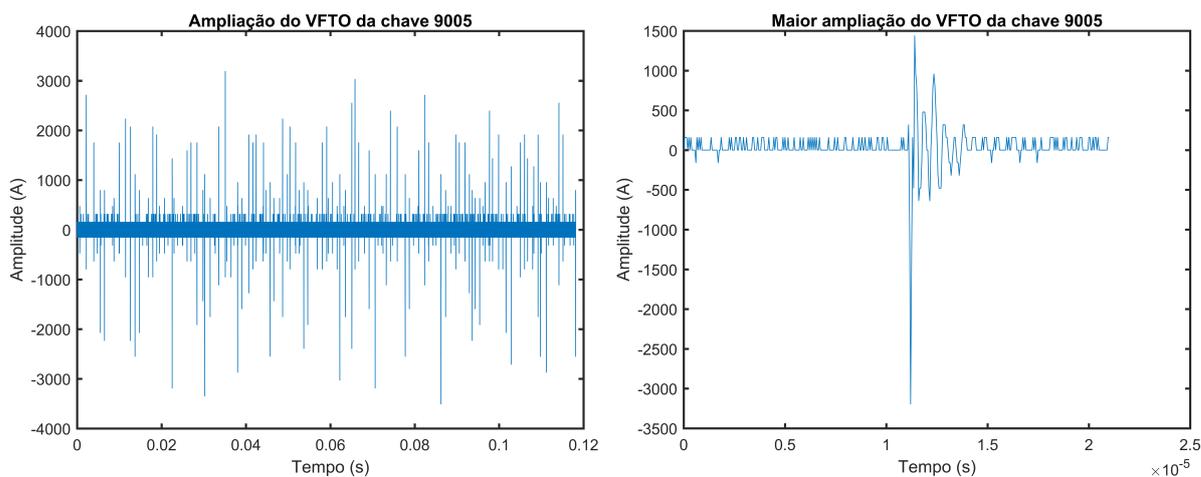


Figura 3.2: Exemplo da técnica de *Data augmentation* por janelamento.

A Figura 3.3 da chave 9005 compara a assinatura espectral de duas amostras janeladas distintas retiradas de uma manobra de abertura da chave 9005. Analisando a Figura em questão, é possível observar um relevante grau de similaridade na assinatura espectral das novas amostras, o que confere confiabilidade ao processo de expansão da base de dados.

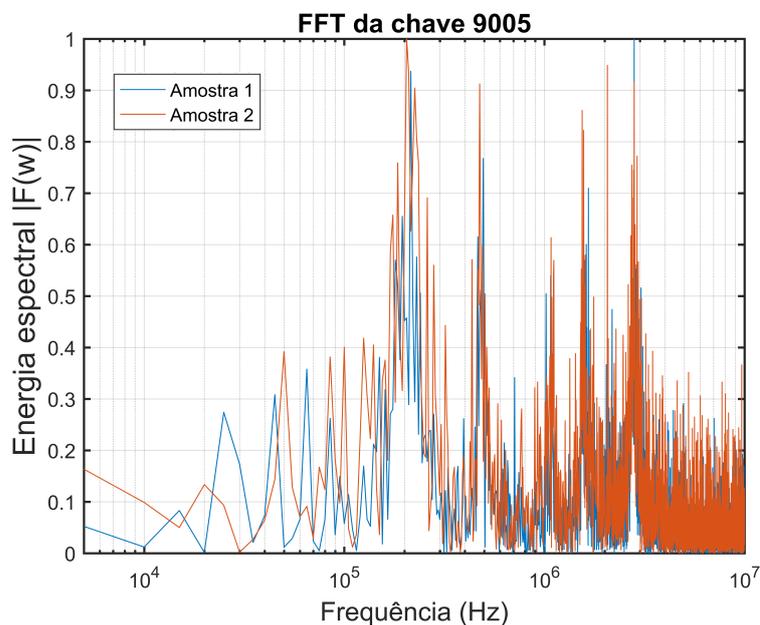


Figura 3.3: Assinatura espectral dos *spikes* gerados pela manobra da chave 9005.

A distribuição do banco de sinais original e após aplicada a técnica de DA são ilustradas nas Figuras 3.4 e 3.5 abaixo:

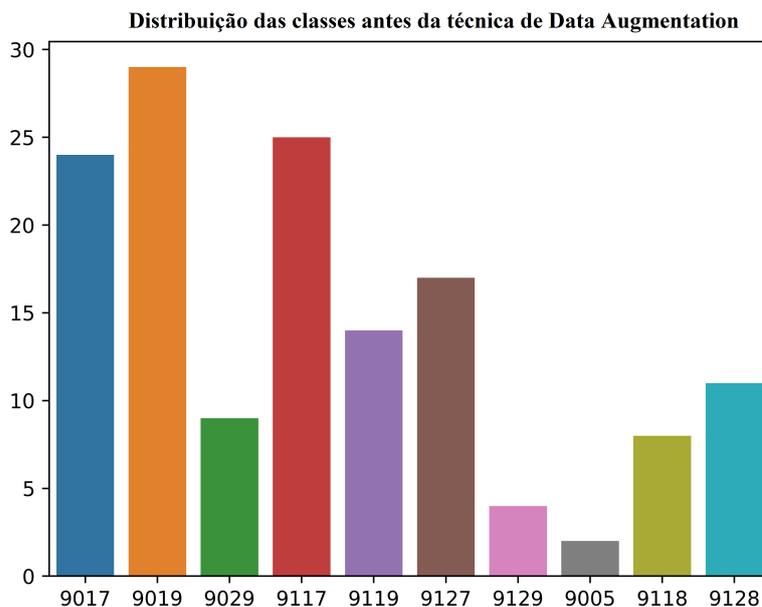


Figura 3.4: Banco de sinais intra-classes de chaves desbalanceado.

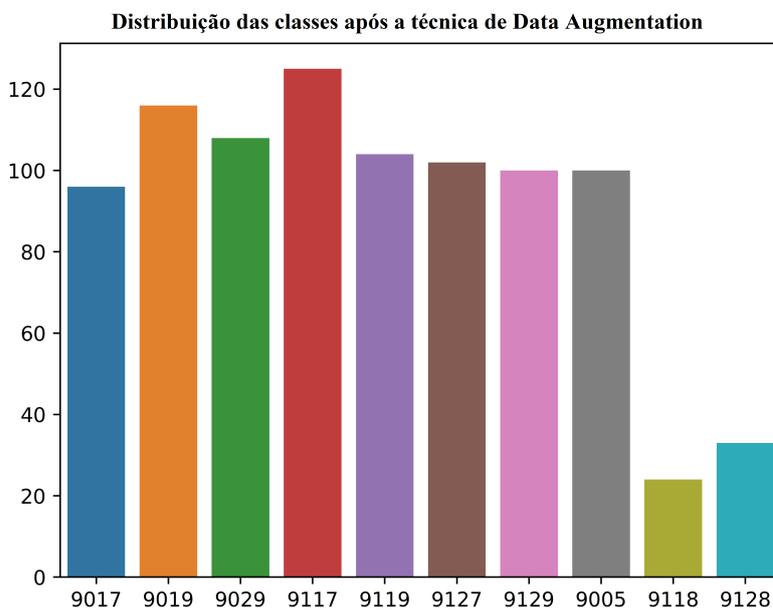


Figura 3.5: Banco de sinais intra-classes de chaves balanceado.

Percebe-se que houve um balanceamento intra-classes das chaves que agora correspondem a 94% do total, distribuídas entre: 9017(10.52%), 9019(13.15%), 9029(11.84%), 9117(13.7%), 9119(11.4%), 9127(11.18%), 9129(10.96%) e 9005(10.96%). Entretanto, o aumento do número de amostras da classe dos disjuntores não foi grande suficiente para um balanceamento inter-classes, representando 6% do total distribuídas entre: 9118(2.63%)

e 9128(3.62%). A limitação do aumento da classe dos disjuntores é devido ao menor número de arcos elétricos gerados durante a manobra desses equipamentos. Apesar do desbalanceamento inter-classes, houve o aumento do número de amostras dos disjuntores e mantendo-se um balanceamento intra-classe. A Tabela 3.2 mostra a nova distribuição do banco de sinais.

Total de amostras após <i>data augmentation</i>	Equipamento	ID	Total por ID	Manobra	Amostras por manobra
912	Chave	9017	96	Abrindo	48
				Fechando	48
		9019	120	Abrindo	52
				Fechando	64
		9029	108	Abrindo	36
				Fechando	72
		9117	125	Abrindo	65
				Fechando	60
		9119	104	Abrindo	44
				Fechando	60
		9127	102	Abrindo	48
				Fechando	54
		9129	100	Abrindo	25
				Fechando	75
	9005	100	Abrindo	100	
			Fechando	0	
Disjuntor	9118	24	Abrindo	6	
			Fechando	21	
	9128	33	Abrindo	21	
			Fechando	12	

Tabela 3.2: Distribuição do banco de dados após *data augmentation*.

Assim, a técnica de DA por janelamento dos sinais VFTOs se mostrou efetiva tanto para o aumento do número de amostras, mantendo-se a assinatura espectral, quanto para o balanceamento da base de dados. Na seção seguinte, explora-se como são extraídas informações espectrais de cada sinal transitório e a utilidade dessas informações como entradas nos algoritmos de aprendizado de máquina.

3.2 Extração de características (*Feature Extraction*)

Uma característica ou *feature* é um dado numérico que representa uma qualidade/característica individual de um sinal transitório. A extração de características de um sinal

oscilográfico, ou *Feature Extraction*, é a etapa onde há a transformação da forma de onda no tempo em um vetor de atributos numéricos que representará as características do sinal [27]. Os dados representativos ou vetores representativos devem ser organizados em uma tabela, ou qualquer outra estrutura capaz de ser acessada pelo algoritmo de ML [27].

As características de uma oscilografia podem ser extraídas tanto no domínio do tempo quanto na frequência. A característica extraída no domínio do tempo pode ser, por exemplo, o valor RMS. Já no domínio da frequência, as transformadas de Fourier e Wavelet podem ser utilizadas [28].

A Transformada de Wavelet (TW) é uma ferramenta matemática para análise de decomposição de frequência de um sinal. A ferramenta opera escalonando e transladando uma função de base chamada de wavelet mãe (*mother wavelet*), como mostra a Equação 3.1.

$$TW(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi \frac{(t - b)}{a} dt \quad (3.1)$$

onde, a é o fator de escalonamento, b é fator de deslocamento, Ψ representa a função Wavelet da família Daubechies (db) e $x(t)$ o sinal transitório a ser analisado.

A transformada decompõe a banda espectral do sinal em várias sub-bandas espectrais de modo a varrer o espectro original por completo [28]. Nas aplicações computacionais é usada a Transformada de Wavelet Discreta (TDW), o que torna os cálculos efetivamente menos onerosos [29]. De maneira prática, a TDW pode ser visualizada como um banco de filtros cuja função de transferência é dependente da *mother wavelet* escolhida. O filtro divide o espectro do sinal pela metade com uma faixa ligeiramente sobreposta [29]-[30], conforme ilustra a Figura 3.6. Por ser uma ferramenta de decomposição espectral, a TDW respeita o teorema da frequência Nyquist, ou seja, um sinal discreto com frequência de amostragem (F_s) tem faixa espectral limitada a $F_s/2$ Hz.

O banco de filtros da TDW é composto por um filtro passa-alta(HP) em paralelo com um filtro passa-baixa(LP) a fim de calcular os coeficientes de Detalhe(D_n) e Aproximação(A_n), respectivamente. De modo a obter-se uma maior resolução em frequência, os coeficientes de aproximação são filtrados sequencialmente, ou seja, para uma nova decomposição de nível superior (nível $n = 2, 3, 4...$). Esta estrutura é organizada em formato de árvore para uma decomposição multi-nível [29]. A Figura 3.7 mostra a decomposição por um banco de filtros para a obtenção dos coeficientes de wavelet.

Observando-se as Figuras 3.6 e 3.7, pode-se verificar que o coeficiente D_1 compreende

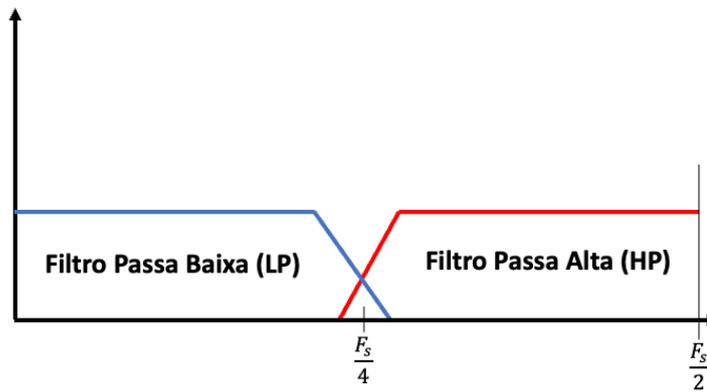


Figura 3.6: Divisão do espectro após passagem pelo nível 1 da TDW.

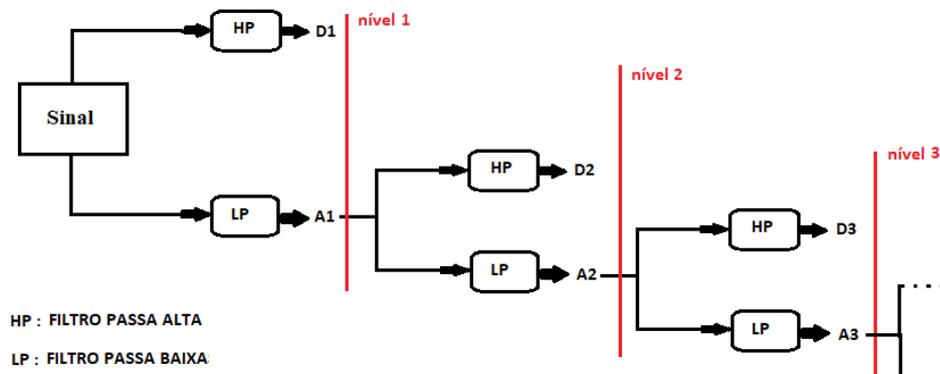


Figura 3.7: TDW visualizada como um banco de filtros.

a banda de frequência entre $F_s/4$ e $F_s/2$ Hz do espectro original. Por outro lado, o nível de aproximação A_1 carrega baixas frequências, cuja banda limita-se a $F_s/4$ Hz [29]-[30]. Após a decomposição multi-nível da TDW, os coeficientes podem ser utilizados para representar a energia carregada por cada banda do sinal [11]-[22]-[31]. Assim, ao utilizar a energia média dos coeficientes de wavelet, espera-se que cada classe de sinal tenha sua própria assinatura espectral, conforme ilustra a Figura 3.8.

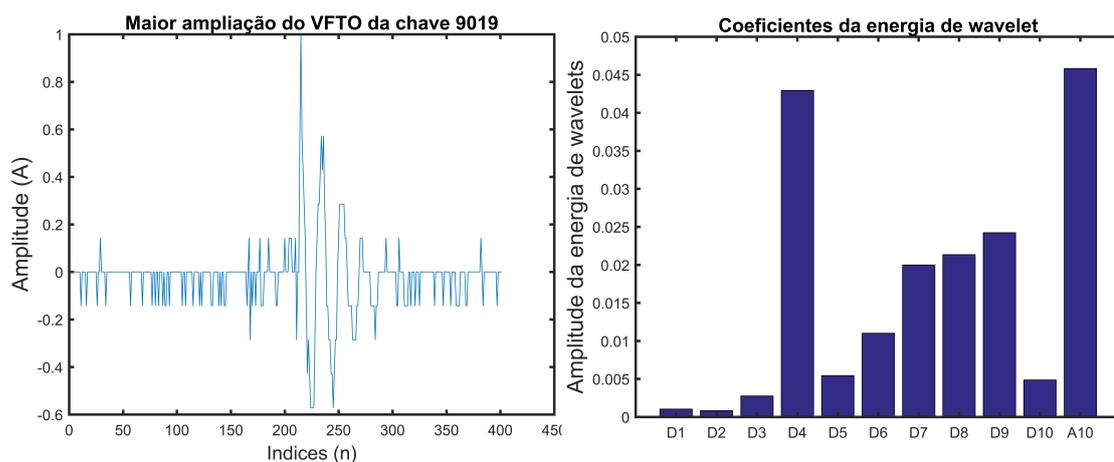


Figura 3.8: Amostra VFTO no tempo versus energia dos coeficientes da TDW.

A *wavelet mother* escolhida para realização deste trabalho foi a da família *Daubechies*, pois sendo essa família de funções Wavelet similar a um sinal transitório e a TDW uma convolução, esta apresenta melhores resultados. Além disso, a *wavelet Daubechies* foi utilizada pelas referências [8] e [11], onde a extração dos coeficientes da TDW foram analisados e utilizados na classificação de distúrbios de energia. A Figura 3.9 ilustra a semelhança visual dos sinais. Entretanto, não há técnicas bem esclarecidas sobre qual variação da família wavelet, ou momento de desvanecimento (*vanishing moments*), resulta em melhores desempenhos de acurácia na etapa de classificação. Portanto, visando uma investigação a respeito de qual momento de desvanecimento resultará em uma maior porcentagem de acurácia na classificação de VFTOs, foram adotados testes empíricos de desempenho (*trial and error*) [32].

Durante a criação das entradas que caracterizam o sinal, foram testadas variações no momento de desvanecimento de ordem 4, 8, 12 e 16 da família *Daubechies*. Ao todo foram testados 23 atributos, compostos por onze energias médias dos coeficientes de detalhe do nível 1 ao 10 e o nível 10 de aproximação. Além das respectivas entropias dos coeficientes (*Wavelet energy entropy-WEE*) e o valor eficaz (RMS) do VFTO [10]-[33], conforme exposto na Tabela de acrônimos 3.3.

A TW, separa o espectro de forma não linear a cada filtragem do sinal de Aproximação. A Tabela 3.4 mostra a banda respectiva de cada coeficiente da transformada.

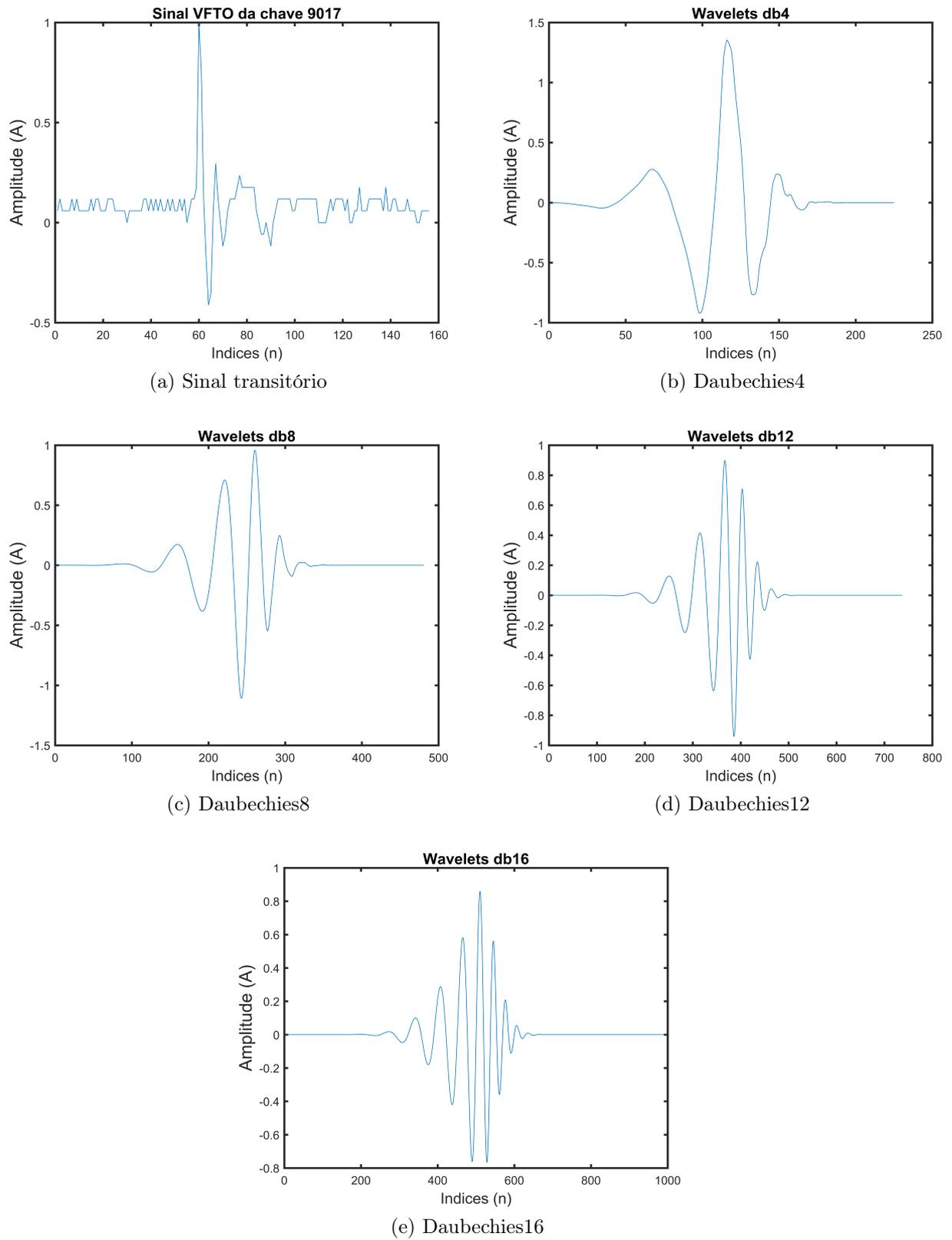


Figura 3.9: Similaridade entre sinal VFTO e wavelet Daubechies

<i>Feature</i>	Acrônimo
Energia média detalhe1	Em1
Energia média detalhe2	Em2
Energia média detalhe3	Em3
Energia média detalhe4	Em4
Energia média detalhe5	Em5
Energia média detalhe6	Em6
Energia média detalhe7	Em7
Energia média detalhe8	Em8
Energia média detalhe9	Em9
Energia média detalhe10	Em10
Energia média aproximação10	EmA10
Entropia de Wavelet coeficiente detalhe1	WEE1
Entropia de Wavelet coeficiente detalhe2	WEE2
Entropia de Wavelet coeficiente detalhe3	WEE3
Entropia de Wavelet coeficiente detalhe4	WEE4
Entropia de Wavelet coeficiente detalhe5	WEE5
Entropia de Wavelet coeficiente detalhe6	WEE6
Entropia de Wavelet coeficiente detalhe7	WEE7
Entropia de Wavelet coeficiente detalhe8	WEE8
Entropia de Wavelet coeficiente detalhe9	WEE9
Entropia de Wavelet coeficiente detalhe10	WEE10
Entropia de Wavelet coeficiente aproximação10	WEEA10
Valor eficaz	RMS

Tabela 3.3: Entradas e seus acrônimos

Coeficiente	Banda (Hz)
Detalhe1	5M até 10M
Detalhe2	2.5M até 5M
Detalhe3	1.25M até 2.5M
Detalhe4	625k até 1.25M
Detalhe5	312.5k até 625k
Detalhe6	156.25k até 312.5k
Detalhe7	78.125k até 156.25k
Detalhe8	39k até 78.125k
Detalhe9	19.53k até 39k
Detalhe10	9.765k até 19.5k
Aproximação10	5k até 9.765k

Tabela 3.4: Bandas dos coeficientes de wavelets

Por fim, a menor resolução espectral da base de dados é igual a 5kHz pelo coeficiente (A_{10}), pois foi feita uma filtragem digital do tipo passa alta, obtendo-se portanto, a faixa de frequência de 5kHz até 10MHz. Tal faixa é composta de média e altas frequências que são o foco de estudo deste trabalho, pois são componentes encontradas nos transitórios de manobras [4]. Todos os atributos foram extraídos utilizando o *software* Matlab.

3.3 Seleção de atributos (*Feature Selection*)

Seleção de atributos, ou *Feature Selection*, é um método cujo objetivo é a seleção dos atributos mais relevantes dada a tarefa específica de classificação [34]-[24]. Nesse sentido, técnicas de seleção de atributos foram desenvolvidas para mitigar os problemas de super dimensionamento e/ou redundância de dados de entrada que implicam diretamente em uma redução de acurácia, além de aumento no tempo de convergência na etapa de treinamento do algoritmo de aprendizado de máquina [34]-[34]-[35]. Deste modo, foram utilizadas métodos de seleção de atributos buscando um subconjunto que otimize a acurácia da classificação dos transitórios VFTOs feita pelos algoritmos de ML, enquanto que, em paralelo, há a redução do número de entradas utilizadas [35]-[36].

3.3.1 Métodos de seleção de atributos utilizados

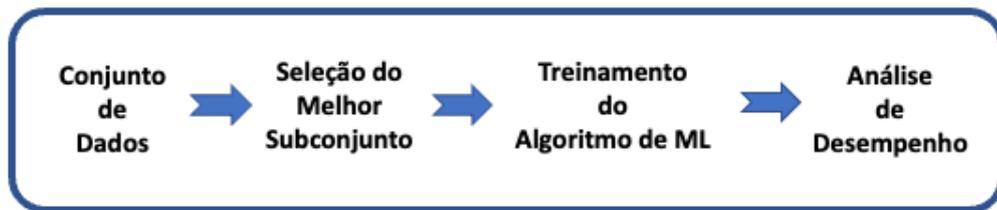
A seleção do subconjunto, em um primeiro momento, foi feita pelo método do tipo filtragem (*filter*) e, em seguida, os melhores resultados, pelo método do tipo enrolamento (*wrapper*). Os melhores resultados apresentados pelos conjuntos de entrada completo ou filtragem foram selecionados através da política "ganhador leva tudo". Em outros termos, foram feitos testes de desempenho de acurácia de uma tarefa específica de classificação e o conjunto que obteve uma maior taxa de acerto foi o escolhido para uma segunda otimização de desempenho pelo método de seleção do tipo enrolamento.

O método de seleção de *features* do tipo filtragem foi utilizado com o propósito de investigar subconjuntos cujo desempenho de classificação fosse melhor ou igual ao conjunto que engloba todas as 23 entradas. Além disso, uma vez que um subconjunto obteve a melhor avaliação de acurácia, este reduz o custo computacional do método enrolamento, que é utilizado em sequência buscando-se obter uma otimização de acurácia e/ou redução no conjunto de atributos utilizados.

A seleção de atributos do tipo filtragem funciona ranqueando a relevância dos atributos para uma determinada tarefa de classificação [37]. Desta maneira, o ranqueamento é feito utilizando-se critérios e regras matemáticas para mensurar a relevância de uma determinada entrada. Entretanto, após feito o ranqueamento dos atributos, deve-se escolher um limiar de corte separando as entradas selecionadas e as que devem ser excluídas do conjunto [38]. Assim, foi escolhido um limiar cuja regra foi a exclusão das 5 entradas (cerca de 20%) como as menos significativas pelos métodos *filter*. Ao todo foram utilizados os algoritmos:

- Correlação de Pearson;
- OneR;
- ReliefF;
- InfoGain;
- GainRatio;
- Symmetrical uncertain

A Figura 3.10 exemplifica o fluxo de como foi utilizado o método *filter* e posteriormente a performance de classificação do subconjunto.

Figura 3.10: Fluxo do método *filter*.

Por outro lado, o método enrolamento utiliza como critério de seleção o próprio algoritmo classificador, isto é, busca-se um subconjunto ótimo dos dados de entrada por exaustão [36]-[38]-[35]. Apesar da desvantagem do custo computacional que cresce quanto maior for o conjunto de características a ser avaliado, o *wrapper* geralmente retorna um subconjunto que desempenha uma melhor acurácia, pois este encontra um subconjunto que otimiza o desempenho de acurácia através de exaustivas combinações de atributos. Assim, a seleção do tipo enrolamento retira ou adiciona atributos até que seja alcançado a condição de parada do algoritmo [38]. A Figura 3.11 abaixo ilustra o mecanismo de funcionamento do método de seleção de atributos tipo wrapper

Figura 3.11: Fluxo do método *Wrapper*.

O método *wrapper* utilizado foi o *WrapperSubsetEval* no ambiente de desenvolvimento WEKA [39]. Os algoritmos de ML de teste para a busca do melhor subconjunto são os mesmos que serão avaliados quanto a seu desempenho de classificação. Logo, ao utilizar ambos os métodos de seleção de atributos em série, filtragem com o enrolamento, ocorre uma redução no custo computacional do segundo método, pois há previamente uma redução do conjunto de entradas pelo método filtragem.

3.4 Aprendizado de máquina (*Machine learning*)

A interpretação de dados está geralmente relacionada com a tarefa de classificação. Nesse sentido, é natural buscar um relacionamento dos dados com classes de observações anteriores [40]. Entretanto, à medida que o conjunto de dados aumenta e o problema de classificação fica mais complexo, a análise por técnicas convencionais já não produzem uma acurácia satisfatória [24]. Já os algoritmos de ML, nesse sentido, podem manipular grandes quantidades de dados e construir modelos representativos de um processo de maneira rápida e efetiva [41].

Os algoritmos de ML testados para fazer parte do sistema autônomo de monitoramento de transitórios foram: redes neurais artificiais do tipo Perceptron múltiplas camadas (PMC), as redes neurais de função base radial (RBF), máquina vetor suporte (MVS) e *Random forest* (RF). Todos estes utilizam o treinamento supervisionado na geração de seus modelos de classificação. Portanto, nesta categoria de treinamento, é necessária uma base de dados rotuladas de modo a relacionar as entradas com as saídas [41].

3.4.1 Redes neurais artificiais

As redes neurais artificiais (RNA) que surgiram na década de 40 são algoritmos que simulam o funcionamento do neurônio biológico. As RNA mais populares são as de propagação do tipo *feedforward*, onde elementos não lineares (neurônios), dispostos em camadas, fazem a criação de um hiperplano para a tarefa de divisão das classes e, assim, a classificação dos dados [40]-[42]. A Figura 3.12 exemplifica a topologia das RNA do tipo PMC e RBF utilizadas.

Na Figura 3.12, observa-se que os sinais de entradas (x_i) dos respectivos neurônios, as quais representam informações sobre os sinais VFTOs são ponderadas pelos pesos sinápticos (W_{ji}) da entrada i e neurônio j . Os valores de entradas em cada neurônio (j) da primeira camada escondida podem ser modelados pela Equação 3.2

$$I_j = \sum_{i=1}^N W_{ji} \cdot x_i \quad (3.2)$$

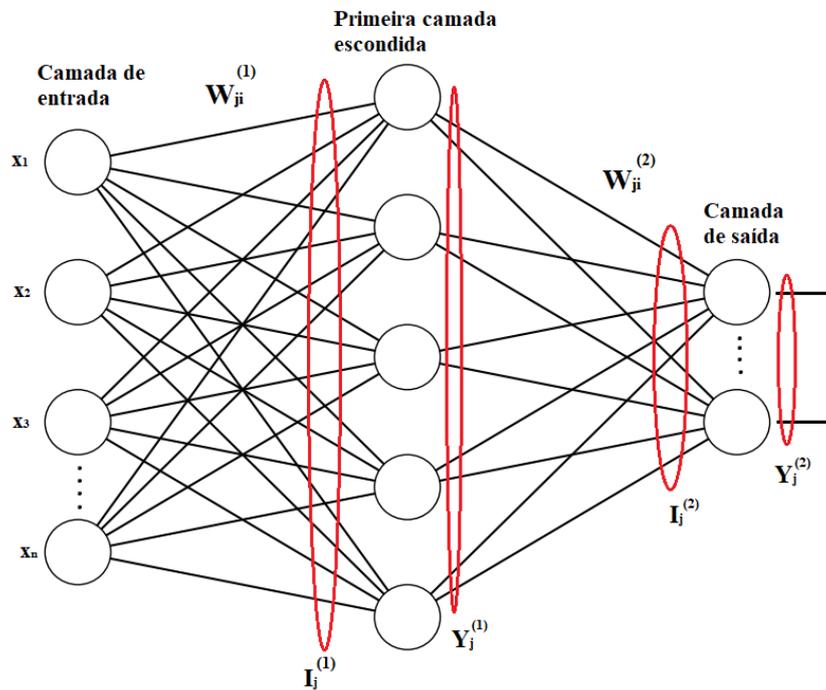


Figura 3.12: Topologia de uma RNA.

Assim, cada neurônio recebe o sinal de entrada e o processa de modo a encaminhar uma resposta excitatória ou inibitória para os neurônios da camada seguinte. O processamento em si é feito pela função de ativação da rede ($g(\cdot)$), onde utiliza-se comumente as funções tangente hiperbólica ou logística nas redes neurais do tipo PMC e gaussiana do tipo RBF. A saída dos neurônios da primeira camada são equacionados conforme a Equação 3.3.

$$Y_j = g(I_j) \quad (3.3)$$

Deste modo, a saída da última camada representa a classificação do sinal VFTO. Entretanto, a rede precisa primeiramente ser treinada de modo a mapear o processo de entrada/saída. Em termos matemáticos, a rede procura minimizar a função de custo dada pela Equação 3.4.

$$E(k) = \frac{1}{2} \sum_{j=1}^K (d_j(k) - Y_j(k))^2 \quad (3.4)$$

onde, $d_j(k)$ representa a saída esperada do neurônio j apresentando-se a k -ésima amostra de treinamento.

Assim, a RNA tem como objetivo do processo de treinamento, ajustar a matriz de pesos sinápticos a fim de minimizar o erro quadrático entre a saída produzida e a saída desejada, conforme mostra a Equação de ajuste dos pesos 3.5.

$$W_{ji}(t+1) = W_{ji}(t) + \eta \delta_j Y_i \quad (3.5)$$

$$\delta_j = (d_j - Y_j)g'(I_j) \quad (3.6)$$

onde W_{ji} é o peso do j -ésimo neurônio ligado a i -ésima entrada, η é a taxa de aprendizado, geralmente variando entre $[0.1, 0.9]$, δ o gradiente local definido por 3.6 e Y_i a i -ésima entrada. Deste modo, em cada camada da rede há um ajuste dos pesos de forma a minimizar o erro. O processo de ajuste da RNA é conhecido como *backpropagation* [43].

3.4.2 Máquina de vetor suporte

A MVS é um algoritmo que pode ser usado para tarefas de classificação de VFTOs e seu funcionamento se inicia com a identificação das amostras de borda de um aglomerado de amostras de mesma classe. Essas amostras de borda, chamadas de vetores suportes, são fundamentais no procedimento de treinamento do algoritmo cujo objetivo é encontrar um hiperplano que seja capaz de dividir as amostras de diferentes classes. O hiperplano é construído perpendicularmente as amostras de vetores suporte distanciados de forma otimizada a fim de garantir que o hiperplano esteja equidistante das amostras vetores suporte. Com isso, delimita-se uma distância de guarda para classificações de futuras amostras, conforme mostra a Figura 3.13 [40].

O hiperplano construído na fase de treinamento é uma função de domínio R^n , onde n é o número de atributos da base de dados que descreve os sinais VFTOs. Então, o algoritmo MVS procura o plano ótimo, descrito pela Equação 3.7.

$$W^T x_i + b = 0 \quad (3.7)$$

onde x_i é o vetor de atributos que irá representar o hiperespaço R^n , o vetor W^T são os coeficientes a serem ajustados no processo de treinamento e o coeficiente b é o bias da

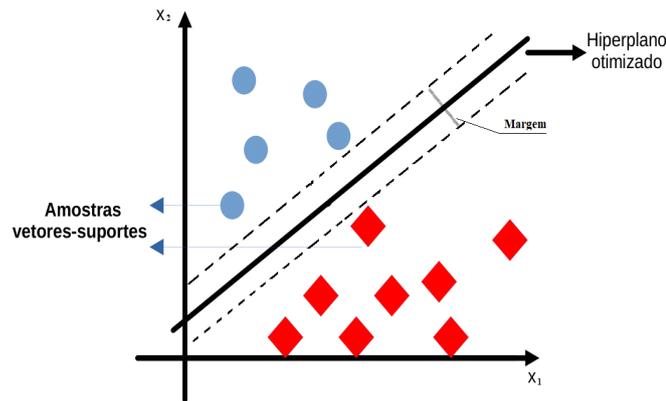


Figura 3.13: Entidades do algoritmo MVS.

função.

O algoritmo MVS também se mostra robusto em casos onde os dados não apresentam uma separação linear. Nesse caso o algoritmo utiliza a transformação dos dados para que haja uma separação linear no novo mapeamento, esta técnica de transformação em um novo domínio é conhecida como truque de kernel (*kernel trick*) [43].

3.4.3 *Random Forest*

O algoritmo Random Forest (RF) é um classificador expresso por um conjunto de n -árvores de decisão, conforme ilustra a Figura 3.14. O algoritmo de árvore de decisão possui uma topologia que consiste de nós. Estes podem ser: nó raiz (topo), nós internos (meio) e os nós folhas (base). O fluxo da árvore de decisão se inicia no topo e termina na base, em um nó folha. Desta forma, a classificação de cada amostra é feita testando-se condições dos atributos dos VFTOs e, então, direcionando-as ao nós internos para que seja feita uma nova testagem de condição até que se chegue a um nó folha, onde a amostra é finalmente classificada [41]. Por ser composta de um conjunto de árvores de decisão, a classificação final do RF é feita pela pelo voto majoritário das n -árvores de decisão do conjunto.

O nome aleatório (*Random*) do algoritmo é devido ao processo de inicialização randômica (*bootstrap*) na geração das árvores de decisão que irão compor o algoritmo. Nesse processo é escolhido ao acaso K amostras da base de dados dos sinais transitórios e J atributos que servirão como condições a serem testadas na divisão dos nós. Essa escolha aleatória é feita de modo a evitar a criação de árvores com alto grau de correlação [41].

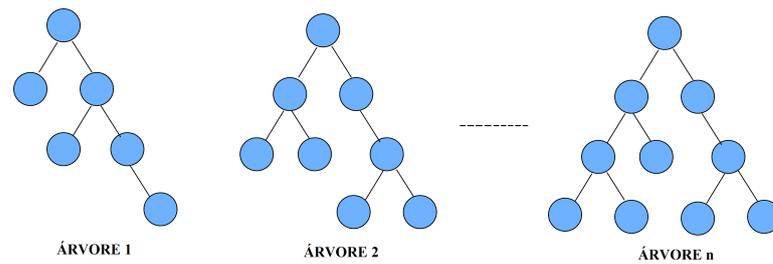


Figura 3.14: Árvores de decisão compondo um *Random Forest*.

Um dos métodos de aprendizado, isto é, criação de regras de separação dos nós, é dado pela otimização do ganho de informação (Ig) na divisão do nó pai (nó acima) e do nós filhos (nós abaixo), como mostra a Equação 3.8.

$$Ig = E(\text{pai}) - \sum_{i=1}^N w_i E(\text{filho}_i) \quad (3.8)$$

onde $E(\cdot)$ é a função entropia, descrito na Equação 3.9, e pode ser calculada pela probabilidade $P(x_i)$ de uma amostra VFTO escolhida ao acaso pertencer a i -ésima classe. Por fim, w_i é a porcentagem do total de amostras disponíveis no nó filho em relação ao nó pai após a divisão dos ramos.

$$E(\text{no}) = \sum_{i=1}^N P(x_i) \log P(x_i) \quad (3.9)$$

A Figura 3.15 exemplifica o processo do cálculo da escolha das condições de separação dos nós pelo atributo que retorna o maior ganho de informação da Equação 3.8. Como pode ser observado, o processo da escolha dos nós inicia-se calculando a entropia do nó pai ($E(\text{pai})$) que representa o conjunto das amostras de VFTOs disponíveis na base de dados. Em seguida, é escolhido um atributo, por exemplo energia dos coeficientes de wavelet, e é aplicado uma condição de separação que irá direcionar a amostra para o nó filho 1 ou nó filho 2 onde são calculadas as entropias E_1 e E_2 e a proporção da quantidade de amostras W_1 e W_2 respectivamente.

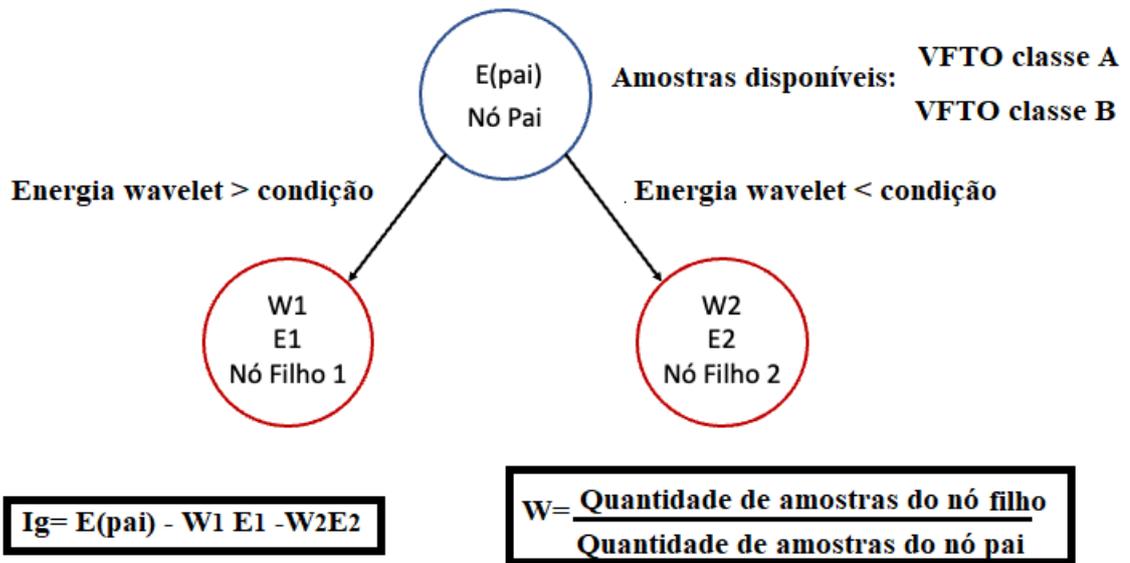


Figura 3.15: Cálculo do ganho de informação após divisão dos nós.

Esse procedimento é repetido para todos os atributos e testando-se diferentes condições de separação dos nós, as melhores condições serão aquelas que retornarão o maior valor da função I_g . Em suma, são realizadas comparações de modo a reduzir a incerteza dos nós filhos.

Capítulo 4

Resultados

Para geração dos resultados apresentados foram avaliados 4 algoritmos de *machine learning*: *Random forest*(RF), Redes Neurais Artificiais do tipo Perceptron múltiplas camadas (PMC) e de base radial (RBF) e máquina vetor suporte (MVS). Juntos esses algoritmos foram testados com diferentes momentos de desvanecimento 4, 8, 12 e 16 da família de Wavelet Daubechies. Após esse primeiro momento de avaliação dos classificadores utilizando todas as *features* de entradas e com um subconjunto selecionado pelo método de seleção de atributos *filter*, os conjuntos de entradas que tiveram melhor desempenho de acurácia foram otimizados pelo método *Wrapper*. Assim, em um segundo momento, foi realizado novamente um teste de desempenho. O roteiro seguido foi elaborado de forma a buscar a maior porcentagem de acerto na classificação pelos algoritmos de ML e redução de custo computacional no método *Wrapper*, uma vez que este funciona por testes de exaustão.

4.1 Análise de desempenho

A tarefa de classificação pode ser simplificada dividindo-se em sub-tarefas. Essa divisão diminui a complexidade de generalização feita pelos algoritmos. As sub-tarefas fora organizadas da seguinte forma:

1. Chave seccionadora *versus* disjuntor.
 - 1.1. Caso seja chave seccionadora:
 - 1.1.1. Identificação da chave seccionadora operada dentro da topologia.
 - 1.1.2. Identificação da manobra realizada pela chave (abertura *vs* fechamento).

1.2. Caso seja disjuntor:

1.2.1. Identificação do disjuntor operado dentro da topologia.

1.2.2. Identificação da manobra realizada pelo disjuntor (abertura *vs* fechamento).

O diagrama em árvore de decisão ilustrado na Figura 4.1 mostra as tarefas e o fluxo proposto para uma completa classificação do sistema de monitoramento proposto. O sistema deve possuir a capacidade de adquirir sinais VFTO no tempo e transformá-los, de modo automático, em dados de entrada no estágio de Extração de atributos. Assim, a completa classificação do evento transitório fornece a equipe de manutenção uma ferramenta de investigação e diagnóstico de falhas pós-manobras, diminuindo o tempo de religamento da subestação após a interrupção no fornecimento de energia 3.3.

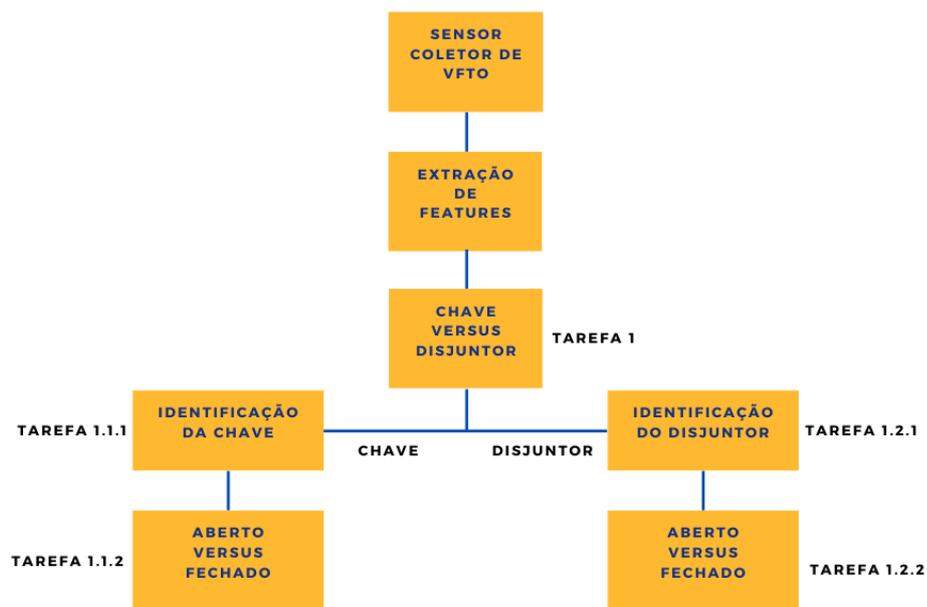


Figura 4.1: Diagrama em árvore do fluxo de classificações.

Os resultados das respectivas tarefas do sistema de monitoramento ilustradas na Figura 4.1 foram avaliados pelo método *K-fold cross validation* com $k = 5$. O método separa os dados de treinamento em 5 partes, onde 4 partes são utilizadas para treinamento [43]. Assim, o método permuta todas as partes por 5 vezes de modo que cada parte seja utilizada no treinamento 4 vezes e 1 vez como teste, o que retorna maior robustez aos resultados pois são apresentados a média dos 5 testes [44]-[45]. Os algoritmos de ML avaliados foram considerando o software WEKA [39].

4.1.1 Chave *versus* Disjuntor (Tarefa 1)

De posse do conjunto de dados de entrada das classes (chaves e disjuntores), a Tarefa 1, conforme mostra a Figura 4.1, consiste em classificar chaves versus disjuntores e encaminhar os dados para a próxima tarefa de classificação a fim de identificar a numeração do equipamento e o tipo de manobra realizada. O conjunto de treinamento utilizado nesta tarefa foi de 912 amostras, sendo 855 de chaves (93.75% do total) e de 57 disjuntores (6,25% do total). A Tabela 4.1 mostra os resultados obtidos para cada algoritmo de aprendizado de máquina e para cada momento de desvanecimento da família Wavelet Daubechies. Os resultados da Tabela 4.1, foram avaliados usando a métrica de *acurácia clássica* (Ac) [46]-[47], Equação 4.1. De maneira qualitativa, os resultados de acurácia desejáveis devem ser superiores a 93.75% de acerto. Esse limiar é obtido classificando-se todas as amostras como sendo parte da classe majoritária (classe modal). Assim, por ser um lógica/algoritmo sem capacidade de aprendizado e/ou adaptação, a métrica é usada como forma de *benchmark* para a avaliação [48] de algoritmos de ML.

$$Ac = \frac{T_P + T_N}{T_P + F_N + F_P + T_N} \quad (4.1)$$

onde T_P , T_N , F_N e F_P são respectivamente os valores das classificações: Verdadeiro positivo, Verdadeiro negativo, Falso negativo e Falso positivo da matriz de confusão mostrada na Figura 4.2.

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

Figura 4.2: Matriz de confusão.

desempenho Tarefa 1 pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	97.1	98	98	97.3
	subconjunto <i>filter</i>	↓	↓	98.1	97.7
RNA (RBF)	conjunto total	94.2	93.6	95.3	95.4
	subconjunto <i>filter</i>	94.6	95.4	↓	95.7
MVP	conjunto total	95.2	96.5	95.7	96.4
	subconjunto <i>filter</i>	↓	↓	↓	↓
Random Forest	conjunto total	97.9	97.4	97.8	97.7
	subconjunto <i>filter</i>	98.2	97.6	98.1	=
Entradas removidas pelo método <i>filter</i>		Em1 Em8 Em9 Em10 WEE10	Em1 Em9 Em10 WEE5 WEE9	Em5 Em9 Em10 WEE7 WEE9	Em1 Em9 Em10 EmA10 WEE2
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.1: Resultados Tarefa 1 - Acurácia clássica.

A Tabela 4.1 mostra que a combinação com melhor acurácia foi o momento de desvanecimento número 4 (db4) junto com o algoritmo *Random Forest*. Essa combinação apresentou ainda uma melhora em sua desempenho após a retirada das *features* Em1, Em8, Em9, Em10 e WEE10 que foram ranqueadas como menos significativas para o processo de classificação. O processo quantitativo do ranqueamento utilizado é similar a de

uma corrida/*regatta* de barco a vela, onde as *features* são como os barcos competidores e a cada corrida, ou neste caso a cada um dos 6 algoritmos do método *filter*, essas *features* acumulam pontos iguais ao da sua colocação/ranqueamento conforme Equação 4.2.

$$Rank = \sum_{n=1}^6 P_i \quad (4.2)$$

onde P_i é a posição do ranqueamento feito pelo método de seleção de atributos do software WEKA [39].

Embora de maneira geral tenha ocorrido um aumento de amostras após a aplicação da técnica de *Data Augmentation*, a distribuição do banco de dados entre as classes *chaves* versus *disjuntores* continua desbalanceada conforme ilustrado na Figura 4.3. Portanto, outras métricas se mostram mais úteis na avaliação de desempenho de *datasets* cujas classes são desbalanceadas [47]-[46].

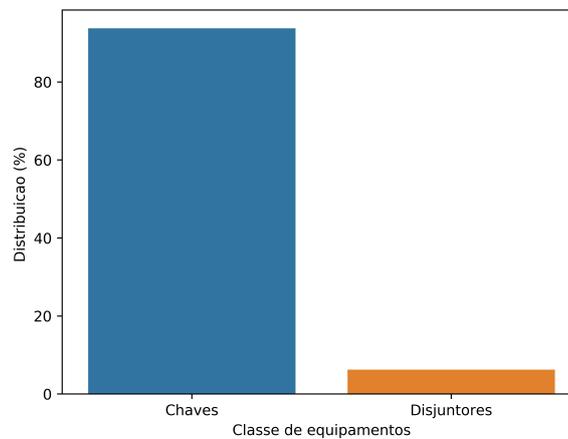


Figura 4.3: Distribuição Chave vs Disjuntores.

Então, em um segundo momento, foi utilizado a métrica F1, conforme pode ser visto na Equação 4.3, também conhecida como média harmônica, que melhor se ajusta a um *dataset* desbalanceado. [47]-[46].

$$F1 = \frac{2(Pr.Re)}{Pr + Re} \quad (4.3)$$

Onde *Precision*(Pr) e *Recall*(Re) são outras métricas de desempenho, que podem ser calculadas pelas Equações(4.4) e (4.5), e carregam informações sobre falsos

negativos e os falsos positivos, respectivamente. A Tabela 4.2 mostra a nova abordagem de análise observando-se a métrica $F1$.

$$Re = \frac{T_P}{T_P + F_N} \quad (4.4)$$

$$Pr = \frac{T_P}{T_P + F_P} \quad (4.5)$$

desempenho Tarefa 1 pela métrica F1 (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	97.1	98.0	97.9	97.1
	subconjunto <i>filter</i>	↓	=	98	97.6
RNA (RBF)	conjunto total	93.8	93.4	94.7	94.7
	subconjunto <i>filter</i>	94.5	95.2	↓	↓
MVP	conjunto total	93.9	95.9	94.9	95.7
	subconjunto <i>filter</i>	↓	↓	↓	↓
Random Forest	conjunto total	97.7	97	97.6	97.5
	subconjunto <i>filter</i>	98.1	97.6	98	=
Entradas removidas pelo método <i>filter</i>		Em1	Em1	Em5	Em1
		Em8	Em9	Em9	Em9
		Em9	Em10	Em10	Em10
		Em10	WEE5	WEE7	EmA10
		WEE10	WEE9	WEE9	WEE2
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.2: Resultados Tarefa 1 - métrica F1.

Em ambas as métricas de avaliação a combinação que continuou apresentando maior porcentagem de acerto foi a de momento de desvanecimento número 4(db4) com o algoritmo *Random Forest* com a retirada de *features* menos significativas.

Por fim, em uma terceira análise, foi adotado a técnica *Undersampling* que consiste na

redução aleatória de amostras da classe que contém o maior número de dados (que nesse caso é a classe das chaves seccionadoras) até que o número de amostras das duas classes se igualem [49]-[50]-[51]. Esta remoção aleatória das amostras foi realizada apenas uma vez. Os resultados da avaliação por *Undersampling* são expostos na Tabela 4.3.

desempenho Tarefa 1 com <i>Undersample</i> pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	80.2	86.5	88.6	88.8
	subconjunto <i>filter</i>	85.4	89	91.2	↓
RNA (RBF)	conjunto total	84.5	82.2	83.3	86.2
	subconjunto <i>filter</i>	↓	↓	=	87
MVP	conjunto total	84.5	83	86.8	88.8
	subconjunto <i>filter</i>	=	84.7	=	↓
Random Forest	conjunto total	91.4	87.3	93	89.6
	subconjunto <i>filter</i>	94	=	94.7	91.4
Entradas removidas pelo método <i>filter</i>		Em1	Em1	Em1	Em1
		Em10	Em10	Em5	Em10
		EmA10	EmA10	EmA10	WEE3
		WEE10	WEE5	WEE3	WEE5
		WEEA10	WEEA10	WEE5	WEEA10
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.3: Resultados Tarefa 1 - *Undersample*

O melhor desempenho entre os resultados da Tabela 4.3 apresenta uma combinação de momento de desvanecimento número 12 (db12) e o algoritmo *Random Forest* com remoção das *features* EM1, EM5, EMA10, WEE3 e WEE5. Este resultado apresenta um número de momento de desvanecimento diferente das Tabelas 4.1 e 4.2 apesar da manutenção da melhor desempenho pelo algoritmo *Random Forest*. A combinação de melhor desempenho da Tabela 4.3 pode ser explicada pela retirada aleatória das amostras da classe

majoritária (classe com maior número de amostras), assim, essa retirada aleatória pode ter beneficiado o *dataset* do experimento com o momento de desvanecimento número 12 (db12). Entretanto, apesar dos resultados divergirem no momento de desvanecimento, o algoritmo *Random Forest* continuou apresentando o melhor desempenho de acurácia. Por fim, o momento de desvanecimento número 4 (db4) continuou apresentando desempenho comparativamente satisfatório na Tabela 4.3, com diferença na acurácia de apenas 0.7 % do momento de desvanecimento número 12 (db12), demonstrando assim, que ambas as soluções são competitivas em uma implementação do projeto de monitoramento.

4.1.2 Identificação da Chave (Tarefa 1.1.1)

A partir da informação que a amostra é uma chave, a Tarefa 1.1.1 tem como objetivo identificar qual foi a chave manobrada. As 851 amostras VFTO das chaves são provenientes de 8 chaves. Estas compõem o *dataset* com a seguinte distribuição: 9017(11.28%), 9019(13.63%), 9029(12.7%), 9117(14.69%), 9119(12.22%), 9127(11.98%), 9129(11.75%), 9005(11.75%). A distribuição percentual dessas amostras é mostrada na Figura 4.4.

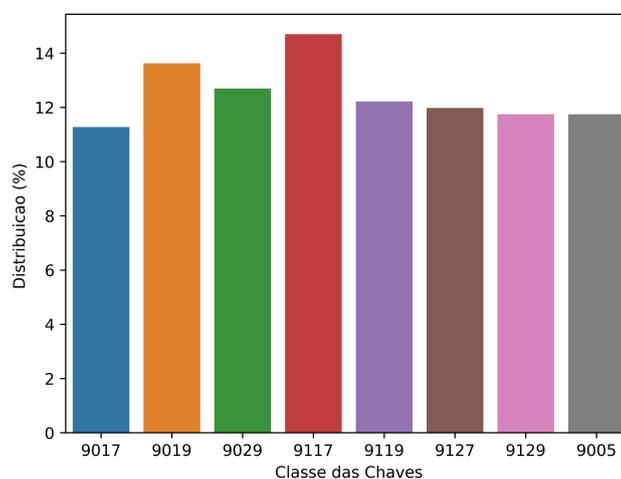


Figura 4.4: Distribuição da classe Chaves Seccionadores.

A maior diferença no percentual entre as distribuições da classe das chaves é de 3.4%, correspondendo à diferença das chaves 9017 e 9117. Além disso, observa-se na Figura 4.4 que as amostras estão distribuídas de forma balanceada, logo, a métrica de avaliação utilizada foi a de acurácia clássica conforme a Equação 4.1. Os resultados da Tarefa 1.1.1 são mostrados na Tabela 4.4 abaixo.

desempenho Tarefa 1.1.1 pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	87	89.2	91.2	89.5
	subconjunto <i>filter</i>	89.7	90.5	91.7	91.1
RNA (RBF)	conjunto total	55	50	48.5	52.6
	subconjunto <i>filter</i>	56	52.6	51.1	53
MVP	conjunto total	76.6	83.9	82.6	82.2
	subconjunto <i>filter</i>	↓	84.4	↓	↓
Random Forest	conjunto total	91	92.4	93.4	92.7
	subconjunto <i>filter</i>	91.8	92.8	↓	93.2
Entradas removidas pelo método <i>filter</i>		EmA10	EmA10	EmA10	EmA10
		WEE8	WEE8	WEE8	WEE8
		WEE9	WEE9	WEE9	WEE9
		WEE10	WEE10	WEE10	WEE10
		WEEA10	WEEA10	WEEA10	WEEA10
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.4: Resultados Tarefa 1.1.1

O resultado da Tabela 4.4 mostra que o algoritmo de melhor combinação de desempenho para realização da Tarefa 1.1.1, pela métrica da acurácia clássica, foi o algoritmo *Random Forest* com utilização do momento de desvanecimento número 12 (db12) sem remoção de nenhuma das entradas/features. Apesar da retirada das features EmA10, WEE8, WEE9, WEE10 E WEEA10 referentes a db12 tornar o treinamento dos algoritmos de ML menos custosos, conforme mostra a Figura 4.5, a seleção da melhor combinação foi adotada pela política de maior desempenho em acurácia.

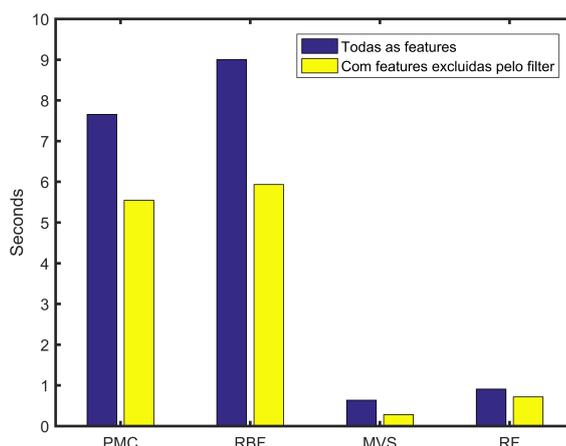


Figura 4.5: Custo computacional dos algoritmos com todas as 23 features *vs* 18 features.

Além disso, outro resultado interessante, é que todos os momentos de desvanecimento do algoritmo *Random Forest* foram acima dos 90%, destacando-o como a melhor alternativa para a Tarefa 1.1.1.

4.1.3 Identificação da manobra da chave (Tarefa 1.1.2)

A partir da informação que a amostra é uma chave e sua respectiva identificação, a Tarefa 1.1.2 tem como objetivo identificar a manobra realizada, isto é, discriminar entre abertura *versus* fechamento. O conjunto de dados da classe de chaves é composto de 851 amostras de VFTO da classe abertura(48.65%) *versus* fechamento(51.35%). As amostras são distribuídas de forma balanceada, conforme ilustra a Figura 4.6. Por se tratar de um *dataset* balanceado, a métrica de avaliação será a acurácia clássica conforme Equação 4.1. Os resultados da Tarefa 1.1.2 são mostrados na Tabela 4.5

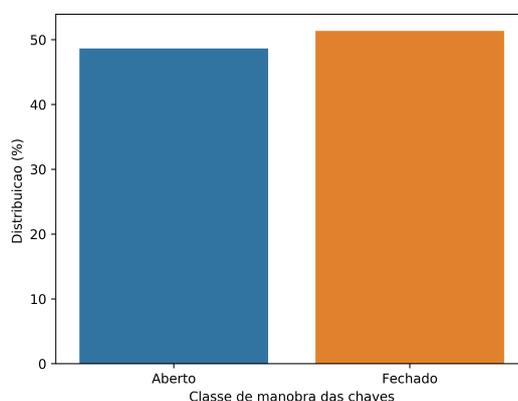


Figura 4.6: Distribuição classe Aberto *versus* Fechado das chaves

desempenho Tarefa 1.1.2 pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	62.7	63.7	66.7	66
	subconjunto <i>filter</i>	65.1	67.8	68.6	68
RNA (RBF)	conjunto total	61.6	61.2	60.1	62.1
	subconjunto <i>filter</i>	↓	63.7	63.3	64.6
MVP	conjunto total	62.2	63.3	62	62.2
	subconjunto <i>filter</i>	↓	↓	=	↓
Random Forest	conjunto total	72	72.2	72.7	71.1
	subconjunto <i>filter</i>	73.3	74.6	76.2	73.2
Entradas removidas pelo método <i>filter</i>		Em7	Em8	EmA10	Em8
		Em10	EmA10	WEE7	EmA10
		WEE7	WEE9	WEE9	WEE9
		WEE8	WEE10	WEE10	WEE10
		WEE9	WEEA10	WEEA10	WEEA10
Legenda					
↓: Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.5: Resultados Tarefa 1.1.2

O resultado da Tabela 4.5 mostra que o algoritmo de melhor desempenho, para realização da Tarefa 1.1.2, pela métrica da acurácia clássica, foi o algoritmo *Random Forest* combinado com o momento de desvanecimento 12(db12). O resultado obteve uma otimização com a remoção das entradas EmA10, WEE7, WEE9, WEE10 e WEEA10 de menor relevância pelos métodos *filter*. O aumento na acurácia após a remoção das entradas foi de 3.5 pontos percentuais, mostrando que as *features* removidas funcionavam como ruído durante o processo de aprendizado do algoritmo. Além disso, os resultados mostram que o algoritmo *Random Forest* obteve desempenho de acerto acima dos 70% em todas as condições. Portanto, na média, independente do momento de desvanecimento, o algoritmo *Random Forest* é o melhor para a Tarefa 1.1.2. Essa manutenção de acerto independente do momento de desvanecimento e obtenção do melhor resultado pela acurácia geral deve-se a robustez dada as múltiplas árvores de decisões implementadas dentro do algoritmo

RF.

4.1.4 Identificação do Disjuntor (Tarefa 1.2.1)

A partir da informação que a amostra é um disjuntor, a Tarefa 1.2.1 tem como objetivo identificar qual disjuntor foi acionado. De posse do conjunto de dados da classe de disjuntores, composto de 57 amostras de VFTO, dividido entre: 9118(42.1%) e 9128(57.9%). As amostras são distribuídas de forma balanceada, conforme ilustra a Figura 4.7. Por se tratar de um *dataset* balanceado, a métrica de avaliação será a acurácia clássica conforme Equação 4.1. Os resultados da Tarefa 1.1.2 é demonstrado através da Tabela 4.6.

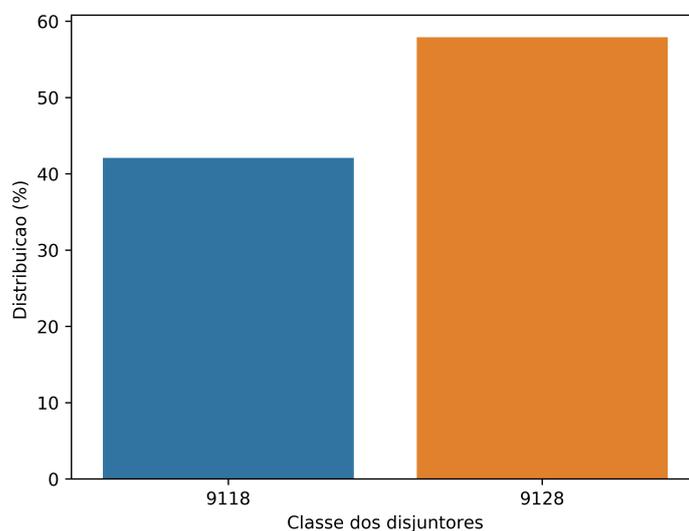


Figura 4.7: Distribuição classe dos disjuntores.

desempenho Tarefa 1.2.1 pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	87.7	87.7	93	93
	subconjunto <i>filter</i>	↓	93	94.7	94.7
RNA (RBF)	conjunto total	85.9	87.7	89.5	91.2
	subconjunto <i>filter</i>	87.7	93	=	↓
MVP	conjunto total	89.5	89.5	84.2	87.7
	subconjunto <i>filter</i>	91.2	↓	87.7	=
Random Forest	conjunto total	91.2	91.2	89.5	93
	subconjunto <i>filter</i>	=	96.5	93	=
Entradas removidas pelo método <i>filter</i>		Em10	Em3	Em3	Em10
		EmA10	WEE4	Em10	WEE4
		WEE1	WEE7	WEE4	WEE7
		WEE9	WEE10	WEE10	WEE10
		WEE10	WEEA10	WEEA10	WEEA10
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.6: Resultados Tarefa 1.2.1

O resultado da Tabela 4.6 mostra que o algoritmo de melhor desempenho, para realização da Tarefa 1.2.1, pela métrica da acurácia clássica, foi o algoritmo *Random Forest* com utilização do momento de desvanecimento 8 (db8) e remoção dos atributos Em3, WEE4, WEE7, WEE10 e WEEA10, considerados menos relevantes pelos métodos *filter*. A retirada dos atributos resultou em um aumento de 5.3% na acurácia, mostrando que os atributos retirados estavam funcionando como ruído no processo de generalização do algoritmo de ML.

4.1.5 Identificação da manobra dos disjuntores (Tarefa 1.2.2)

A partir da informação que a amostra é um disjuntor e identificado qual disjuntor foi manobrado, a Tarefa 1.2.2 tem como objetivo identificar qual manobra foi realizada. De posse do conjunto de dados da classe de disjuntores, composto de 57 amostras de VFTO divididos entre aberto(47.36%) *versus* fechado(52.6%). As amostras são distribuídas de forma balanceada, conforme ilustra a Figura 4.8. Por se tratar de um *dataset* balanceado, a métrica de avaliação será a acurácia clássica conforme Equação 4.1. Os resultados da Tarefa 1.1.2 são demonstrados através da Tabela 4.7.

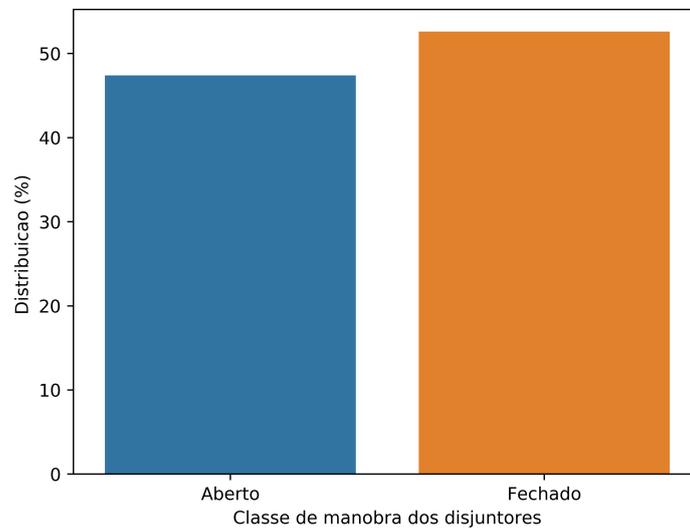


Figura 4.8: Distribuição classe aberto *versus* fechado dos disjuntores.

desempenho Tarefa 1.2.2 pela acurácia clássica (%)					
//		db4	db8	db12	db16
RNA (PMC)	conjunto total	61.4	77.2	82.4	80.7
	subconjunto <i>filter</i>	82.4	=	=	=
RNA (RBF)	conjunto total	66.7	73.7	61.4	70.2
	subconjunto <i>filter</i>	=	↓	70.2	=
MVP	conjunto total	75.4	66.7	73.7	72
	subconjunto <i>filter</i>	=	68.4	↓	79
Random Forest	conjunto total	71.9	84.2	75.5	72
	subconjunto <i>filter</i>	77.2	=	80.7	79
Entradas removidas pelo método <i>filter</i>		Em10	Em10	Em10	Em10
		WEE3	WEE3	EmA10	WEE3
		WEE4	WEE4	WEE3	WEE4
		WEE5	WEE10	WEE4	WEE10
		WEE10	WEEA10	WEEA10	WEEA10
Legenda					
↓ : Redução no desempenho					
= : desempenho se manteve constante					

Tabela 4.7: Resultados Tarefa 1.2.2

O resultado da Tabela 4.7 mostra que o algoritmo de melhor desempenho, para realização da Tarefa 1.2.2, pela métrica da acurácia clássica, foi o algoritmo *Random Forest* com utilização do momento de desvanecimento número 8 (db8). Além disso, observa-se que a remoção das entradas EM10, WEE3, WEE4, WEE10, WEEA10, ranqueadas como menos significativas no momento de desvanecimento número 8 (db8) não resultaram em um aumento de desempenho combinada com o algoritmo *Random Forest*. Entretanto, houve uma otimização, pois reduziu-se cerca de 20% do vetor de entradas utilizado para um mesmo desempenho de acurácia.

4.2 Otimização de desempenho pelo método *wrapper*

Após avaliação de desempenho, as combinações dos algoritmos de ML e momentos de desvanecimento que obtiveram melhores resultados são conduzidos para uma próxima etapa visando uma otimização no desempenho de classificação pelo método de seleção de atributos do tipo enrolamento *Wrapper*.

Entretanto, apesar de algumas Tarefas de classificação já apresentarem um ganho em acurácia e um conjunto reduzido pelo subconjunto do tipo filtragem (*filter*), o método do tipo (*wrapper*) procura, dado o conjunto de entradas, o subconjunto ótimo que maximize o desempenho de acurácia. O método de busca pelo subconjunto é feito por exaustão, ou seja, o algoritmo adiciona ou remove atributos de modo a cobrir todas permutações, e então, é feita a testagem e avaliação de desempenho de acurácia e escolhido o subconjunto que apresentar o melhor desempenho.

A Tarefa 1 é a única tarefa onde valeu-se dois cenários, pois foram realizados testes de desempenho com diferentes métricas e ambos se mostraram competitivos.

Apesar do método *wrapper* retornar o melhor subconjunto para uma tarefa específica de classificação, os subconjuntos previamente selecionados pelo método filtragem(*filter*) foram válidos para diminuir o custo computacional do método *wrapper*, uma vez que este funciona por exaustão, testando inúmeros conjuntos de entradas e, então, avaliando seu desempenho de acerto na classificação [37]-[38]. Além disso, nos casos em que a retirada das *features* resultou em uma otimização de desempenho, há indicação de que existem entradas irrelevantes.

Os resultados obtidos com o método *wrapper* tem a vantagem de retornar o melhor subconjunto. Porém, há limitações deste método, como: *overfitting* e alto custo computacional [38]. O resultado da otimização pelo método é apresentado na Tabela 4.8. A Tabela mostra que houve reduções no número de *features* em todos os casos e aumento na porcentagem de acurácia na maioria dos casos. As exceções de casos foram as Tarefas 1.1.2 e 1.2.1 que serão especialmente comentadas.

Tarefa/Algoritmo	desempenho com método wrapper (% de aumento)	Novo subconjunto de features	Antigo nº de features Novo nº de features (% de redução)
1(desbalanceado)/RF	98.5% (+0.3%)	Em2, Em5 Em7, WEE1 WEE2, WEE4 WEE7, RMS	18 8 (-55.5%)
1(balanceado)/RF	95.6% (+0.9%)	Em2, Em3 Em6, Em9 WEE4, WEE7 WEEA10, RMS	18 8 (-55.5%)
1.1.1/RF	94.2% (+0.8%)	Em1, Em3 Em4, Em6 Em8, WEE1 WEE2, WEE3 WEE4, WEE5 WEE6, WEE7 WEE8, RMS	23 14 (-39%)
1.1.2/RF	76.2% (0%)	Em1, Em2 Em3, Em4 Em5, Em7 Em8, WEE1 WEE2, WEE3 WEE5, WEE6 RMS	18 13 (-27.8%)
1.2.1/RF	94.7% (-1.8%)	Em4, Em7 WEE5, WEE6 WEE8, RMS	18 6 (-66.7%)
1.2.2/RF	89.5% (+5.3%)	Em6, WEE4 RMS	23 3 (-87%)

Tabela 4.8: desempenho com novos subconjuntos pelo método *wrapper*

A Tarefa 1.1.2 não apresentou aumento de acurácia, entretanto, o número de *features* foi reduzido em 27.8% mantendo-se a mesma acurácia. Já a Tarefa 1.2.1, apesar de não apresentar melhora na acurácia, houve uma redução de *features* na ordem de 67% com uma redução na acurácia de 1.8%. Deste modo, em caso de uma implementação do sistema de monitoramento, pode-se dizer que há um *trade-off* entre modelos a respeito da acurácia *versus* número de entradas/custo computacional.

4.3 Resumo de resultados

Em resumo a Tabela 4.9 expõe os resultados de maior acurácia entre todos os conjuntos de testes (subconjunto de todas as entradas, subconjunto pelo método *filter*, subconjunto pelo método *wrapper*). A Tarefa 1 (Chave vs Disjuntor) foi a única a apresentar duas soluções em paralelo, pois essa foi avaliada por duas métricas diferentes para os datasets desbalanceados e balanceados, métricas F1 e Acurácia Clássica (Ac) respectivamente.

Nota-se, que na coluna algoritmo, o *Random Forest* foi unânime em todas as Tarefas apesar de outros algoritmos apresentaram-se competitivos. Destaca-se que o algoritmo que apresentou a melhor acurácia não necessariamente é o aconselhável para aplicação em campo. Para isto, deve-se considerar o custo computacional de implementação e operação, pois em aplicações críticas, o tempo de resposta de um algoritmo de ML pode ser crucial para a operação de monitoramento.

Na terceira coluna constata-se que o momento de desvanecimento (db12) foi mais adequado para a maioria das Tarefas de classificação do sistema de monitoramento. Apresenta-se também, na última coluna, as *features* utilizadas para se obter a desempenho de acurácia relatada na Tabela. Por fim, Destaca-se a Tarefa 1.2.2 que reduziu para apenas 3 *features* (cerca de 87% de redução) após o processo de seleção de *features Wrapper*.

Tarefa	Algoritmo	Momento de desvanecimento	desempenho em acurácia(%)	Features
1(desbalanceado)	Random Forest	db4	98.5	Em2, Em5 Em7, WEE1 WEE2, WEE4 WEE7, RMS
1(balanceado)	Random Forest	db12	95.6	Em2, Em3 Em6, Em9 WEE4, WEE7 WEEA10, RMS
1.1.1	Random Forest	db12	94.2	Em1, Em3 Em4, Em6 Em8, WEE1 WEE2, WEE3 WEE4, WEE5 WEE6, WEE7 WEE8, RMS
1.1.2	Random Forest	db12	76.2	Em1, Em2 Em3, Em4 Em5, Em7 Em8, WEE1 WEE2, WEE3 WEE5, WEE6 RMS
1.2.1	Random Forest	db8	96.5	Em1, Em2, Em4 Em5, Em6, Em7 Em8, Em9, Em10 EmA10, WEE1, WEE2 WEE3, WEE5, WEE6 WEE8, WEE9, RMS
1.2.2	Random Forest	db8	89.5	Em6, WEE4 RMS

Tabela 4.9: Combinações com melhor desempenho em acurácia.

Assim, de posse das escolhas das combinações dos algoritmos, momento de desvanecimento e vetores de atributos otimizados, observa-se que o sistema classificador de VFOTs proposto, contribui para a automação do monitoramento da subestação e rapidez em diagnósticos de ativos da rede.

4.4 Trabalhos correlatos

De modo a contextualizar o presente trabalho com os principais artigos da área de classificação de distúrbios no setor elétrico, a Tabela 4.10 faz uma comparação qualitativa de trabalhos correlatos datados do ano de 2001 até o mais atual publicado em 2021.

Referencia	Base de dados	Janela de tempo	Frequência de amostragem	Data Augmentation?	Classificação (detalhes)	Algoritmos testados	Desempenho geral
Trabalho atual (2022)	Real	0.2ms	20MHz	sim	DS vs CB (Identificação, Localização e manobra)	RF MVS RBF PMC	94%
Asman, et al. [32] (2021)	Simulado	350ms	11.28kHz	Não	Distúrbios vs Não-distúrbios (Afundamento, elevação, transitórios vs Condições normais)	SVM	93.8%
Khokhar, et al. [30] (2016)	Simulado	200ms	10kHz	Não	Distúrbios vs Não-distúrbios (16 tipos de distúrbios vs condições normais)	SVM RBF MLP PNN	99%
Perera, et al. [33] (2011)	Simulado	15ms	10kHz	Não	Falta vs Não-falta (identificação da fase e da falta vs carregamento e manobra de banco de capacitor)	DT HMM PNN	95%
Zhengyou, et al. [9] (2011)	Simulado	250ms	100kHz	Não	manobra de disjuntor manobra de capacitores falta de curto na LT arco primário descarga atmosférica perto da LT descarga atmosférica na LT	MLP	95%
Benemar, et al. [4]-[10] (2010)	Simulado e Real	16.66ms	15.36kHz	Não	Transitório vs Não-Transitório (Faltas, afundamento, manobras vs condição normal)	árvore de decisão	99%
Hamzah, et al. [20] (2009)	Simulado	0.75 seg	não mencionado	Não	Classificação de distúrbio (manobra de banco de capacitor vs descarga elétrica)	MVS	80%
Hong, et al. [31] (2005)	Simulado	16.66ms	30.72kHz	Não	manobra de carga vs manobra de banco de capacitor	SOMN	93%
Mao, et al. [8] (2001)	Simulado	5ms	25kHz	Não	Falta interna vs Corrente de magnetização do transformador	PMC	99%

Tabela 4.10: Comparação qualitativa de trabalhos da área.

É interessante observar que este trabalho foi o único a abordar a técnica de *Data augmentation* aplicada diretamente em sinais VFTOs. Além disso, destaca-se que, dentro dos trabalhos citados, foi o que amostrou os distúrbios transitórios com maior frequência de amostragem e menor janela de tempo para sua análise e classificação pelos algoritmos de aprendizado de máquina e único a classificar chave seccionadora *versus* disjuntor, mostrando-se relevante e atual com o estado da arte.

Capítulo 5

Conclusão

O trabalho apresentado propôs uma metodologia para a classificação de VFTOs gerados por chaves seccionadores e disjuntores que incidem em ativos no interior de uma subestação elétrica. O sistema autônomo de monitoramento é pautado pela aquisição da forma de onda transitória e pelo uso de algoritmos de ML que são capazes de caracterizar o evento. Além disso, foi descrito o procedimento de janelamento aplicado diretamente na forma de onda do sinal VFTO visando a criação de novas amostras e então, o balanceamento da base de dados para um melhor desempenho do algoritmo de aprendizado de máquina e também a utilização de métricas de avaliação menos complexas e análises mais objetivas, como por exemplo, a métrica de acurácia clássica (A_{cc}). Ainda sobre a descrição da metodologia, foram apresentadas rotinas de extração de características das formas de onda dos sinais utilizando a TDW e otimizações do vetor de atributos com as técnicas de seleção do tipo filtragem(*filter*) e enrolamento(*wrapper*).

Os resultados mostraram que diversos algoritmos poderiam ser implementados junto ao sistema de monitoramento, porém, o algoritmo de ML que apresentou melhor desempenho em todos os cenários foi o *Random Forest*. Ainda, os resultados mostraram que para cada tarefa classificadora, um momento de desvanecimento da Wavelet apresentou melhor desempenho de acurácia. Entretanto, em um sistema embarcado *in loco*, cujo processamento computacional seja lento, a escolha do algoritmo de ML e de seu vetor de atributos torna-se relevante a medida em que o monitoramento se torna crítico, demandando rapidez na classificação do sinal VFTO adquirido. Além disso, outra limitação a ser observada, é a co-dependência da etapa de extração de atributos, onde as entropias das energias de wavelet(WEE) são dependentes e derivadas da extração das energias dos coeficiente de wavelet (Em). Apesar disso, a redução do número de entradas continua sendo uma otimização válida, pois um algoritmo que necessita de menos entradas tende

a ter um tempo de resposta mais curto [38].

Desta forma, esse trabalho procurou preencher a lacuna da falta de uma metodologia de classificação de eventos VFTOs de alta frequência gerados por chaves seccionadores e disjuntores no interior de subestações. Pesquisas na área de *Power Quality* (PQ) que investigam a classificação de distúrbios transitórios pela forma de onda de tensão/corrente foram comparados situando este trabalho com outros artigos atuais.

Como etapas futuras trabalho propõe-se a implementação do sistema proposto embarcado em uma subestação do setor elétrico em parceria com o Laboratório de Diagnóstico em Equipamentos e Instalações Elétricas(LABDIG) do CEPEL-Eletrbras. Assim, o sistema irá monitorar e auxiliar no diagnóstico de falhas potencialmente causadas por transitórios do tipo VFTOs. O monitoramento e caracterização do evento, pode ser utilizada para a geração de *logs* de erro e/ou relatórios sobre o distúrbio. Esses podendo reportar: valor de pico do evento, assinatura espectral, informação quanto ao tipo de equipamento, sua identificação na topologia interna da rede, manobra realizada. As informações geradas podem ser auxiliares para diagnósticos ágeis e eficazes nos casos de falhas ocasionadas pelos VFTOs. Por fim, uma outra proposta seria a utilização de amostrador com frequência de amostragem superior a 20MHz e a utilização da metodologia descrita para classificação de chaves seccionadoras em bom estado *versus* degradada pelo VFTO gerado, visando uma manutenção preditiva da rede.

Referências

- [1] PAULILLO MATEUS DUARTE TEIXEIRA, I. B. G. Cap xi: Transitórios. O setor elétrico, p. p.42–48, 2013.
- [2] PAULILLO, G. Cap i: Conceitos gerais sobre qualidade da energia. O setor elétrico, p. p.28–35, 2013.
- [3] D’AJUZ FONSECA, F. A. F. D. A. Transitórios elétricos e coordenação de isolamento – aplicação em sistemas de potencia de alta tensão. In: . [S.l.]: FURNAS Centrais Elétricas e Universidade Federal Fluminense, 1987.
- [4] IEEE Recommended Practice for Monitoring Electric Power Quality. In: . [S.l.: s.n.], 2019. p. 1–98.
- [5] PATHAK, N.; BHATTI, T. S.; IBRAHEEM. Study of very fast transient overvoltages and mitigation techniques of a gas insulated substation. In: *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]*. [S.l.: s.n.], 2015. p. 1–6.
- [6] ELÉTRICO, O. N. do S. Diretrizes para a elaboração de projetos básicos para empreendimentos de transmissão. In: *Estudos Elétricos, Especificação das Instalações, de Equipamentos e de Linhas de Transmissão*. [S.l.: s.n.], 2013.
- [7] MA, G.-m.; LI, C.-r.; QUAN, J.-t.; TANG, H.-x.; JIANG, J.; YIN, Y.-h.; ZHAO, H.-g.; CHAO, H. Measurement of very fast transient overvoltages at the transformer entrance. In: *2008 International Conference on Electrical Machines and Systems*. [S.l.: s.n.], 2008. p. 788–792.
- [8] PROBERT, S.; SONG, Y. Detection and classification of high frequency transients using wavelet analysis. In: *IEEE Power Engineering Society Summer Meeting*. [S.l.: s.n.], 2002. v. 2, p. 801–806 vol.2.
- [9] MAO, P.; AGGARWAL, R. A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network. *IEEE Transactions on Power Delivery*, v. 16, n. 4, p. 654–660, 2001.
- [10] ZHENGYOU, H.; SHIBIN, G.; XIAOQIN, C.; JUN, Z.; ZHIQIAN, B.; QINGQUAN, Q. Study of a new method for power system transients classification based on wavelet entropy and neural network. *International Journal of Electrical Power Energy Systems - INT J ELEC POWER ENERG SYST*, v. 33, p. 402–410, 03 2011.
- [11] SOUZA, B.; BRITO, N.; COSTA, F.; LEITAO, J.; CAUPONI, S.; SILVA, S. Evolution of oscillograph - how did chesf transform data into information? *Latin America Transactions, IEEE (Revista IEEE America Latina)*, v. 8, p. 45 – 50, 04 2010.

- [12] ABDULLAH, M.; SAID, A.; BADRAN, E. High-frequency spectrum analysis of vfto generated inside gas insulated substations. *World Academy of Science, Engineering and Technology, ISSN: 2010376X, 20103778*, v. 79, 06 2013.
- [13] VASQUES RÔMULO DOS SANTOS DELGADO, L. E. D. S. D. S. D. S. I. F. D. N. G. S. R. M. T. F. D. S. C. M. R. Sistema de mediÇÃo para solicitaÇÕes transitórias em transformadores de corrente para alta tensÃo. *XXV SNPTEE SEMINÁRIO NACIONAL DE PRODUÇÃO E TRANSMISSÃO DE ENERGIA ELÉTRICA*), 10 2019.
- [14] TAVAKOLI, M.; NAFAR, M. Human reliability analysis in maintenance team of power transmission system protection. In: . [S.l.: s.n.], 2020. v. 5.
- [15] PINTO, J. S. d. S. F. A.; ALVES, M. E. G. Application of On-line Monitoring Systems in the view of Maintenance Engineering of Eletrobras Furnas. In: . [S.l.: s.n.], 2015.
- [16] FILHO, J. M. Instalações Elétricas Industriais. In: *SmartGridComm*. [S.l.]: Instalações Elétricas Industriais, 2017. p. 422–427.
- [17] PAULINO, M. Cap i: Princípios básicos de transformadores de potência. In: . [S.l.]: O setor elétrico, 2014. p. p.52–61.
- [18] PUENTE A. FUCHS, J. A. F. e. a. A. P. Disjuntores e chaves. In: *Aplicação em sistemas de potência*. [S.l.]: Eduff, Niteroi-RJ, 1995.
- [19] VERY fast transients in GIS. *IEEE Transactions on Power Delivery*, v. 4, n. 1, p. 223–233, 1989.
- [20] XIN, W.; ZHANG, G.; WANG, J.; LIU, K.; GENG, Y. Simulation of very fast transient overvoltage caused by disconnecter switch operation in sf6 gas insulated switchgear using multi-restrike arcing model. In: . [S.l.: s.n.], 2013. p. 1–4. ISBN 978-1-4799-3817-9.
- [21] POVH, D.; SCHMITT, H.; VOLCKER, O.; WITZMANN, R.; CHEWDHURI, P.; IMECE, A.; IRAVANI, R.; MARTINEZ, J.; KERI, A.; SARSHAR, A. Modeling and analysis guidelines for very fast transients. *Power Engineering Review, IEEE*, v. 17, p. 71 – 71, 11 1996.
- [22] HAMZAH, N.; ANUWAR, F.; ZAKARIA, Z.; TAHIR, N. Classification of transient in power system using support vector machine. In: . [S.l.: s.n.], 2009.
- [23] QIU, J.; WU, Q.; DING, G.; XU, Y.; FENG, S. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, v. 2016, 05 2016.
- [24] L'HEUREUX, A.; GROLINGER, K.; YAMANY, H. E.; CAPRETZ, M. Machine learning with big data: Challenges and approaches. *IEEE Access*, PP, p. 1–1, 04 2017.
- [25] WEN LIANG SUN, F. Y. X. S. J. G. X. W. H. X. Q. Time series data augmentation for deep learning: A survey. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 2021.

- [26] KLEINFELD, D.; MITRA, P. P. Spectral methods for functional brain imaging. *Cold Spring Harbor Protocols*, Cold Spring Harbor Laboratory Press, v. 2014, n. 3, p. pdb-top081075, 2014.
- [27] RAWAT, T.; KHEMCHANDANI, V. Feature engineering (fe) tools and techniques for better classification performance. 05 2019.
- [28] GRAPS, A. "an introduction to wavelets". *IEEE Comp. Sci. Engi.*, v. 2, p. 50–61, 02 1995.
- [29] WEEKS, M. *Digital Signal Processing Using MATLAB amp; Wavelets, Second Edition*. 2nd. ed. USA: Jones and Bartlett Publishers, Inc., 2010. ISBN 0763784222.
- [30] OLKKONEN, H. Discrete wavelet transforms. In: *Algorithms and Applications*. [S.l.]: IN-TECH, 2011.
- [31] PINHO EMANOEL G, A. G. A. C. Novas métricas para a classificação de distúrbios da qualidade da energia elétrica. 02 2021.
- [32] AHUJA, N.; LERTRATTANAPANICH, S.; BOSE, N. Properties determining choice of mother wavelet. *iee proc vis image signal process. Vision, Image and Signal Processing, IEE Proceedings -*, v. 152, p. 659 – 664, 11 2005.
- [33] JAYASREE, T.; DEVARAJ, D.; SUKANESH, R. Classification of transients using wavelet based entropy and radial basis neural networks. *International Journal of Computer and Electrical Engineering*, p. 590–595, 2009.
- [34] DANASINGH, A. A.; BALAMURUGAN, S.; EPIPHANY, J. L. Literature review on feature selection methods for high-dimensional data. *International Journal of Computer Applications*, v. 136, 02 2016.
- [35] U.R., A.; PAUL, S. Feature selection and extraction in data mining. In: . [S.l.: s.n.], 2016. p. 1–3.
- [36] CAI, J.; LUO, J.; WANG, S.; YANG, S. Feature selection in machine learning: A new perspective. *Neurocomputing*, v. 300, p. 70–79, 2018. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231218302911>>.
- [37] CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. *Computers Electrical Engineering*, v. 40, n. 1, p. 16–28, 2014. ISSN 0045-7906. 40th-year commemorative issue. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045790613003066>>.
- [38] SUTO, J.; ONIGA, S.; SITAR, P. P. Comparison of wrapper and filter feature selection algorithms on human activity recognition. In: *2016 6th International Conference on Computers Communications and Control (ICCC)*. [S.l.: s.n.], 2016. p. 124–129.
- [39] HORNIK, K.; BUCHTA, C.; ZEILEIS, A. Open-source machine learning: R meets Weka. *Computational Statistics*, v. 24, n. 2, p. 225–232, 2009.
- [40] MUHAMEDYEV, R. Machine learning methods: An overview. *CMNT*, v. 19, p. 14–29, 01 2015.

- [41] NASTESKI, V. An overview of the supervised machine learning methods. *HORIZONS.B*, v. 4, p. 51–62, 12 2017.
- [42] MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, v. 52, p. 99–115, 1990.
- [43] HAYKIN, S. *Redes Neurais: Princípios e Prática*. Artmed, 2007. ISBN 9788577800865. Disponível em: <<https://books.google.com.br/books?id=bhMwDwAAQBAJ>>.
- [44] JUNG Y., . H. J. A k-fold averaging cross-validation procedure. *journal of nonparametric statistics*. p. 167–179, 2015.
- [45] RASCHKA, S. *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*. 11 2018.
- [46] JAPKOWICZ, N. Imbalanced learning: Foundations, algorithms, and applications. In: *Assessment Metrics for Imbalanced Learning*. [S.l.]: Wiley, 2013.
- [47] LUQUE, A.; CARRASCO, A.; MARTÍN, A.; de las Heras, A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, v. 91, p. 216–231, 2019. ISSN 0031-3203. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320319300950>>.
- [48] CHOUDHARY, R.; GIANEY, H. K. Comprehensive review on supervised machine learning algorithms. In: *2017 International Conference on Machine Learning and Data Science (MLDS)*. [S.l.: s.n.], 2017. p. 37–43.
- [49] VIMALRAJ, S.; DR.R, P. A review on handling imbalanced data. 12 2018.
- [50] KOTSIANTIS, S.; KANELLOPOULOS, D.; PINTELAS, P. Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, v. 30, p. 25–36, 11 2005.
- [51] MOHAMMED, R.; RAWASHDEH, J.; ABDULLAH, M. Machine learning with over-sampling and undersampling techniques: Overview study and experimental results. In: . [S.l.: s.n.], 2020. p. 243–248.