

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Guilherme Luiz da Cunha

**Uma Abordagem Baseada em Árvores
de Decisão para a Análise da
Estabilidade Angular do Rotor**

Guilherme Luiz da Cunha

**Uma Abordagem Baseada em Árvores
de Decisão para a Análise da
Estabilidade Angular do Rotor**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Sistemas Elétricos e Eletrônicos

Orientador: Tatiane Cristina da Costa Fernandes

Coorientador: Ricardo Augusto Souza Fernandes

São Carlos - SP

2021



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Engenharia Elétrica

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Guilherme Luiz da Cunha, realizada em 06/12/2021.

Comissão Julgadora:

Profa. Dra. Tatiane Cristina da Costa Fernandes (UFSCar)

Prof. Dr. Edson Luis Geraldi Junior (IFSP)

Prof. Dr. Roman Kuiava (UFPR)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Engenharia Elétrica.

Agradecimentos

Agradeço à minha família, especialmente, aos meus avós Clarice e Luiz que sempre me incentivaram e fizeram muito com pouco;

Agradeço aos meus orientadores Tatiane Cristina da Costa Fernandes e Ricardo Augusto Souza Fernandes por terem contribuído muito com meu crescimento, pela confiança, oportunidade e paciência durante todo o mestrado;

Aos meus amigos, principalmente, ao Marcão e a Fernanda, pelo apoio e conselhos nos últimos anos.

Resumo

A análise de segurança do sistema elétrico de potência é fundamental para que se possa garantir o fornecimento de energia elétrica e a viabilidade da operação. Dentre estas análises, o estudo da estabilidade angular do rotor visa assegurar que os modos eletromecânicos do sistema estão bem amortecidos e, para tanto, é comum que os métodos empregados façam uso de modelos matemáticos do sistema elétrico de potência. Entretanto, a dependência dos modelos matemáticos, potencialmente, torna os métodos de avaliação sensíveis às variações dos parâmetros e mudanças na rede que podem não ser adequadamente representadas, principalmente, em aplicações em tempo real. Atualmente, as *Smart Grids* se propõem a oferecer maior capacidade de monitoramento e taxas de amostragem que viabilizam análises em tempo real. Concomitantemente, os avanços da computação distribuída e em nuvem tem fomentado o emprego de técnicas de aprendizagem de máquina para a resolução de diversos problemas a partir do uso da quantidade massiva de dados disponíveis. Nesse sentido, esse trabalho se propõe a fazer uso das medições disponibilizadas por unidades de medição fasorial para criar um *framework* composto por árvores de decisão, que avalie a estabilidade angular do rotor. Para análise da estabilidade a pequenas perturbações, uma abordagem descentralizada baseada em árvores de decisão individuais e dados advindos de unidades de medição fasorial alocadas nas barras de geração é aplicada. Nessa abordagem, cada árvore de decisão utiliza apenas medidas locais para a avaliação da estabilidade angular do rotor a pequenas perturbações, de modo que a classificação pode ser realizada mesmo quando há perda de informação de geradores específicos ou falha no sistema de comunicação. Quando o sistema é submetido a uma grande perturbação, um segundo método é utilizado, o qual se baseia em uma árvore centralizada e fasores de tensão mensurados nas barras de geração de todo o sistema em estudo. Esse último classificador é capaz de identificar instabilidade na porção pós-falta da resposta do sistema, estabelecendo um *tradeoff* entre o número de ciclos de medições utilizado e o desempenho do classificador. Os resultados obtidos no sistema IEEE 68 barras mostram a eficácia da abordagem proposta. Na classificação da estabilidade a

pequenas perturbação, uma acurácia de 93% é alcançada pela árvores distribuídas mesmo em cenários de contingência e variações de carga. Já com relação a grandes perturbações, foi possível classificar com precisão a estabilidade transitória mesmo com apenas 1 ciclo de medição (96,4%) e com um apenas 3 pontos de medição ao longo do sistema teste.

Palavras-chave: Estabilidade a pequenas perturbações. Estabilidade transitória. Unidades de medição fasorial. Aprendizagem de máquina. Aplicações em tempo real.

Abstract

The power system security assessment is essential to ensure the supply of electrical energy and the feasibility of the operation. Among these analyses, the study of the rotor angle stability aims to ensure that the electromechanical modes of the system are well damped and, for that, it is common that the methods employed make use of mathematical models of the electric power system. However, the dependence on mathematical models, potentially, makes the evaluation methods sensitive to variations in parameters and changes in the network that have not been adequately represented, mainly in real-time applications. Currently, Smart Grids proposes to offer greater monitoring capacity and sample rates that allow real-time analysis. Concomitantly, the advances in distributed and cloud computing have encouraged the use of machine learning techniques to solve various problems using the massive amount of data available. In this sense, this work proposes to make use of the measurements made available by phasor measurement units to evaluate the feasibility of using decision trees in the analysis of the rotor angle stability. For stability analysis at small disturbances, a decentralized approach based on individual decision trees and data from phasor measurement units allocated in the generator buses is applied. In this approach, each decision tree uses only local measurements to evaluate the rotor angle small-signal stability, in this way, the classification can be carried out even when there is loss of information from specific generators or failure in the communication system. When the system is subjected to a large disturbance, a second method is employed, which is based on a centralized decision tree and voltage phasors measured at generator buses from the whole studied system. This last classifier is able to identify instability in the response post fault portion of the system, establishing a trade-off between the number of measurement cycles used and the classifier performance. The results obtained on the IEEE 68-buses system showed the efficiency of the proposed approach. In the classification of small-signal stability, an accuracy of 93% is reached by the distributed trees even in scenarios with contingencies and load variations. About large disturbances was possible to classify with precision the transient stability even with

only 1 measurement cycle (96.4%) and with only 3 measurement points along the test system.

Keywords: Small signal stability. Transient stability. Phasor measurement units. Machine learning. Real-time applications.

Lista de ilustrações

Figura 1 – Classificação de estabilidade em SEP	30
Figura 2 – Resposta do ângulo do rotor diante de distúrbio transitório.	33
Figura 3 – Gráfico polar <i>mode shape</i>	40
Figura 4 – Abordagem para a análise de estabilidade a pequenas perturbações . .	54
Figura 5 – Abordagem para a análise de estabilidade transitória	54
Figura 6 – Árvore de decisão	56
Figura 7 – Abordagem descentralizada proposta para análise da estabilidade a pe- quenas perturbações	58
Figura 8 – Panorama do método proposto para pequenas perturbações.	59
Figura 9 – Abordagem centralizada proposta para análise da estabilidade transitória	64
Figura 10 – Panorama do método proposto para estabilidade transitória.	65
Figura 11 – Ciclos de medição do pós falta	67
Figura 12 – Ângulos dos rotores no pós-falta	68
Figura 13 – Fluxograma de seleção de features	69
Figura 14 – Método K-folds	71
Figura 15 – 68 barras NETS-NYPS	74
Figura 16 – Variações de curva diária	76
Figura 17 – Distribuição de casos - Cenário 1	76
Figura 18 – Distribuição de casos - Cenário 2	77
Figura 19 – Distribuição de casos - cenário 3	78
Figura 20 – Distribuição de casos - Cenário 4	78
Figura 21 – Árvore de decisão da barra 15.	79
Figura 22 – Acurácia em diferentes cenários	83
Figura 23 – Faltas trifásicas	87
Figura 24 – Distribuição de casos x tempo de eliminação	87
Figura 25 – Teste F - Tempo de eliminação de 1,5 segundos	89
Figura 26 – Desempenho da DT para diferentes profundidades	89

Figura 27 – Desempenho de DT e SVM para diferentes quantidades de ciclos de medição	91
Figura 28 – Desempenho das DTs distribuídas para diferentes quantidades de ciclos de medição	92
Figura 29 – Desempenho da DT 16 para diferentes quantidades de ciclos de medição	93
Figura 30 – Desempenho da DT centralizada para o ciclo 1 em função da quantidade de <i>features</i>	94
Figura 31 – Desempenho da DT centralizada para o ciclo 2 em função da quantidade de <i>features</i>	95
Figura 32 – Desempenho da DT centralizada para o ciclo 3 em função da quantidade de <i>features</i>	95
Figura 33 – Desempenho da DT centralizada para o ciclo 4 em função da quantidade de <i>features</i>	96
Figura 34 – Desempenho da DT centralizada para o ciclo 5 em função da quantidade de <i>features</i>	97
Figura 35 – Desempenho da DT centralizada para o ciclo 6 em função da quantidade de <i>features</i>	97
Figura 36 – Desempenho da DT centralizada para o ciclo 7 em função da quantidade de <i>features</i>	98
Figura 37 – Desempenho da DT centralizada para o ciclo 8 em função da quantidade de <i>features</i>	98

Lista de tabelas

Tabela 1 – Matriz de Confusão	63
Tabela 2 – Matriz de frequência de <i>features</i> mais importantes por ciclo	69
Tabela 3 – Modos eletromecânicos do sistema teste.	74
Tabela 4 – Performance no primeiro cenário	80
Tabela 5 – Performance no segundo cenário.	81
Tabela 6 – Performance no terceiro cenário.	82
Tabela 7 – Performance no quarto cenário.	82
Tabela 8 – Importância de <i>features</i> por ciclo - Seleção por <i>Random forest</i>	93

Lista de siglas

ANN Artificial Neural Network

BCU Boundary Controlling Unstable Equilibrium Point

CEPEL Centro de Pesquisas de Energia Elétrica

DSA Dynamic Security Assessment

DT Decision Tree

IRF Iterative Random Forest

LSMYW Least Square Modified Yule Walker

OBP Operating Base Points

PCA Principal Component Analysis

PCC Pearson Correlation Coefficient

PDC Phasor Data Concentrator

PEBS Potential Energy Boundary Surface

PMI Partial Mutual Information

PMU Phasor Measurement Unit

SCADA Supervisory Control and Data Acquisition

SEP Sistema Elétrico de Potência

SI Subspace Identification

SSA Small-signal Security Assessment

SVM Support Vector Machine

TKE Total Kinetic Energy

TPE Transient Potential Energy

TSA Transient Security Assessment

TSI Transient Instability Index

TSO Transmission System Operator

Sumário

1	INTRODUÇÃO	21
1.1	Técnicas de aprendizagem de máquina para estudo da estabilidade angular do rotor	23
1.2	Objetivos	26
1.3	Organização dos capítulos	26
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Introdução ao conceito de Estabilidade e as suas categorias	29
2.2	Estabilidade transitória	32
2.2.1	Ferramentas para estudo da estabilidade transitória	34
2.3	Estabilidade a pequenas perturbações	36
2.3.1	Representação em espaço de estados	37
2.3.2	Resposta no tempo do sistema linearizado	39
3	REVISÃO BIBLIOGRÁFICA	43
3.1	Aprendizagem de máquina na análise de segurança	44
3.2	Aprendizagem de máquina na análise de estabilidade angular	47
3.3	Análise Geral	50
4	METODOLOGIA	53
4.1	Árvores de decisão	55
4.1.1	ID3	57
4.1.2	C4.5	57
4.1.3	CART	57
4.2	Método 1: Classificação da estabilidade angular do rotor a pequenas perturbações	58
4.2.1	Formação do banco de dados e Monitoramento do Modos Eletromecânicos	60

4.3	Treinamento e testes com árvores de decisão	62
4.3.1	Métricas para análise de desempenho	63
4.4	Método 2: Classificação da Estabilidade Angular do Rotor a Grandes Perturbações	64
4.4.1	Formação do Banco de Dados	65
4.4.2	Seleção de <i>features</i>	67
4.4.3	Treinamento e Validação da Árvore de decisão	69
4.4.4	Relação entre desempenho e a quantidade de ciclos de medição	71
4.4.5	Classificadores distribuídos	72
4.4.6	Identificação do conjunto mínimo de medidores necessários	72
5	RESULTADOS - CLASSIFICAÇÃO DE ESTABILIDADE AN- GULAR DO ROTOR A PEQUENAS PERTURBAÇÕES A PARTIR DE ÁRVORES DE DECISÃO DISTRIBUÍDAS . . .	73
5.1	Sistema Teste	73
5.2	Banco de dados do sistema teste de 68 Barras	74
5.2.1	Cenário 1	75
5.2.2	Cenário 2	75
5.2.3	Cenário 3	77
5.2.4	Cenário 4	78
5.3	Desempenho dos classificadores	79
5.3.1	Cenário 1 - Variações diária de carga e vetor de crescimento unidirecional	79
5.3.2	Cenário 2 - Variações de carga, vetor de crescimento unidirecional e contingências	80
5.3.3	Cenário 3 - Variações de carga e vetor de crescimento multidirecional . .	80
5.3.4	Cenário 4 - Variações de carga, vetor de crescimento multidirecional e contingências	81
5.4	Discussões	82
6	RESULTADOS - CLASSIFICAÇÃO DE ESTABILIDADE AN- GULAR DO ROTOR TRANSITÓRIA A PARTIR DA ÁR- VORE DE DECISÃO CENTRALIZADA	85
6.1	Composição do Banco de Dados	86
6.2	Critério de simulação dos testes propostos	88
6.2.1	Seleção de <i>Features</i> - tipo de Variável	88
6.2.2	Escolha da profundidade	88
6.3	Resultados	89
6.3.1	Robustez à redução de ciclos de medição	90
6.3.2	Abordagem distribuída	91
6.3.3	Seleção de <i>features</i> combinadas - Avaliação das barras de medição . . .	92

7	CONCLUSÕES	99
	Conclusão	103
	REFERÊNCIAS	103

Capítulo 1

Introdução

No planejamento e operação dos sistemas elétricos de potência (SEPs), a análise e predição da segurança é essencial para assegurar o fornecimento ininterrupto de energia aos consumidores e garantir a confiabilidade do sistema (MUKHERJEE; DE, 2020; MASSERAN; GOMES; AMARAL, 2017). Uma operação segura está relacionada com a capacidade do sistema em operar dentro de limites aceitáveis, ou seja, em um ponto de operação viável no espaço de estados e, da habilidade de diante de uma perturbação retomar o regime permanente, sem exceder esses mesmos limites (Balu et al., 1992).

Segundo Morison, Lei Wang e Kundur (2004), para garantir que o sistema opere de forma suficientemente confiável, de modo que os riscos de falhas no suprimento de energia elétrica aos consumidores a longo prazo sejam minimizados, dois principais aspectos são importantes: 1) O planejamento da operação do sistema de potência e da sua expansão devem ser desenvolvidos tomando como um dos aspectos primordiais a análise de segurança e 2) o sistema deve ser continuamente monitorado para assegurar que a sua operação ocorra dentro de uma margem suficiente de segurança, uma vez que os pontos de operação da rede bem como as contingências podem ser diferentes daquelas planejadas.

Dentre as avaliações de segurança essenciais ao sistema, encontra-se a Avaliação de Segurança Dinâmica (em inglês, *Dynamic Security Assessment* - DSA) que é composta pela Avaliação de Segurança Transitória (em inglês, *Transient Security Assessment* - TSA) e Avaliação de Segurança a Pequenos Sinais (em inglês, *Small-signal Security Assessment* - SSA). Entre os aspectos analisados pela TSA está avaliação da estabilidade angular transitória, onde avalia-se o sincronismo das máquinas síncronas do sistema após uma grande perturbação, como a perda de uma linha. A SSA, por sua vez, avalia a estabilidade a pequenas perturbações analisando, principalmente, a frequência e o amortecimento dos modos do sistema diante de pequenos distúrbios, como as variações de carga diária

(KUNDUR, 1994a; MASSERAN; GOMES; AMARAL, 2017).

Nas *smart grids*, em que há maior interação de tecnologias distintas tais como fontes alternativas intermitentes e cargas a base de eletrônica de potência, além de diferentes perfis de consumo, garantir a operação segura de rede elétrica interconectada têm se tornado cada vez mais desafiador. Neste cenário, o monitoramento da rede para atendimento da demanda e gestão adequada dos recursos energéticos é de extrema importância. Diante das informações desse monitoramento, o operador poderá antever riscos e acionar as devidas medidas de segurança, permitindo que a energia elétrica seja fornecida continuamente e sem interrupções (Babakmehr et al., 2016).

Apesar das complexidades adicionais, as *smart grids* oferecem um ambiente extremamente rico em medições, apresentando um alto potencial para monitoramento em tempo real e otimização (MCNABB; WILSON; BIALEK, 2013). As medições coletadas por Unidades de Medição Fasorial (em inglês, *Phasor Measurement Unit* - PMUs) permitem o uso de análise de dados para obter as previsões relacionadas ao problema de estabilidade angular do rotor e podem auxiliar nas tomadas de decisão, por parte do operador do sistema. Nesse contexto, o uso de técnicas de aprendizagem de máquina na avaliação e previsão da estabilidade angular se apresenta como uma possibilidade bastante promissora.

O uso da PMU possibilita a medição de diversos pontos da rede com alta taxa de amostragem, normalmente, entre 30 e 60 amostras por segundo. Os dados coletados pela PMU são encaminhados ao concentrador de dados fasoriais (em inglês, *Phasor Data Concentrator* - PDC) e, posteriormente, enviados ao operador do sistema de transmissão. Dessa forma, o operador tem disponível um panorama de toda a rede e com o auxílio das ferramentas adequadas, pode atuar para garantir a operação segura da mesma. É importante, portanto, que se faça uso do ambiente rico em dados que a inserção das PMUs proporciona, para não se subutilizar o potencial dos PDCs (WERBESTON, 2017).

Na análise de estabilidade angular do rotor, independente da excursão da perturbação, tem-se como objetivos definir se o sistema é capaz de alcançar um novo ponto ou permanecer no mesmo estado de equilíbrio no cenário pós-distúrbio e, findada essa verificação, investigar se o amortecimento dos modos eletromecânicos se encontra dentro do limite estipulado para um sistema de transmissão, que é acima de 5% (SENE SOULIN; HONGESOMBUT; DECHANUPAPRITTHA, 2019). Se as oscilações eletromecânicas não forem devidamente amortecidas, elas podem causar instabilidade e potencialmente levar o sistema a um *blackout*.

Para mitigar os efeitos da presença de modos eletromecânicos mal amortecidos, em tempo real, é necessário que se monitore as principais características dessas oscilações como fator de amortecimento, frequência e amplitude. A partir desse monitoramento, potenciais cenários de instabilidade do sistema podem ser prontamente identificadas, possibilitando o envio de um alerta prévio aos operadores para que os mesmos possam tomar as medidas necessárias, sejam preventivas ou corretivas, para assegurar que o sistema

opere dentro de sua margem segura (GHAHREMANI; KAMWA, 2011).

Tradicionalmente, no contexto de pequenas perturbações, para analisar as oscilações eletromecânicas utiliza-se uma abordagem analítica baseada na linearização das equações algébrico-diferenciais que representam o modelo conceitual do SEP. Pela análise dos autovalores da matriz de estados do modelo linearizado, as causas e efeitos das oscilações mal amortecidas são investigados (KUNDUR et al., 2004).

Na análise de grandes perturbações, como a linearização deixa de ser válida, é comum que se faça uso de simulações no domínio do tempo para resolução do modelo matemático não linear que representa o sistema elétrico (KUNDUR, 1994a). Para esta análise, os cenários com maior probabilidade de ocorrência são avaliados por meio de simulações, normalmente, a partir da integração numérica do modelo não linear, que permitem um adequado planejamento da operação da rede. No entanto, as simulações no domínio do tempo podem demandar um alto custo computacional, restringindo a sua aplicação para tomada de decisão em tempo real.

Além disso, métodos que dependem diretamente do modelo matemático do sistema, tornam-se sensíveis aos erros de modelagem e incertezas nos parâmetros, o que de acordo com Zhang et al. (2016) pode restringir sua capacidade de representar o comportamento dinâmico do SEP. Dessa forma, as incertezas referentes ao estado da topologia da rede, ao carregamento vigente e os estados dos diversos elementos que compõe o sistema se fazem necessários na modelagem, dificultando a análise em tempo real da estabilidade angular do rotor via modelo matemático (LEANDRO et al., 2015). Em suma, a necessidade de atualização manual dos modelos matemáticos e o tempo de processamento excessivo tornam a simulação no domínio do tempo pouco viável para análises *online*.

Diante desses desafios, pesquisas para melhorar o ferramental disponível que garante a operação segura e facilitar a tomada de ações de controle corretivo, em tempo real, por parte do operador do sistema de transmissão (em inglês, *Transmission System Operator* - TSO) são fundamentais (MCNABB; WILSON; BIALEK, 2013).

No contexto apresentado, este trabalho se propõe a investigar o uso de técnicas de aprendizagem de máquina que dependam, apenas, de medições provenientes de PMUs para classificar a estabilidade angular do rotor. Em específico, serão empregadas árvores de decisão, que apresentam, entre outras vantagens, baixo custo computacional e fácil interpretabilidade para auxiliar o operador do sistema na análise de estabilidade em tempo real.

1.1 Técnicas de aprendizagem de máquina para estudo da estabilidade angular do rotor

Dadas as dificuldades encontradas nos métodos convencionais de análise de estabilidade angular do rotor em tempo real e considerando os avanços no uso da inteligência

artificial e na capacidade computacional disponíveis pelo processamento em paralelo e da nuvem, técnicas de aprendizagem de máquina tem sido amplamente utilizadas para esse estudo. Dentre estas, árvores de decisão (em inglês, *Decision Tree* - DT), máquinas de vetor de suporte (em inglês, *Support Vector Machine* - SVM) e as redes neurais artificiais (em inglês, *Artificial Neural Network* - ANN) se apresentam como propostas promissoras no desenvolvimento de ferramentas robustas que auxiliem na análise de estabilidade, em tempo real. Tais ferramentas para serem aplicadas são usualmente treinadas a partir de simulações *offline* que relacionam variáveis medidas nos SEPs com a classificação de estabilidade do sistema (MCNABB; WILSON; BIALEK, 2013; LIU et al., 2014). Alguns exemplos de trabalhos reportados na literatura que utilizam técnicas de aprendizagem de máquina para classificação da estabilidade do SEP a partir de dados medidos são mencionados abaixo.

Uma estrutura baseada em árvore de decisão é desenvolvida para avaliar a estabilidade transitória em Liu et al. (2014). Para o desenvolvimento dessa estrutura, primeiro um banco de dados é formado para treinamento da DT, a partir de simulações no modelo do sistema teste para diferentes cenários de operação (níveis distintos de carga e contingências). Em cada simulação, as variáveis de fluxo de carga, a classificação do sistema quanto à estabilidade são registradas. Esse conjunto de informação é adotado para o treinamento da DT. Além disso, neste trabalho, uma segunda DT também é estabelecida para definir as estratégias de controle preventivas mais adequadas mediante o ponto de operação da rede.

Já em Guo e Milanovi (2016), um método de dois estágios é desenvolvido. No primeiro estágio, a DT e SVM são aplicadas para classificar a estabilidade da rede. No segundo estágio uma rede neural é aplicada para identificar o padrão da resposta dinâmica dos geradores conectados no sistema em estudo, pós-perturbação. Para essa identificação, medidas da resposta do ângulo do rotor dos geradores conectados no sistema em estudo são necessárias. Na classificação, uma acurácia de 99,75% é alcançada para 60 ciclos de medição (1,002 segundos), enquanto que para 10 ciclos de medição (0,167 segundos) a acurácia é de 98,7%.

Recentemente, em An et al. (2020), os autores utilizam uma rede neural convolucional para classificar estabilidade transitória, a partir de medições de PMU e do sistema de informação geográfico (em inglês, *Geographic information system* - GIS). Os resultados obtidos mostram que é possível inferir sobre a topologia da rede elétrica a partir da metodologia proposta, obtendo alta acurácia nos sistemas teste IEEE 39 barras e IEEE 118 barras.

Dentre as técnicas de aprendizagem de máquina mais utilizadas na análise de estabilidade, a DT se distingue pela facilidade de interpretação dos seus resultados, estabelecendo limites de operações factíveis para variáveis conhecidas do sistema elétrico de potência (KRISHNAN et al., 2011; WERBESTON, 2017). Por exemplo, em (MCNABB; WIL-

SON; BIALEK, 2013), a árvore de decisão além de classificar os modos eletromecânicos nas categorias estável ou instável, também fornece informação sobre possíveis combinações de variáveis do fluxo de potência que resultarão em baixo amortecimento do modo eletromecânico. A simplicidade de interpretação e acurácia fornecida pelas DTs pode, portanto, facilitar o processo de tomada de decisão, em tempo real, pelo operador.

A avaliação da estabilidade transitória bem como da segurança dinâmica da rede é predita por um modelo a base de DTs de múltiplos caminhos (*multiway*) em WerbestonPHD2017. Diferentes topologias e contingências são usadas como uma variável categórica, responsável pela primeira divisão da árvore, dessa forma cada ramo especializa-se na classificação de uma topologia ou contingência única. Os testes são realizados com o IEEE 39 barras e com a porção Noroeste do sistema interligado Nacional (SIN), sendo que para o último são usados dados reais, obtendo alta acurácia, com exceção do treinamento e teste realizados com dados históricos, resultado este que o autor atribuiu ao desbalanceamento de classes no banco de dados.

Um *ensemble* de árvores de decisão, o qual corresponde ao uso em conjunto de diversas árvores em um esquema de votação, é utilizado para a classificação de segurança dinâmica em SEPs de grande porte em (MUKHERJEE; DE, 2020), novamente, envolvendo a classificação de estabilidade transitória. Testes são realizados com o IEEE 118 barras e o IEEE 300 barras, em que acurácias de 94,86% e 91,44% são obtidas, respectivamente. Os autores reduzem os vetores de entradas através de um processo de seleção de *feature*, sendo que os vetores finais são compostos por potência ativa de geração, fluxo de potência ativa, potência reativa de geração, fluxo de potência reativa e magnitudes de tensão do pré-contingência.

Neste contexto, este trabalho se propõe a utilizar o algoritmo da árvore de decisão para análise de estabilidade angular do rotor, a partir de medições disponíveis em campo e com foco na viabilidade de implementação. Uma estrutura composta por duas abordagens, a pequenas e grandes perturbações, é desenvolvida para que as características intrínsecas de cada categoria de estabilidade angular do rotor sejam contempladas.

Entre os trabalhos com árvore de decisão anteriormente citados, não se explora a possibilidade de se utilizar classificadores distribuídos, que dependem de medições individualizadas. Na proposta sugerida para a estabilidade a pequenas perturbações, pretende-se explorar a redundância que classificadores distribuídos pode oferecer, o que tornaria o sistema mais robusto a perdas de medições. Além disso, a simplicidade computacional desses classificadores permite que os mesmos possam ser embarcados nas unidades de medição fasorial, fazendo uso da estrutura das *smart grids*.

Complementar às árvores distribuídas, também é proposta a construção de uma árvore de decisão centralizada que utiliza as medições de todas as barras geradoras. Esse classificador é capaz de analisar a estabilidade a partir do sinal pós-falta e oferece um pequeno *trade-off* entre o número de ciclos de medição utilizado e a acurácia. Sendo

assim, o operador poderá optar por predições com poucos ciclos de medição, mas ainda assim com alta acurácia, o que lhe disponibilizará mais tempo para tomada de decisão.

1.2 Objetivos

Diante do exposto, esta pesquisa tem como objetivo principal o desenvolvimento de classificadores de estabilidade angular do rotor, a pequenas e a grandes perturbações, que utilizem apenas dados referentes às medições de PMUs para realizar a análise da estabilidade. Dessa forma, pretende-se auxiliar na tomada de decisão por parte do operador do sistema de transmissão, em tempo real.

Para a realização desse trabalho, os seguintes objetivos específicos devem ser atendidos:

- ❑ Produzir um banco de dados com diferentes cenários de operação para o estudo de estabilidade angular do rotor;
- ❑ Em específico a classificação da estabilidade a pequenas perturbações: avaliar o desempenho de classificação das árvores distribuídas, quando comparadas a uma árvore centralizada;
- ❑ Em específico a classificação da estabilidade a grandes perturbações: investigar o conjunto mínimo de medidas necessárias para que o classificador de estabilidade responda com precisão, em um curto intervalo de tempo;
- ❑ Avaliar o desempenho dos classificadores frente a diferentes cenários e identificar possíveis dificuldades e melhorias.

1.3 Organização dos capítulos

A dissertação está organizada em sete capítulos principais, os quais, com exceção da introdução, são definidos em sequência.

- ❑ **Capítulo 2:** Os conceitos fundamentais envolvendo estabilidade em sistemas elétricos de potência, com ênfase em estabilidade angular do rotor são apresentados nesse capítulo.
- ❑ **Capítulo 3:** Neste capítulo é apresentada uma revisão bibliográfica sobre o uso de algoritmos de aprendizagem de máquina na análise de segurança do SEP, não se restringindo apenas à estabilidade angular.
- ❑ **Capítulo 4:** A metodologia utilizada no treinamento e avaliação dos algoritmos propostos para a análise de estabilidade angular, é detalhada nesse capítulo.

- ❑ **Capítulo 5:** Resultados para pequenas perturbações são apresentados para quatro diferentes cenários de operação, que incluem combinações entre variações de carga e contingências $N-1$.
- ❑ **Capítulo 6:** Resultados para estabilidade transitória são apresentados. Neste caso, combinações de três níveis de carregamento, faltas em diferentes linhas, trechos e tempos de abertura são exploradas.
- ❑ **Capítulo 7:** Conclusões sobre os resultados a pequenas e grandes perturbações são discutidas e, o potencial para novas pesquisas é abordado.

Capítulo 2

Estabilidade do ângulo do rotor

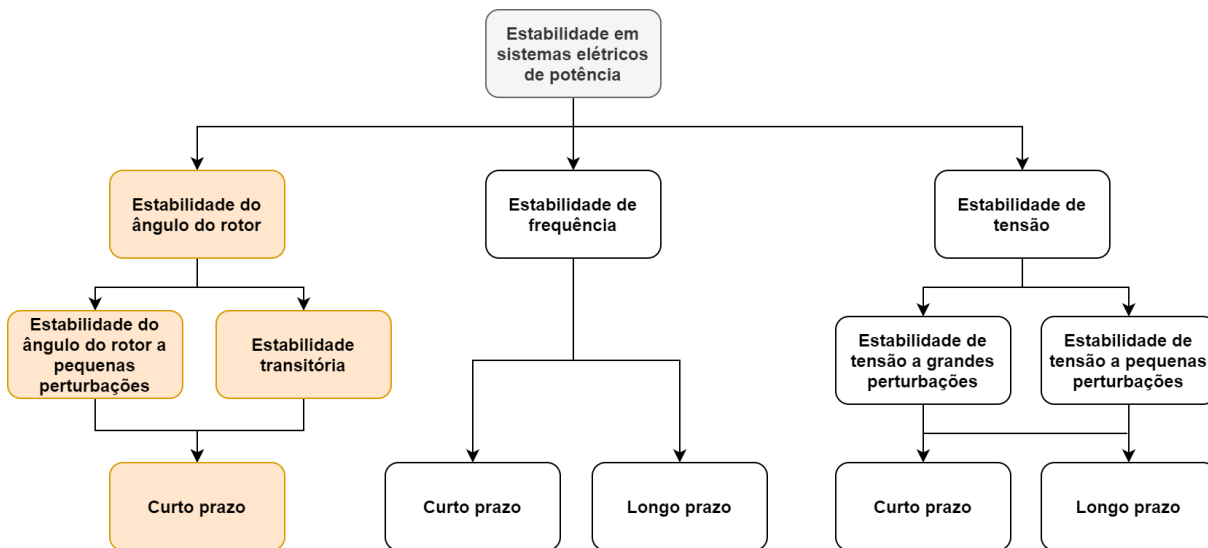
Avaliar a estabilidade de um sistema interconectado e identificar estratégias e mecanismos para que o sistema permaneça estável mesmo na ocorrência de eventos inesperados, como a saída de uma linha principal ou de contingências é essencial para garantir a operação segura de uma rede interconectada (GOMEZ-EXPOSITO; CONEJO; CANIZARES, 2011). Diante da importância desse conceito e levando em conta o embasamento teórico necessário para o desenvolvimento desse trabalho, neste capítulo é feita uma breve introdução ao conceito de estabilidade, com ênfase na categoria analisada neste trabalho, a estabilidade angular do rotor. Em sequência, as principais ferramentas clássicas utilizadas para avaliação da estabilidade angular do rotor são apresentadas.

2.1 Introdução ao conceito de Estabilidade e as suas categorias

Conforme já mencionado no capítulo de introdução, a estabilidade de SEP refere-se à habilidade da rede em permanecer em estado de equilíbrio e retomar as condições normais de operação após ter sido submetida a uma perturbação (KUNDUR et al., 2004; GOMEZ-EXPOSITO; CONEJO; CANIZARES, 2011). Uma condição normal de operação é estabelecida se os fluxos de potência nas linhas, a magnitude da tensão e ângulos das barras estão dentro de limites aceitáveis (em uma região viável do espaço de estado), os quais são definidos pelos operadores do sistema (Balu et al., 1992).

Idealmente, o sistema deve ser resiliente a distúrbios de diversas naturezas e intensidades, desde pequenas variações aleatórias de carga, as quais ocorrem continuamente ao longo da operação do sistema; curtos-circuitos nas linhas de transmissão ; perda de gran-

Figura 1 – Classificação de estabilidade em SEP



Fonte: Adaptado de (KUNDUR, 1994a)

des geradores ou cargas ou, ainda, a perda de conexão entre dois subsistemas. Portanto, avaliar a estabilidade de uma rede interconectada é essencial para que ações preventivas e corretivas sejam tomadas com intuito de assegurar a operação em condições normais, permitindo o fornecimento ininterrupto de energia elétrica aos consumidores, mesmo após a ocorrência de uma perturbação.

O estudo da estabilidade em SEP está dividido em três categorias principais, as quais estão relacionadas com a natureza física do mecanismo que resulta nesta instabilidade. As três variáveis que regem essa classificação são: ângulo do rotor dos geradores, magnitude das tensões nas barras e frequência do sistema. Já as subcategorias que compõem o problema estão relacionadas a dois principais critérios (KUNDUR et al., 2004):

- ❑ Magnitude do distúrbio, pequena ou grande perturbação, o qual irá influenciar no método de cálculo e de previsão;
- ❑ Os dispositivos de atuação, os processos envolvidos e janela de tempo, necessários para avaliar a estabilidade.

A Figura 1 exhibe essas categorias e as subcategorias que compõem o estudo da estabilidade de um SEP. A categoria que é investigada nessa pesquisa aparece em destaque nesta figura, a estabilidade angular do rotor.

A estabilidade de tensão refere-se a capacidade do sistema de, quando submetido a perturbações, manter os níveis de tensão em todas as barras dentro de limites aceitáveis, definidos pelo Operador do Sistema. Para isso, o sistema deve ser capaz de manter ou reestabelecer o equilíbrio entre a potência fornecida pelos geradores e demanda das cargas. Quando não é possível manter esse equilíbrio, a instabilidade ocorre como queda ou aumento progressivo do nível das tensões de algumas barras. Potencialmente, quando

não corrigida, a instabilidade dessa natureza pode resultar em corte de cargas de uma área inteira, abertura de linhas de transmissão e outros elementos por dispositivos de segurança, provocando um efeito em cascata e o ao colapso do sistema em situações extremas (KUNDUR, 1994a).

A estabilidade de frequência está vinculada à capacidade do sistema em manter ou reestabelecer a frequência do sistema mediante a uma grande perturbação, a qual resulte em desequilíbrio significativo da potência ativa entre a geração e a demanda. Nesta análise, estuda-se as variações na frequência do sistema devido às grandes mudanças no equilíbrio geração-demanda (GOMEZ-EXPOSITO; CONEJO; CANIZARES, 2011). A instabilidade da frequência é caracterizada por oscilações sustentadas da frequência, podendo ocasionar corte dos geradores e cargas (KUNDUR, 1994a).

Por definição, a estabilidade do ângulo do rotor refere-se à habilidade das máquinas síncronas de um sistema interconectado permanecerem em sincronismo após serem submetidas a um distúrbio (KUNDUR et al., 2004). A perturbação leva ao desequilíbrio entre a potência mecânica de entrada e a potência elétrica de cada gerador. Devido a esse desequilíbrio temporário, que pode resultar em um excesso ou *déficit* de energia nas máquinas, os rotores das máquinas conectadas no sistema oscilam uns contra os outros, trocando energia por meio das linhas de transmissão. Os problemas de estabilidade ficam, então, evidentes por meio das oscilações na potência/frequência de alguns geradores, que não irão operar de forma síncrona com os demais (GOMEZ-EXPOSITO; CONEJO; CANIZARES, 2011).

Especificamente, quando há desequilíbrio entre o torque mecânico e elétrico provocado por algum distúrbio, há aceleração ou desaceleração das máquinas conectadas no sistema. No entanto, quando um gerador acelera em relação aos demais, o ângulo de seu rotor avança em relação aos ângulos dos geradores mais lentos. A diferença angular resultante gera a transferência de parte da carga do gerador mais lento para o gerador mais rápido, levando a desaceleração da máquina mais veloz e, portanto, reduzindo a diferença angular dos geradores. Contudo, a relação potência-ângulo é altamente não linear, de modo que a partir de um determinado limite, o incremento no ângulo irá resultar na redução da transferência de potência dos geradores mais lentos para o mais veloz. Nesta condição, a diferença angular entre geradores irá se agravar, resultando na instabilidade do sistema (KUNDUR, 1994a). Também é válido ressaltar que a perda de sincronismo não se restringe a geradores únicos. Grupos de geradores podem perder o sincronismo em relação a outros grupos de geradores (KUNDUR et al., 2004).

Em sua operação, o sistema está sujeito a uma grande variedade de distúrbios que podem levar ao desequilíbrio temporário entre a potência elétrica e a potência mecânica de cada uma das máquinas síncronas conectadas no sistema. O sistema está susceptível a pequenas perturbações, como as variações diárias de carga, bem como a eventos de natureza severa, como a perda de uma linha de transmissão, ou de um grande gerador.

De acordo com a magnitude desse distúrbio existem técnicas e estratégias mais adequadas para investigar e mitigar os problemas relacionadas com a estabilidade do ângulo do rotor. Dessa forma o estudo da estabilidade angular do rotor pode ser dividida em duas subcategorias segundo a magnitude da perturbação (KUNDUR et al., 2004): estabilidade do ângulo do rotor a grandes perturbações, mais conhecida como estabilidade transitória, e a pequenas perturbações. As características dessas duas classes são apresentadas nas seções a seguir.

2.2 Estabilidade transitória

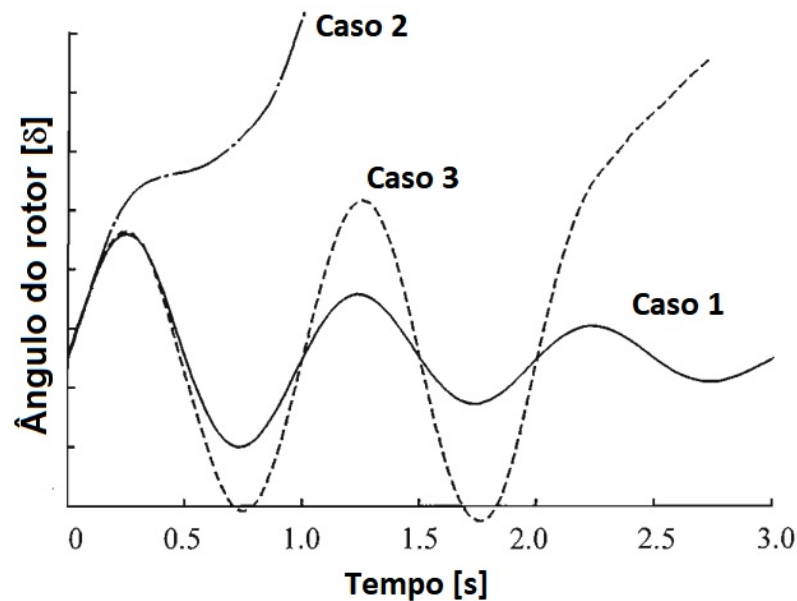
Segundo Kundur et al. (2004), a estabilidade transitória refere-se à capacidade dos geradores, conectados ao SEP, em manterem o sincronismo, quando submetidos a distúrbios de natureza severa, ou seja, uma grande perturbação. Um distúrbio é classificado como grande perturbação quando é capaz de afastar de forma significativa a trajetória de um sistema dinâmico do seu ponto de equilíbrio, de modo que as não linearidades inerentes de um sistema de potência não podem ser desprezadas e, conseqüentemente, o sistema linearizado em torno desse ponto não representa satisfatoriamente a resposta do sistema original (THEODORO, 2017). Neste cenário, a resposta dinâmica do sistema será dominada pelas não linearidades do sistema de potência, resultando em excursões de elevada amplitude do ângulo dos geradores.

Para este estudo, o modelo matemático utilizado é um conjunto de equações diferenciais não lineares que caracterizam as principais dinâmicas do sistema de potência. Uma das principais preocupações da área tem sido a manutenção do sincronismo entre as máquinas após a ocorrência do distúrbio (ALBERTO; BRETAS, 2000). O intervalo de tempo de interesse é usualmente de 3 a 5 segundos após a eliminação da falta, período no qual a resposta dos controladores usualmente não é capaz de provocar efeitos significativos no comportamento do sistema. Para grandes sistemas de potência com modos de oscilação entre áreas, este intervalo pode ser estendido para 10 a 20 segundos (Hatziargyriou et al., 2020).

A Figura 2 apresenta 3 possíveis cenários desencadeados por distúrbios transitórios. No caso 1, o sistema é estável, e a resposta do ângulo do rotor é representada por oscilações decrescentes em amplitude. No caso 2, o ângulo do rotor apresenta crescimento monotônico, levando a perda de sincronismo e, conseqüentemente, a instabilidade. Essa forma de instabilidade normalmente ocorre na primeira oscilação da resposta do gerador, período conhecido pelo termo em inglês como *first-swing*. Finalmente, o caso 3 apresenta um sistema estável no *first-swing*, mas cujas oscilações crescem em amplitude e ocasionam instabilidade. Esta última não é necessariamente resultado de uma grande perturbação, podendo também ser ocasionada por uma pequena perturbação.

A instabilidade transitória pode resultar emilhamento não intencional, provocar inter-

Figura 2 – Resposta do ângulo do rotor diante de distúrbio transitório.



Fonte: Adaptado de (KUNDUR, 1994a).

rupções em cascata e a propagação de *blackouts* ao longo do sistema. Devido à gravidade das consequências provocadas pela instabilidade transitória, é importante que o sistema opere com uma margem de estabilidade. Para garantir essa margem de estabilidade na operação, é essencial que o sistema seja monitorado e que informações sobre indícios de sobrecarregamento da rede, que podem resultar em uma resposta inadequada do sistema sejam prontamente levados ao operador do sistema de transmissão.

As características da resposta do sistema a uma grande perturbação transitória e se o sistema será estável ou instável após essa perturbação são influenciadas por uma série de fatores, dos quais pode se citar (KUNDUR, 1994a):

- Quão carregados estão os geradores;
- A potência de saída do gerador durante a falta;
- O tempo de abertura, ou seja, o tempo necessário para eliminação da falta;
- A reatância do sistema de transmissão no período pós-falta;
- A inércia do gerador. Quanto maior for a inércia, mais lenta é a variação do ângulo. Isso reduz o ganho de energia cinética na falta e as chances de instabilidade;
- A magnitude da tensão de excitação do gerador;
- A magnitude de tensão na barra infinita.

Em um problema multimáquinas, para definir se o sistema irá operar de forma estável após a ocorrência de uma perturbação, existem algumas ferramentas consolidadas como

simulação no domínio do tempo das equações algébrica-diferenciais não lineares que representam o sistema, métodos energéticos e índices de estabilidade (KUNDUR, 1994a). Essas ferramentas serão discutidas, brevemente, com um pouco mais de ênfase aos índices de estabilidade, que serão usados nesse trabalho para efetuar a classificação.

2.2.1 Ferramentas para estudo da estabilidade transitória

Na prática, o sistema elétrico de potência apresenta estrutura complexa, composta por equações diferenciais não lineares de alta dimensão. Dessa forma, a análise precisa da estabilidade transitória, normalmente, requer modelagem detalhada das unidades geradoras, outros equipamentos, além da topologia da rede. Historicamente, métodos energéticos e simulações no domínio do tempo tem sido empregadas para solucionar esse problema (YU et al., 2018).

Os métodos energéticos, tais como método de Lyapunov, PEBS, BCU se apresentam como abordagens diretas para a análise de estabilidade transitória, permitindo caminhos para a estimação da área de atração ou região de estabilidade do sistema, o que é uma vantagem fora da capacidade de técnicas de linearização. Um dos grandes problemas com tais técnicas é a necessidade de escolha de funções auxiliares, o que não é uma tarefa trivial para sistemas multimáquinas de grande porte. De forma que a escolha incorreta das funções auxiliares pode prejudicar a estimativa da área de atração (ALBERTO; BRETAS, 2000).

A simulação no domínio do tempo consiste da resolução das equações diferenciais não lineares por métodos de integração numérica, tais como Euler, Euler modificado, Newton-Raphson e Runge-Kutta (R-K) (TELEGINA, 2015). Para o estudo da estabilidade transitória, avalia-se a resposta do sistema para um determinado conjunto de contingências, com grande probabilidade de ocorrência, a partir de simulações *offline* do modelo do sistema (YU et al., 2018).

O modelo mais simples para representar as oscilações eletromecânicas no sistema elétrico de potência é descrito pela equação de *swing* (KUNDUR, 1994a):

$$J \frac{d^2 \delta_m}{dt^2} = T_m - T_e \quad (1)$$

em que J é o momento de inércia total da máquina síncrona, δ_m é o ângulo mecânico do rotor, T_m é o torque mecânico do rotor e T_e é o torque elétrico do rotor.

Se o SEP se encontra em estabilidade, então o balanço de energia gerada e consumida é mantido e todos os geradores síncronos estão operando com a mesma velocidade angular elétrica. Porém, um distúrbio grave pode causar o desequilíbrio entre o torque mecânico e o torque eletromagnético do rotor, levando a máquina a acelerar, se mais potência é gerada do que consumida, ou desacelerar, se a potência gerada não é o suficiente para cobrir a demanda. No caso de distúrbios severos, pode-se atingir um estado em que uma

ou mais máquinas perdem o sincronismo em relação às demais, o que ocasionaria danos aos equipamentos, além de perdas econômicas e interrupção do fornecimento de energia (KUNDUR et al., 2004).

Para análises de segurança em tempo real, a simulação no domínio do tempo não é atrativa, devido ao seu alto custo computacional, que ocasiona, também, alto tempo de processamento, podendo levar até algumas dezenas de segundos para ser executada, além de apresentar forte dependência na modelagem adequada do sistema e ser suscetível a variações nos valores de parâmetros reais em relação às suas contrapartes na modelagem (YU et al., 2018; KUNDUR et al., 2004). Uma forma alternativa de avaliação de estabilidade transitória é a utilização de índices de estabilidade.

Um índice de estabilidade transitória pode ser definido como a quantificação da capacidade do sistema em manter sua estabilidade transitória. Nessa categoria de indicadores, pode-se ressaltar o índice de estabilidade transitória baseado na diferença angular do rotor (TSI) (YU et al., 2018), energia transitória cinética (TKE) e energia potencial transitória (TPE) (SAJADI; PREECE; MILANOVIĆ, 2020).

No índice TKE, calcula-se a energia cinética transitória do sistema logo após a falta. Para isso, a seguinte expressão é utilizada:

$$TKE = \sum_{i=1}^n \frac{J_i \Delta\omega_i^2}{2} \quad (2)$$

em que J_i é o momento angular do i -ésimo gerador e $\Delta\omega_i$ é a variação da velocidade angular do mesmo.

O TPE é a energia potencial transitória dos geradores logo após a eliminação da falta, o qual é calculado a partir da potência ativa injetada pelos geradores, conforme a expressão:

$$TPE = \int_{t_{falta}}^{t_{eliminação}} [\Delta P_{G_i} - \Delta P_{G_j}] \Delta f_{ij} dt \quad (3)$$

em que ΔP_{G_i} e ΔP_{G_j} são a potência ativa de um par de geradores i e j , Δf_{ij} é a diferença de frequência entre esses geradores. t_{falta} e $t_{eliminação}$ são os tempos em que se inicia a falta e se elimina a mesma.

No TSI é avaliada a diferença máxima entre quaisquer dois geradores do sistema, segundo a equação (4), em que δ_{max} é a diferença máxima entre os ângulos de quaisquer dois geradores do sistema, imediatamente, após a falta.

$$TSI = \frac{360 - \delta_{max}}{360 + \delta_{max}} \quad (4)$$

Em (4), quanto mais próximo o valor do índice a unidade, maior a probabilidade do sistema ser estável. Em alguns trabalhos, para o monitoramento em tempo real, estabelece-se um valor mínimo para o TSI, acima do qual o sistema é considerado estável (YU et al., 2018). Neste trabalho, o TSI é usado como índice de estabilidade.

2.3 Estabilidade a pequenas perturbações

O estudo da estabilidade a pequenas perturbações, também denominada como estabilidade a pequenos sinais, avalia a capacidade do sistema de se manter em sincronismo, diante de pequenas perturbações (KUNDUR, 1994b). Assim como mencionado no estudo da estabilidade transitória, neste caso, a pequena perturbação irá provocar um desequilíbrio temporário entre o torque mecânico e eletromagnético (o qual atua em sentido oposto) associado a cada máquina síncrona. A estabilidade, então, irá depender da capacidade das máquinas síncronas em manter ou restaurar o equilíbrio entre esses dois torques de oposição (Hatziargyriou et al., 2020).

De acordo com Anderson e Fouad (1993), se as oscilações na potência e na velocidade resultante do desequilíbrio temporário provocado pela pequena perturbação são amortecidas, de modo que o sistema seja capaz de retornar a um ponto de equilíbrio, o sistema é classificado como estável. Segundo Martins (1986), o amortecimento das oscilações eletromecânicas está relacionado a algumas características do próprio sistema, tais como a sua estrutura, o nível de carregamento, aos efeitos da ação dos controles automáticos, bem como à condição inicial de operação dos estados do sistema (KUNDUR et al., 2004).

O termo pequenas perturbações refere-se aos distúrbios que não resultam em um afastamento significativo do sistema em relação ao seu ponto de equilíbrio (KUNDUR, 1994b). As variações de carga que o sistema está sujeito em sua operação diária e curto-circuitos de pequena intensidade são exemplos de pequenas perturbações. Desse modo, nesta categoria a linearização do conjunto de equações algébrico-diferenciais não lineares que representam equações do sistema é válida e, portanto, a estabilidade local do sistema pode ser avaliada a partir de uma abordagem linearizada.

A instabilidade provocada por uma pequena perturbação pode ser de duas formas: aumento do ângulo do rotor de forma aperiódica devido à falta de torque sincronizante ou aumento na amplitude das oscilações do rotor, devido à falta de torque de amortecimento. Com a atuação dos reguladores de tensão de alto ganho nos geradores, a instabilidade aperiódica foi praticamente eliminada dos sistemas elétricos de potência. Portanto, nos SEPs atuais, estabilidade a pequenas perturbações é majoritariamente um problema de amortecimento das oscilações eletromecânicas (KUNDUR, 1994a).

As oscilações eletromecânicas podem ser classificadas de acordo com a frequência e conjunto de máquinas envolvidas, da seguinte forma (KUNDUR, 1994a; KUNDUR; ROGERS; KLEIN, 1991):

- Modos locais: Associados às oscilações das unidades geradoras em relação ao restante do sistema. Normalmente, possuem frequências que se estendem de 0,7 a 2,0 Hz;
- Modos interáreas: Associados às oscilações de grupos de máquinas em uma área contra outro grupo de máquinas em uma área distinta. Normalmente, apresentam

frequências entre 0,1 e 0,8 Hz;

- Modos torsionais: Associados aos componentes rotacionais do sistema do eixo turbina-gerador. Podem ser causados pela interação entre reguladores de velocidade, controle de excitação, controle de HVDC, capacitores de compensação;

De forma geral, os modos locais são mais simples de se modelar e controlar, necessitando de detalhamento do sistema na vizinhança da planta. Por outro lado, modos interárea são mais complexos e, normalmente, necessitam de informação do sistema interconectado completo, tornando-se um ponto de maior preocupação nos SEPs (KUNDUR; ROGERS; KLEIN, 1991).

Nesse trabalho, será investigada a estabilidade dos modos interárea, ou seja, modos que configuram a oscilação de grupos de geradores de uma determinada área contra as demais áreas.

2.3.1 Representação em espaço de estados

O comportamento de um sistema dinâmico, como os sistemas elétricos de potência podem ser descritos como um sistema de n equações algébrico-diferenciais não lineares de primeira ordem.

O comportamento dinâmico não linear de um SEP é normalmente descrito a partir do seguinte conjunto de equações algébrico-diferenciais (5a), (5b) e (5c) (KUNDUR, 1994a):

$$\dot{x} = f(x, u, z) \quad (5a)$$

$$0 = h(x, u, z) \quad (5b)$$

$$y = g(x, u, z) \quad (5c)$$

em que $x \in \mathbb{R}^n$ é um vetor constituído pelas variáveis de estados do sistema, $u \in \mathbb{R}^p$ são as entradas de controle, $y \in \mathbb{R}^q$ é um vetor formado pelas saídas do sistema e $z \in \mathbb{R}^m$ corresponde ao vetor das variáveis algébricas.

A variável f representa o conjunto de n equações diferenciais de primeira ordem que descreve o vetor de estados x . h refere-se ao conjunto de m equações não lineares algébricas que relacionam o vetor de variáveis algébricas z com os vetores de entrada e de estados, u e x . O conjunto de equações algébricas não lineares representado por g , refere-se às equações que descrevem as variáveis de saída do sistema, com dimensão q .

Com a finalidade de aplicar técnicas de análise de caráter linear, as equações (5a) – (5c) são usualmente linearizadas em torno de um ponto de equilíbrio. Substituindo a equação (5b) em (5c), aplicando a expansão em série de Taylor e desprezando os termos superiores à primeira ordem, o modelo linearizado resultante no ponto de operação $\alpha_e = (x_e, z_e, u_e)$ é descrito pelas seguintes equações em espaço de estados (KUNDUR, 1994a):

$$\Delta \dot{x} = A\Delta x + B\Delta u \quad (6)$$

$$\Delta y = C\Delta x + D\Delta u \quad (7)$$

em que as matrizes A , B , C e D se referem às matrizes de estados, de controle, de saída e de transmissão direta, respectivamente. Já o símbolo Δ , caracteriza o desvio da variável em análise em relação ao seu valor de equilíbrio.

Os modos do sistema linearizado são as raízes do polinômio característico associado à matriz A , ou seja, seus autovalores. De posse dos modos, a estabilidade de um sistema pode ser avaliada da seguinte forma (KUNDUR, 1994a):

- ❑ O sistema é assintoticamente estável, quando a parte real dos autovalores é negativa;
- ❑ O sistema é instável quando, ao menos, um dos autovalores tem parte real positiva;
- ❑ Quando a parte real de um modo é nulo, não se pode afirmar algo sobre a estabilidade.

Desta forma, conforme evidenciado por Kundur (1994a), a existência de um autovalor com parte real positiva indica a instabilidade na resposta do sistema. Se este autovalor é real e positivo, a resposta do sistema será uma exponencial crescente. Caso seja um autovalor complexo com parte real positiva, a resposta será oscilatória com amplitude crescente.

Entretanto, do ponto de vista da segurança de operação do SEP, a não existência de um autovalor com parte real positiva, associado a um modo de natureza eletromecânica, não permite que a resposta dinâmica do sistema seja, automaticamente, considerada adequada. Para um autovalor complexo, a parte real muito pequena, pode resultar em uma saída excessivamente oscilatória, conduzindo a variações intoleráveis do módulo e ângulo da tensão no sistema. Para quantificar o amortecimento dessa oscilação, utiliza-se o fator de amortecimento ζ .

Considere o k -ésimo autovalor complexo associado à matriz de estados A de um sistema, representado na equação (8). O fator de amortecimento desse autovalor, em porcentagem, é calculado pela expressão definida em (9).

$$\lambda_k = \sigma_k \pm j\omega_k \quad (8)$$

$$\zeta_k = \frac{-\sigma_k}{\sqrt{\sigma_k^2 + \omega_k^2}} \cdot 100 \quad (9)$$

No sistema de transmissão, além de estarem posicionados no semi-plano esquerdo, os modos eletromecânicos também devem apresentar fator de amortecimento superior a

5%. Dessa forma, como já citado, evitam-se oscilações demasiadas em variáveis importantes do sistema, como o módulo e ângulo da tensão (SENEOULIN; HONGESOMBUT; DECHANUPAPRITTHA, 2019).

Dentre os autovalores complexos associados à matriz de estado de um sistema de potência, os modos de interesse neste trabalho são os eletromecânicos. Tais modos recebem essa alcunha por estarem associadas às interações dinâmicas entre as partes mecânicas e elétricas do geradores no sistema. Na análise da segurança dinâmica a pequenas perturbações, o fator de amortecimento e a frequência dessas oscilações devem ser continuamente monitorados, para que ações preventivas e corretivas possam ser estabelecidas caso o sistema opere com um baixo fator de amortecimento.

2.3.2 Resposta no tempo do sistema linearizado

A matriz de estados A , apresentada na seção anterior, determina a parcela de resposta do sistema que é excitada pelas condições iniciais, além da relação entre estados e suas variações no tempo. A partir dos autovalores da matriz A , que são os modos do sistema dinâmico, e dos autovetores a resposta do mesmo pode ser caracterizada.

A resposta no tempo do sistema pode ser representada pela seguinte expressão (KUNDUR, 1994b):

$$\Delta x(t) = \sum_{i=1}^n \phi_i c_i \exp(\lambda_i t) \quad (10)$$

em que c_i corresponde ao produto escalar $c_i = \psi_i^T \Delta x(0)$, no qual ψ_i^T são os autovetores à esquerda associados à matriz A e $\Delta x(0)$ refere-se à condição inicial do vetor de estados. ϕ_i são os autovetores à direita e λ_i são os autovalores do sistema.

A equação (10) fornece a resposta no tempo do movimento livre, ou seja, sem a atuação de uma entrada externa em um sistema dinâmico, em função dos autovetores e autovalores. Dessa forma, a resposta do sistema se apresenta como a combinação linear de n autovalores distintos associados à matriz A , conhecidos também como modos de resposta.

O autovalor λ_i é responsável por caracterizar a resposta do sistema e portanto denominado como modo de resposta. Por meio dos modos, como já foi explorado na seção anterior, é possível classificar a estabilidade do sistema e seu comportamento no tempo.

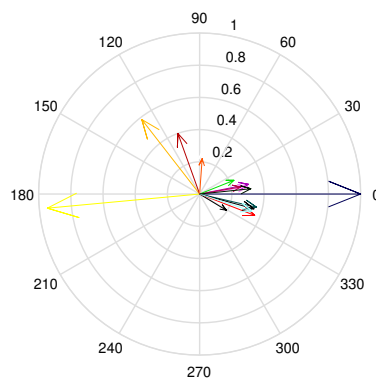
Quanto à ação dos autovetores, pode-se verificar em c_i que o autovetor à esquerda modula a contribuição das condições iniciais nos estados. Portanto, a constante c_i e o autovetor à esquerda ψ_i fornecem a magnitude da excitação inicial de cada modo.

Por fim, os autovetores à direita, determinam a intensidade com que cada modo influencia na resposta dos estados do sistema. Dessa forma, a magnitude do k -ésimo elemento do autovetor ϕ_i fornece o grau de atividade da k -ésima variável de estado em relação ao autovalor λ_i . Os autovetores à direita também fornecem o *mode shape*, que é uma ferra-

menta importante para identificar a influência de um modo em determinadas variáveis de estado do sistema, como as velocidades ou ângulos dos rotores dos geradores.

O *mode shape* ou fator de observabilidade pode ser definido como a representação gráfica composta por amplitude e fase do elemento ϕ_{ki} , da matriz de autovetores à direita, relacionando um determinado modo i com variável de estado k de interesse. Essa ferramenta possibilita definir grupos de geradores que oscilam coerentemente em relação a um determinado modo, ou seja, que partilham de respostas semelhantes frente à excitação desse modo (HENRIQUES, 2009; ZHOU et al., 2009). Para exemplificar, a Figura 3 apresenta o *mode shape* dos ângulos dos rotores de um sistema composto por 16 geradores em relação a um dos modos eletromecânicos desse sistema. Dentre os 16 geradores do sistema teste, é nítido um grupo de 12 geradores oscilando coerentemente, próximos ao vetor de módulo 1 e ângulo 0, enquanto isso, há 4 outros geradores que oscilam contra esse grupo, destacando-se o vetor de 185 °.

Figura 3 – Gráfico polar *mode shape*.



Fonte: Autoria própria.

Entretanto, a análise direta por meio dos autovetores ou *mode shape* na determinação da natureza física dos modos não é adequada, devido à dependência de unidade e escala dos autovetores. Dessa forma, se faz necessário o uso de uma ferramenta cujos elementos sejam adimensionais, para relacionar, diretamente, os modos aos estados do sistema (KUNDUR, 1994b).

2.3.2.1 Fatores de participação

Os fatores de participação são uma ferramenta importante na identificação da natureza dos modos do sistema, tratando-se de escalares que possibilitam medir a contribuição relativa dos modos do sistema nas variáveis de estado e a contribuição das variáveis de estado na resposta dos modos. São definidos a partir dos autovetores à direita e à esquerda normalizados do sistema. Antes de prosseguir é válido retomar, brevemente, as características de tais autovetores.

A matriz de autovetores à direita é composta por n vetores colunas, em que cada coluna está associada a um autovalor i e possui n linhas. Tal matriz pode ser representada por:

$$\Phi = [\phi_1 \quad \phi_2 \quad \dots \quad \phi_n]. \quad (11)$$

A matriz de autovetores à esquerda é composta por n vetores linhas, em que cada autovetor linha está associado a um autovalor i e possui n colunas, sendo descrita por:

$$\Psi = [\psi_1^T \quad \psi_2^T \quad \dots \quad \psi_n^T]^T \quad (12)$$

Para cada i -ésimo autovalor associado a matriz A , é válido ressaltar que o i -ésimo autovetor à direita e à esquerda podem ser calculados, respectivamente por meio de:

$$A\phi_i = \lambda_i\phi_i \quad (13)$$

$$\psi_i A = \lambda_i\psi_i \quad (14)$$

Os autovetores à esquerda e à direita são ortogonais para autovalores distintos. Dessa forma, para quaisquer autovalores j e i , em que $j \neq i$, tem-se a equação (15) em que o produto é sempre zero. Em contra partida, para o mesmo autovalor esse produto torna-se uma constante não nula (16).

$$\psi_i\phi_j = 0 \quad (15)$$

$$\psi_i\phi_i = C_i \quad (16)$$

É comum que os autovetores sejam normalizados por essas constantes, de forma que seu produto torne-se 1 (KUNDUR, 1994a). Dessa maneira, o produto das matrizes de autovetores resulta na identidade (17).

$$\Psi\Phi = I \quad (17)$$

Partindo dessas informações é possível apresentar com maior clareza a matriz de fatores de participação (18) (KUNDUR, 1994b; HENRIQUES, 2009; FERNANDES, 2017) que é o produto elemento a elemento das matrizes de autovetores à direita e à esquerda.

$$P = \Psi \odot \Phi \quad (18)$$

Cada elemento da matriz de participação é definido pela expressão abaixo:

$$p_{ki} = \phi_{ki} \cdot \psi_{ik} \quad (19)$$

em que p_{ki} representa a participação relativa da variável de estado k , no modo i e vice-versa (KUNDUR, 1994a). ϕ_{ki} é o elemento da linha k e da coluna i da matriz de autovetores à direita e ψ_{ik} é o elemento da linha i e coluna k da matriz de autovetores à esquerda. Diferente do mode shape, os fatores de participação são adimensionais e normalizados, permitindo uma avaliação bastante confiável das relações entre os modos e as variáveis de estado, a partir da qual é possível caracterizar os modos.

No contexto de estabilidade a pequenas perturbações, a matriz de fatores de participação é comumente utilizada para identificar os modos eletromecânicos do sistema. Tais modos são aqueles que apresentam maiores participações nas variáveis de estado de natureza eletromecânica, como ângulo do rotor e velocidade angular.

Capítulo 3

Revisão Bibliográfica

As técnicas clássicas, introduzidas no capítulo anterior, fornecem a fundamentação teórica necessária para analisar e investigar o comportamento dinâmico de um SEP e, portanto, são essenciais para o planejamento da operação da rede. No entanto, a necessidade de conhecer o estado topológico da rede, das cargas e de todos os elementos em operação (injeção de potência da geração, cargas conectadas, ajuste dos controladores, entre outros) para atualização do modelo, torna difícil o monitoramento e a identificação em tempo real a partir do modelo matemático (LEANDRO et al., 2015; MCNABB; WILSON; BIALEK, 2013). Neste contexto, torna-se cada vez mais necessário o estudo de técnicas capazes de extrair informações a partir dos dados medidos, de aprender o seu comportamento e, atualizar esse conhecimento à medida que novos conjuntos de dados sejam gerados.

Neste sentido, ao longo dos últimos anos, com o aprimoramento dos sistemas de medição, técnicas de aprendizagem de máquina têm sido amplamente utilizadas para monitorar o comportamento dinâmico da rede elétrica e fornecer subsídios para a decisão do operador. Dada a importância dessas técnicas, neste capítulo, uma breve revisão bibliográfica sobre a utilização de técnicas de aprendizagem de máquina em análise de segurança será efetuada, iniciando com um escopo amplo, que contempla análise de segurança em SEP utilizando redes neurais, SVM e árvores de decisão e, concluindo com trabalhos que empregam esses mesmos três algoritmos no contexto de estabilidade angular do rotor.

3.1 Aprendizagem de máquina na análise de segurança

Os métodos clássicos disponíveis para a avaliação da segurança em um SEP normalmente apresentam forte dependência com a precisão da modelagem do sistema e alto custo computacional (MCNABB; WILSON; BIALEK, 2013). Dessa forma, esses métodos podem não ser tão atrativos para análises em tempo real, em que se faz necessária uma resposta rápida e robusta às incertezas da operação. Vislumbrando um contexto de alta disponibilidade de dados, como as *smart grids*, justifica-se a adoção de ferramentas de mineração de dados e aprendizagem de máquina para auxiliar em tarefas como a análise de estabilidade, não se restringindo apenas ao problema da estabilidade angular do rotor.

As principais técnicas de aprendizado de máquina empregadas para avaliar a estabilidade de um rede elétrica reportadas na literatura são: redes neurais artificiais, SVM e árvores de decisão (KRISHNAN et al., 2011; ALIMI; OUAHADA; ABU-MAHFOUZ, 2020). Sendo assim, o principal objetivo dessa seção será realizar uma breve revisão acerca de trabalhos que utilizem essas técnicas.

Chung e Ying (1998) abordam o problema de avaliação da segurança de tensão, ressaltando o alto tempo necessário nos cálculos dos métodos convencionais que dificultaria o seu uso em tempo real e as limitações da utilização de métodos de equivalentes externos, em especial do tipo Ward (AMERONGEN; MEETEREN, 1982). Como alternativa ao equivalente Ward, é proposto o uso de uma rede neural de 3 camadas para estimar as injeções de potência de um equivalente externo do SEP, a partir dos fasores de tensão das barras de fronteira, auxiliando na análise de segurança do sistema reduzido. O método é testado no IEEE 39 barras, isolando uma área delimitada pelas barras 1, 4 e 16 e simulando cenários que envolvem variação de carga e perdas de linhas. A rede treinada obtém maior acurácia que o método de Ward tradicional, mantendo a qualidade da regressão mesmo diante de mudanças de topologia e, demonstrando o potencial da aplicação de redes neurais na resolução de um problema do setor elétrico.

Com relação a SVM, Kalyani e Swarup (2009) usam essa técnica para avaliação de segurança estática e transitória, a partir de limites de tensão, potência de geração, potência nas linhas e estabilidade angular do rotor. O classificador é testado nos sistemas IEEE 57 barras e IEEE 118 barras. Para a avaliação da estabilidade transitória os autores utilizam variáveis estáticas e dinâmicas, incluindo fasores de tensão das barras, potências de geração e carga das barras, potências das linhas, e os ângulos relativos do rotor no tempo da falta e no tempo de abertura. Antes da etapa de treino dos classificadores, o trabalho realiza seleção de características reduzindo os vetores de entrada entre 9% e 3,5% do tamanho original. O SVM proposto é comparado com o método dos mínimos quadrados e obtém acurácias superiores, atingindo 98,29% contra 77,71% para o IEEE 57 barras e, 96,42% contra 93,01% para o IEEE 118 barra, na tarefa de classificação de

estabilidade transitória.

Subramani et al. (2016) utiliza duas arquiteturas de redes neurais, sendo elas CFBNN (em inglês, *Cascade Feed-forward Back Propagation Neural Network*) e RBFNN (em inglês, *Radial Basis Function Neural Network*), para a análise de estabilidade de tensão, no sistema teste IEEE 14 barras. As redes são treinadas tendo como entradas as impedâncias e reatâncias de linha, potências reativa de linha, módulo de tensão na barra transmissora e defasagem angular entre barras. Como alvo, as redes tentam prever um índice global de estabilidade de tensão definido em inglês como *Global Voltage Stability Index* - GVSI por linha. Como métrica os autores utilizam MSE (em inglês, *Mean Square Error*) obtendo resultados com ordem entre 10^{-4} e 10^{-5} a depender da linha testada.

Werboston (2017) cita o alto tempo de processamento necessário para a análise de segurança dinâmica e estática, quando se utilizam simulações no domínio do tempo ou fluxo de potência continuado, como uma dificuldade no seu emprego para análises em tempo real. Dessa forma, os autores propõem a utilização de uma árvore de decisão com *split multiway* para classificação de estabilidade estática e dinâmica de um SEP. A árvore proposta realiza a divisão em função da topologia da rede, dada pela contingência, gerando n ramos que avaliam cada qual a estabilidade para uma topologia específica. Para isso, além das informações das variáveis de fluxo de carga, uma variável categórica indicando a topologia da rede no caso avaliado deve compor o vetor de variáveis de entrada para árvore de decisão. Nos testes realizados uma elevada acurácia é obtida, tanto para a estabilidade de tensão, quanto para a estabilidade angular transitória. Porém, como trabalho futuro, o autor ressalta a possibilidade da investigação da clusterização de topologias, para evitar que se torne necessário o treino da árvore para cada contingência diferente.

Os autores em Gavgani e Eftekharnajad (2017) exploram a análise de segurança a partir do estudo do impacto da perda de múltiplos elementos da rede, sem a necessidade de análises de contingência $n-k$. Neste trabalho, é proposta a elaboração de um grafo de dependência de linhas (em inglês, *Line Dependency Graph* - LDG), cujos vértices são as linhas do sistema analisado e as arestas são variações de fluxo definidas por uma função não linear dependente da topologia, carregamento e redespacho da rede. Para calcular os pesos das arestas, é proposta a utilização de uma árvore de regressão, devido o seu baixo tempo de processamento e custo computacional que a fazem um algoritmo interessante para aplicações *online*. O grafo LDG fornece uma visualização do impacto em aumento de fluxo de potência que a perda de determinadas linhas terá nas demais linhas do sistema. Além disso, o mesmo trabalho propõe o uso de um segundo grafo - grafo de ameaças imediatas (em inglês, *Immediate Threats Graph* - ITG), cujo o objetivo é definir os sobrefluxos causados pelas perdas das linhas e que depende, diretamente, do LDG. Para os testes, os autores utilizam o IEEE 30 barras e o IEEE 118 barras, atingindo tempos de 10,3 segundos e 86,7 segundos, respectivamente, para a obtenção dos grafos. Dessa forma, a análise dos impactos da perda de determinadas linhas torna-se muito

mais rápida e viável durante a operação, auxiliando na tomada de decisão por parte do operador.

Uma rede convolucional (em inglês, *Convolutional Neural Network*) é proposta por An et al. (2020) para classificar a estabilidade transitória. Os autores exploram a possibilidade de representar o estado da rede como uma imagem formada a partir da localização geográfica dos elementos, usando GIS (em inglês, *geographic information system*) e informações do fluxo de potência, que são mapeadas em diferentes cores. A representação gráfica da rede elétrica permite que a CNN utilize as informações do fluxo de potência e a topologia do sistema, como entradas. Testes são realizados nos sistemas IEEE 39 barras e IEEE 118 barras, alcançando acurácias superiores a 98%. Casos de perdas de medição não são explorados.

Recentemente, para a análise de segurança dinâmica, contemplando estabilidade angular transitória, estabilidade de tensão de curto prazo e estabilidade a pequenas perturbações, Liu et al. (2020) propõe o uso de uma *Random Forest*, que é constituída por um conjunto de árvores de decisão (Ensemble) com o intuito de reduzir as chances de *overfitting* e aumentar o poder de generalização do algoritmo. Em técnicas do tipo *ensemble*, algoritmos individuais são treinados para realizar a mesma tarefa e, a predição é composta por um esquema de votação de todos os modelos que compõem o *ensemble*. Os autores sugerem o uso de técnicas de seleção de características como informação mútua parcial (em inglês, *Partial Mutual Information - PMI*), que captura correlações não lineares e coeficiente de correlação de *Pearson* (em inglês, *Pearson Correlation Coefficient - PCC*), para capturar correlações lineares. O vetor de características é selecionado por meio desses dois índices PMI e PCC.

Ainda com relação ao trabalho proposto em (LIU et al., 2020), sobre o sistema IEEE 39 barras utilizando a seleção de características, o vetor de entrada é composto por 32 variáveis e tem como saída a margem de estabilidade transitória, que é calculada a partir do tempo de abertura crítico e o tempo de abertura atual. Para avaliação do desempenho do método proposto, uma das métricas de avaliação utilizada é o erro residual quadrado (R^2), em que se obtêm uma acurácia de 0,9877. Também são realizados testes para um sistema de 1648 barras, em que se obtêm a acurácia (R^2) de 0,9802. O algoritmo é treinado com as contingências com maior probabilidade de ocorrência. Os autores também ressaltam a possibilidade da inserção de novas florestas no IRF treinado, quando os operadores observarem contingências que não fazem parte da base de conhecimento do algoritmo.

De forma geral, os trabalhos que abordam o uso de aprendizagem de máquina na predição de margens ou classificação de segurança apresentam arquiteturas centralizadas que partem de medições de todo o SEP, mesmo que ocorra, posteriormente, a redução da quantidade de entradas por meio de algoritmos de seleção de características. As acurácias apresentadas, normalmente, estão acima de 98%, o que evidencia o potencial das técnicas

empregadas na análise de segurança. Todavia, poucos trabalhos exploram a possibilidade de perdas de medição e alternativas para manter a classificação nestes cenários.

3.2 Aprendizagem de máquina na análise de estabilidade angular

Análogo à análise de segurança dinâmica, as técnicas de aprendizagem de máquina também são amplamente utilizadas em escopos mais específicos, como a estabilidade angular do rotor. Neste contexto, encontram-se desde trabalhos com classificação de estabilidade e predição de velocidade angular dos rotores, até clusterização de grupos de geradores coerentes. A seguir são elencados alguns trabalhos exemplificando o uso de tais técnicas.

Em Angel, Glavic e Wehenkel (2003), DEL ANGEL et al. (2007) apesar de não se avaliar diretamente a estabilidade angular do rotor, os autores propõem a estimação das variáveis de estado mecânicas da máquina, as quais podem ser utilizadas na análise de segurança dinâmica. Angel, Glavic e Wehenkel (2003) consideraram a disponibilidade de medições em tempo real e de diversos pontos da rede fornecidas pelas PMUs para propor a utilização de séries temporais dos fasores de corrente e tensão na barra terminal conectada à máquina a qual se quer estimar velocidade e ângulo. Com esse intuito, a estrutura proposta contém duas redes neurais *feed-forward*. A primeira rede utiliza a amplitude e os ângulos da tensão e corrente da barra de geração com uma janela temporal de três medições, totalizando doze entradas e tendo como objetivo a predição do ângulo. A segunda rede utiliza as mesmas doze entradas da primeira, mais três predições de ângulo da rede neural anterior e tem como objetivo a predição da velocidade angular do rotor.

Já os mesmos autores em DEL ANGEL et al. (2007) voltam a utilizar redes neurais para a estimação do ângulo do rotor, porém com o uso de duas redes distintas, sendo uma para a predição do seno do ângulo do rotor e a segunda para a predição do cosseno do ângulo do rotor. Posteriormente, há uma última etapa em que se calcula o ângulo a partir dessas duas predições.

Com relação a classificação da estabilidade transitória, em relação ao estudo da estabilidade angular do rotor, usualmente, o modelo de predição é estabelecido a partir de uma estrutura centralizada, em que informações da resposta do sistema medida em pontos diferentes do sistema são necessárias.

Por exemplo, os autores Amraee e Ranjbar (2013) utilizam uma árvore de decisão do tipo C4.5 para identificar o *status* de estabilidade da rede. Para realizar a classificação, o modelo de predição têm como variáveis de entrada as seguintes medidas: a variação de energia cinética (KE), a potência mecânica (P_m), a aceleração média (α_{ave}), potência elétrica (P_e) no momento da eliminação da falta, e o tempo de falta (TC) associada a cada um dos geradores conectados no sistema teste. A eficácia da técnica é avaliada por

simulações realizadas em dois sistemas testes: sistema IEEE 9 barras e no 1969 barras que representa a rede nacional do Irã. Uma alta acurácia é atingida pela DT, alcançando uma acurácia de 99,91% para o maior sistema teste, no entanto, variáveis de difícil medição são utilizadas.

Recentemente, um ponto bastante avaliado com relação a essa classificação é quantidade de ciclos de medição da resposta transitória do sistema, em que se investiga a acurácia da técnica de acordo com o número de medições da série temporal da resposta do sistema, no período pós-falta.

Por exemplo, Guo e Milanovi (2016) propõem uma metodologia composta por dois estágios, um para predição de estabilidade transitória e, em cenários de instabilidade, a identificação de grupos de geradores com comportamentos dinâmicos similares. Seis bancos de dados com diferentes quantidades de ciclos de medição pós-falta, indo de 10 à 60 ciclos, são gerados a partir de combinações de cenários de variações de carga e contingências $N-1$. Testes são realizados com o sistema IEEE 68 barras, tendo os ângulos dos rotores de todos os geradores conectados no sistema teste. A DT obtêm o melhor resultado para o vetor de 60 ciclos, atingindo uma acurácia de 99,75%. Dessa forma, o trabalho aborda o efeito de diferentes quantidades de ciclos de medição na acurácia do classificador.

Já os autores em Yu et al. (2018) propõem o uso de uma rede neural do tipo LSTM (*Long Short-Term Memory*) para a análise de estabilidade angular transitória, explorando o *tradeoff* entre quantidade de ciclos de medição e acurácia do modelo. Cenários de teste são compostos por três níveis de carregamento, contingências $N-1$, curto-circuitos trifásicos variando entre diferentes trechos da linha e variações de tempo de abertura. Como entrada, o modelo recebe medições de fasores de tensão de todas as barras do sistema, as saídas são calculadas a partir de índice de estabilidade derivados dos ângulos dos rotores no pós-falta. Testes são realizados no IEEE 39 barras, IEEE 145 barras e IEEE 162 barras, obtendo-se acurácias de 100%, 99,98% e 100% e tempos de resposta médio em ciclos de 1,448, 2,047 e 1,901, respectivamente. Os autores também exploram o efeito do emprego de diferentes quantidades de PMUs na acurácia e tempo de resposta, sugerindo que é possível obter resultados semelhantes aos obtidos quando se usa todas as barras, com uma quantidade consideravelmente reduzida de PMUs.

No que tange o estudo da estabilidade a pequenas perturbações, técnicas de aprendizagem de máquina têm sido aplicadas para classificar o amortecimento dos modos eletromecânicos em estável e instável bem como na predição da parte real e parte imaginária desses modos (TEEUWSEN; ERLICH; EL-SHARKAWI, 2003; MCNABB; WILSON; BI-ALEK, 2013; WANG et al., 2015; CAO; FAN, 2018). Por exemplo, em Teeuwsen, Erlich e El-Sharkawi (2003), uma abordagem baseada em rede neural artificial é proposta para a classificação da estabilidade angular a pequenas perturbações do SEP bem como para identificação de regiões onde autovalores críticos podem ser encontrados. Neste trabalho

os testes são realizados no sistema 68 barras. As entradas são obtidas a partir de seleção de características em fasores de tensão nas barras, potências nas linhas de transmissão e energia de rotação dos geradores.

Uma DT é proposta por McNabb, Wilson e Bialek (2013) para a classificação de fator de amortecimento de modo interárea como estável ou instável. O modelo recebe como entradas variáveis de fluxo de potência (Como potência ativa e reativa de geração, tensão, ângulo de fase, fluxo de potência ativa e reativa). Para mostrar sua eficiência, testes são realizados com um sistema de 140 barras e dados coletados no sistema elétrico de potência do Equador, atingindo-se acurácias de 96,0% e 96,9%, respectivamente.

Sulla et al. (2015) utiliza dados de dois anos de medição de cinco PMUs distribuídas entre o sul da Finlândia e a Suécia contendo também as medições de frequência e desvio de frequência, com taxa de amostragem de 50 Hz, para o treinamento de três tipos de redes neurais visando a classificação do amortecimento de um modo eletromecânico de 0,35 Hz. É interessante ressaltar que os autores propõem um método que utiliza apenas dados de medição e, para tanto o amortecimento do modo eletromecânico é calculado por dois algoritmos *mode-meter*, que são o LSMYW (em inglês, *Least Square Modified Yule Walker*) e SI (em inglês, *Subspace Identification*). Os autores também exploram uma etapa bastante importante na construção de classificadores que é a extração de características, com o objetivo de reduzir a dimensão do vetor de entrada e assim reduzir o tamanho das redes, chances de *overfitting* e tempo de processamento a partir da técnica PCA (em inglês, *Principal Component Analysis*).

Wang et al. (2015) aplica DT para a classificação de amortecimento dos modos eletromecânicos em uma versão modificada do IEEE 68 barras com gerações eólica e fotovoltaica. Medições de PMU relativas a variáveis que apresentam alta observabilidade são escolhidas como entradas das DTs, que tem como saídas a classificação do amortecimento dos modos em muito inadequado, inadequado e marginalmente adequado.

Cao e Fan (2018) obtêm alta acurácia na classificação de estabilidade a pequenas perturbações usando uma versão modificada do IEEE 68 barras com alta penetração de energia renovável. Nesse trabalho, os autores investigam a viabilidade da aplicação de uma rede neural profunda (DNN). Na criação do *dataset*, são consideradas variações de carga, contingências $N-1$, e os dados de operação de 7 anos de energia renovável do sistema de transmissão alemão. Nos testes, a DNN é capaz de obter uma acurácia de 99,89%. Todavia, uma rede neural simples (não profunda) também é usada para critério de comparação e, obtêm uma acurácia de 99,52%, bastante similar à DNN.

Recentemente, Asvapoositkul e Preece (2021) usa uma árvore de decisão do tipo *Extreme Gradient Boosting* (XGBoost) para a predição de fator de amortecimento de modos inter-área para diferentes condições de operação. Além disso, baseado em cenários de baixo amortecimento, os autores propõem a otimização do redespacho para amortecimento de modos críticos do SEP. Em (ASVAPOOSITKUL, 2021) o autor investiga o

uso de clusters de contingências com padrões dinâmicos parecidos para gerar uma entrada categórica e melhorar a predição de fator de amortecimento, aplicando o algoritmo XGBoost.

3.3 Análise Geral

Os trabalhos citados ao longo do capítulo são uma parte do esforço feito na avaliação de segurança do SEP e, em específico, na análise de estabilidade angular do rotor, para mitigar as dificuldades dos modelos clássicos diante das incertezas relacionadas ao modelo matemático de SEP, em constante mudança, e para reduzir o ônus computacional da aplicação em tempo real.

Apesar do alto desempenho obtido pelos classificadores, nota-se que algumas pesquisas utilizam medições que dificilmente estão disponíveis ou que apresentam pouca precisão em ambiente prático, como o ângulo do rotor, velocidade angular da máquina ou potência mecânica. Nesse sentido, visando propor métodos que possam ser facilmente aplicados, é interessante que se explore o uso exclusivo de medições de PMU, que apresentam alta disponibilidade, em ambiente real, como proposto em diversos trabalhos previamente citados.

Um ponto pouco explorado é a perda de medições durante as avaliações de segurança em tempo real. Nesse cenário, esse trabalho propõe especificamente para a análise de estabilidade a pequenas perturbações, um método baseado em classificadores distribuídos, que apresenta alto grau de redundância. Nesse método, será explorado o potencial de classificadores individuais e independentes, sendo que cada um deles usa apenas a medição de uma barra geradora para identificar casos de instabilidade. A redundância proposta implica em maior robustez e segurança do sistema, de forma que, em caso de perda de medição por parte de um dos classificadores, qualquer outro classificador será capaz de avaliar a estabilidade do sistema.

No contexto de estabilidade transitória, este trabalho propõe a investigação do impacto de diferentes quantidades e ciclos de medição do pós-falta no desempenho de classificadores centralizados e distribuídos, combinando-se medições de magnitude de tensão e ângulo de fase das barras geradoras. Dessa forma, busca-se encontrar a quantidade mínima de ciclos de medição que possa oferecer a acurácia desejada ao operador do sistema, disponibilizando mais tempo para atuação de controle e mitigando eventuais danos. Por fim, também se investiga conjuntos mínimos de medição que preservem o bom desempenho dos modelos, visando a economia com o uso de uma quantidade reduzida de PMUs.

A escolha da árvore de decisão como algoritmo de aprendizagem de máquina, como ressaltado nos trabalhos revisados, reside na simplicidade de implementação, baixo custo computacional, que facilita o processo de embarcar o classificador e na interpretabilidade das regras geradas pelas árvores, que pode auxiliar na tomada de decisão por parte do

operador. É importante destacar também que a baixa complexidade do algoritmo implica na necessidade de menos amostras para o treinamento.

Capítulo 4

Método Proposto

Com o objetivo de desenvolver uma estrutura que possibilite ao operador avaliar em um curto intervalo de tempo a estabilidade angular do rotor a partir de medições disponíveis em ambiente real, a estrutura proposta é composta por dois métodos distintos. Ambos métodos têm como premissa classificar a estabilidade angular do rotor a partir de dados que podem ser facilmente amostrados em uma *smart grid*, tais como o fasor da tensão, a potência consumida pelas cargas e a injeção de potência dos geradores. O algoritmo baseado em árvores de decisão é utilizado nos dois métodos propostos para classificação, o que torna possível compreender as regras adotadas pelo classificador e, conseqüentemente, auxiliar no processo de adoção de medidas corretivas ou preventivas para garantir a segurança da operação.

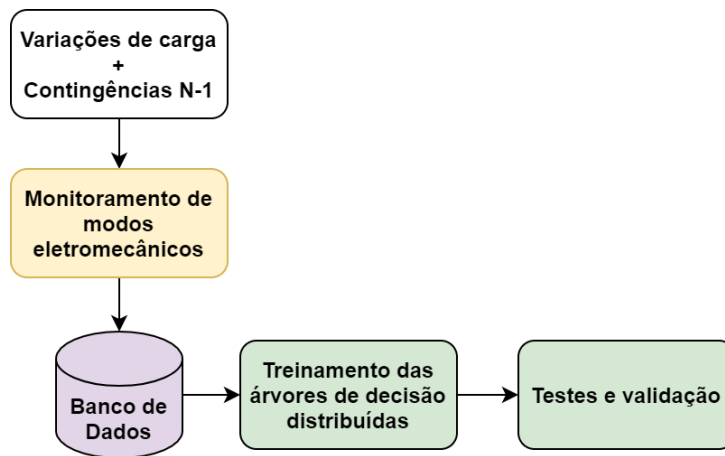
O primeiro método têm como finalidade monitorar a estabilidade a pequenas perturbações durante a operação diária do SEP. Para isso, uma abordagem descentralizada é proposta, a qual utiliza informações da barra de geração, em que o medidor está conectado, por exemplo, de uma PMU ou um *smart meter*. Essa estrutura irá fornecer suporte à decisão do operador do sistema quanto à classificação do amortecimento dos modos eletromecânicos, identificando se os mesmos têm um comportamento estável ou instável para uma determinada condição de operação.

Apesar do primeiro método fornecer subsídios para operação diária da rede, o mesmo não é capaz de fornecer informações sobre a estabilidade angular do rotor, quando o sistema está sujeito a eventos de elevada magnitude. Nestas condições de operação, o segundo método deverá ser utilizado para fornecer informações ao operador quanto à estabilidade do sistema. Para isso, uma abordagem centralizada é aplicada, na qual as medidas realizadas nas barras de geração são enviadas para uma única árvore. O classificador irá identificar se a resposta do sistema pós-perturbação será estável ou instável,

investigando, principalmente, a ocorrência de instabilidade no *first swing*.

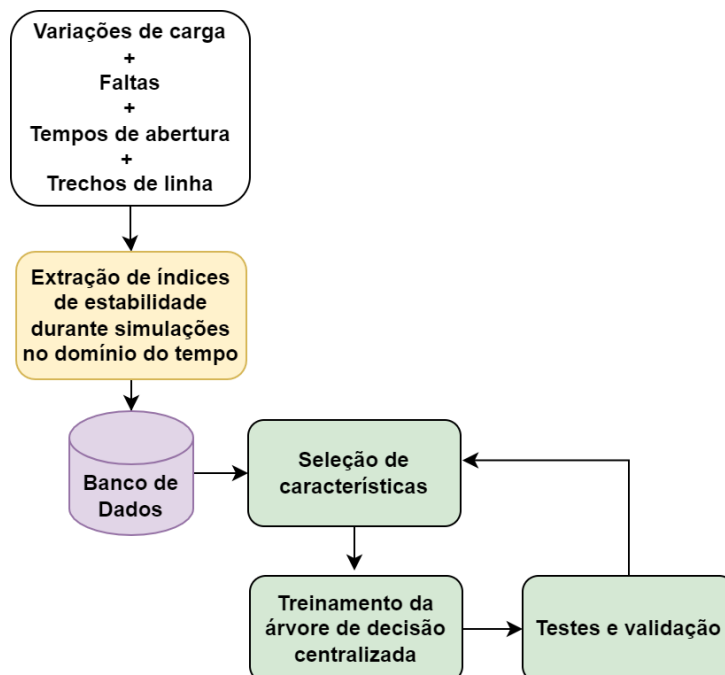
A Figura 4 apresenta a abordagem proposta na análise de estabilidade a pequenas perturbações, enquanto que a Figura 5 apresenta a proposta para análise transitória. De forma geral, os fluxos dos dois diagramas de blocos são semelhantes, compreendendo etapas de montagem de banco de dados a partir de simulações com diferentes perturbações e extração de indicadores de estabilidade. Na análise a pequenas perturbações monitora-se os modos eletromecânicos e, a partir deles, se define a estabilidade do sistema, enquanto que na análise de estabilidade transitória são utilizados índices de estabilidade.

Figura 4 – Abordagem para a análise de estabilidade a pequenas perturbações



Fonte: Autoria própria.

Figura 5 – Abordagem para a análise de estabilidade transitória



Fonte: Autoria própria.

Inicialmente, são apresentados conceitos relativos às árvores de decisão, que será o classificador empregado nas duas abordagens. Posteriormente, as etapas que compõem a análise a pequenas perturbações são discutidas. Algumas técnicas e métricas empregadas durante as fases de treino e teste são comentadas. Por fim, as etapas que compõem a análise de estabilidade transitória são apresentadas.

4.1 Árvores de decisão

Na predição e análise da avaliação da segurança de um SEP, técnicas de aprendizado de máquina como redes neurais, SVM e árvores de decisão têm sido amplamente utilizadas para auxílio na tomada de decisão do operador (LIU et al., 2014). Dentre esses algoritmos, as árvores de decisão se distinguem pela simplicidade de interpretação, expondo limiares de estabilidade para variáveis conhecidas do sistema elétrico de potência e, dessa forma, facilitando o processo de tomada de decisão, em tempo real, por parte operador (KRISHNAN et al., 2011; OLIVEIRA et al., 2017; PATIDAR; TIWARI, 2013).

A árvore de decisão também apresenta como outros benefícios a eficiência no tempo de processamento e um baixo esforço computacional para realizar uma classificação ou regressão, quando comparada com outras técnicas de aprendizagem de máquina (CZAJKOWSKI; JURCZUK; KRETOWSKI, 2020).

Vale ressaltar que na classificação, a partir dos dados de entrada o modelo tem como objetivo estimar a qual classe a amostra pertence. Sendo que as classes são categorias de interesse do problema. No caso deste trabalho, as classes as quais as amostras podem pertencer são estável ou instável.

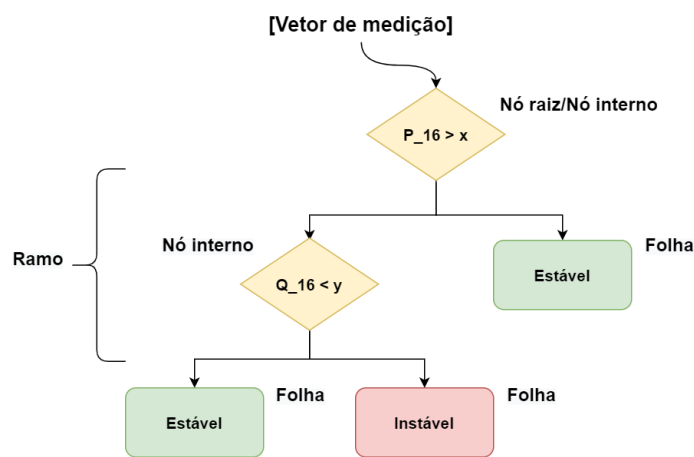
Quanto a sua estrutura, a árvore de decisão pode ser decomposta em:

- **Nó raiz:** O ponto de partida da árvore, responsável pela primeira divisão e que contém todas as classes de treinamento juntas;
- **Nós internos:** São compostos por regras que definem a divisão das amostras em novos ramos, a partir das entradas da árvore. Podem gerar divisão binárias ou múltiplas, dependendo do tipo de árvore empregada;
- **Ramos:** São compostos por conjuntos de regras (nós) e as divisões das amostras;
- **Folhas:** Nós terminais em que se atingiu um critério de parada ou a presença de uma única classe no nó.

As árvores são geradas a partir de uma estratégia *top-down*, que inicia a partir de um nó raiz, contendo todas as classes possíveis. Após a identificação do nó raiz, o algoritmo segue buscando regras de divisão que melhor separem as classes em cada nó filho, recursivamente. O processo de divisão de nós perdura até que não seja mais possível reduzir a diversidade nos próximos nós ou que se tenha nós com apenas uma classe. Os nós onde

não ocorrem novas divisões recebem o nome de folha (MARCO; MARTINS; FERRAZ, 2013). A Figura 6 ilustra uma árvore recebendo um vetor de medições, como entrada, e realizando a classificação do caso a partir de regras referentes a duas variáveis desse vetor. Neste exemplo, em específico, P_{16} e Q_{16} representam as potências ativa e reativa de um gerador conectado a barra 16, em cada nó interno é feita uma comparação entre o valor dessas variáveis e limiares x e y definidos pela árvore. Caso a potência ativa seja menor que x a primeira regra ($P_{16} > x$) classificaria a amostra como estável, caso contrário seria necessária mais uma verificação para definir a classe da amostra, na qual Q_{16} menor que y define a amostra como pertencente a classe estável.

Figura 6 – Árvore de decisão



Fonte: Autoria própria.

Como a maioria dos algoritmos de aprendizagem de máquina supervisionados, a construção da árvore é feita a partir da divisão de dados em dois conjuntos: um conjunto utilizado no treinamento e outro para teste. A base de treinamento é adotada para a geração da árvore e identificação de regras de decisão mais adequadas para representar as classes que compõem os dados em análise. As regras de seleção são definidas a partir do critério de impureza.

Quando um nó contém apenas uma classe, a função que caracteriza a impureza apresenta seu valor mínimo, enquanto que quando há a mesma quantidade de exemplos de todas as classes, a função apresenta o seu valor máximo. A melhor divisão é aquela que minimiza essa função nos nós filhos, quando comparados ao nó pai. O critério de impureza pode ser computado por diferentes funções, as mais utilizadas são: entropia e o índice de Gini (MCNABB; WILSON; BIALEK, 2013; QUINLAN, 1993).

A entropia é representada pela equação (20) e o índice de Gini pela equação (21), em que p_{mk} representa a probabilidade de ocorrência da classe k avaliada no nó m , dadas N

classes diferentes (MCNABB; WILSON; BIALEK, 2013; GERON, 2017).

$$H_{entropia} = - \sum_{i=1}^N p_{mk} \cdot \log_2 p_{mk} \quad (20)$$

$$H_{gini} = \sum_{k=1}^N p_{mk} \cdot (1 - p_{mk}) \quad (21)$$

Ambas as funções atingem seus máximos quando $p_{mk} = 1/N$ e esse valor pode ser encontrado a partir de multiplicadores de Lagrange com restrição dada pela equação (22).

$$\sum_{k=1}^N p_k = 1 \quad (22)$$

As três árvores de decisão mais utilizadas são a ID3, C4.5 e CART (em inglês, *Classification And Regression Trees*) (PATIDAR; TIWARI, 2013) e, portanto, são abordadas brevemente abaixo.

4.1.1 ID3

A ID3 foi apresentada por Quinlan Ross (PATIDAR; TIWARI, 2013) e, Quinlan (1993) afirma fundamentar-se nas ideias do trabalho de Hoveland e Hunt, em meados dos anos 50, sobre o processo de geração de árvores de decisão a partir de um grupo de amostras de treino. O algoritmo é do tipo *multiway*, em que cada nó pode gerar diversos nós filhos. O algoritmo busca ótimos locais, tomando como medida o ganho de informação na escolha de um determinado atributo para realizar a divisão. Tem como limitações não ter a capacidade de operar com variáveis contínuas, ter baixa acurácia com grandes *datasets* e ser difícil de escalar.

4.1.2 C4.5

C4.5 é um sucessor do ID3, também desenvolvido por Quinlan Ross. Esse algoritmo supera certas restrições do anterior, sendo capaz de operar com atributos contínuos a partir de uma discretização interna, na qual as divisões são feitas utilizando segmentos do atributo. Outra vantagem da C4.5 reside em poder utilizar amostras com atributos faltantes, diferente de sua antecessora. Além disso, esse algoritmo também apresenta a possibilidade da poda, em que são descartados nós mais próximos das folhas com o objetivo de reduzir o *overfitting* e, assim, aumentar o poder de generalização.

4.1.3 CART

O algoritmo CART apresenta muitas das vantagens já presentes na C4.5, no entanto, gera uma árvore com divisões binárias e que trabalha com valores contínuos diretamente,

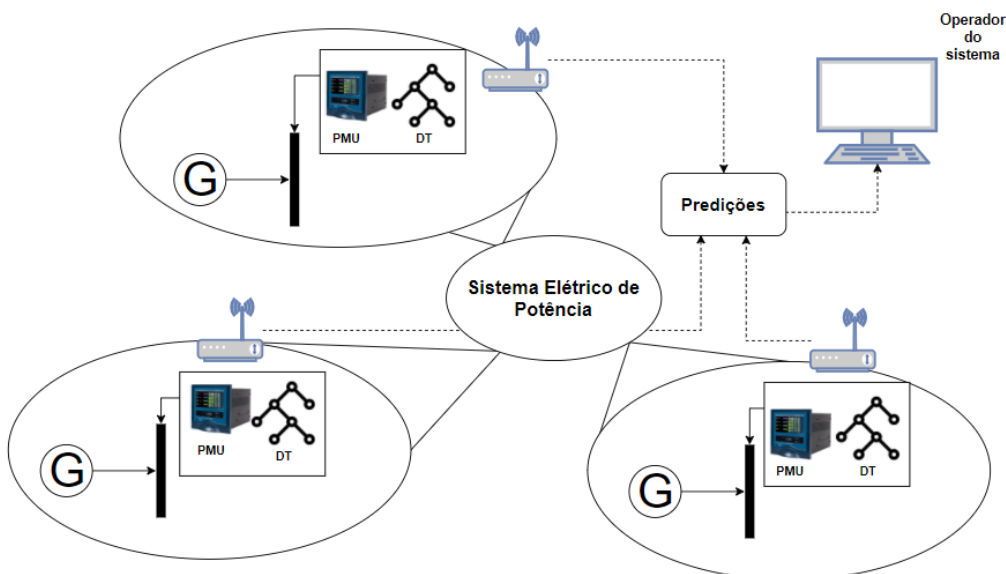
sem a necessidade de segmentação. Assim como as árvores apresentadas anteriormente, pode sofrer *overfitting*, porém esse problema pode ser mitigado pela poda ou redução da dimensão do vetor de características. Diferente da C4.5 e ID3, em sua forma base CART usa o critério de Gini para executar as divisões nos seus nós.

Dentre os diferentes tipos de algoritmos, o CART é um dos mais utilizados para construção de uma árvore de decisão (WILCOX, 2016). As principais vantagens desse algoritmo em relação aos demais consiste em suas habilidades em realizar classificações a partir de dados numéricos e realizar regressões. A partir de um conjunto de valores numéricos, o algoritmo constrói árvores binárias usando a característica e o limiar que trazer maior ganho de informação, em cada nó. Sendo assim, CART será o tipo de árvore utilizada nessa proposta.

4.2 Método 1: Classificação da estabilidade angular do rotor a pequenas perturbações

A metodologia proposta para classificação da estabilidade angular a pequenas perturbações se baseia em uma estrutura descentralizada de árvores de decisão, as quais encontram-se distribuídas nas PMUs alocadas junto aos geradores, conforme mostrado na Figura 7. Uma vez treinado, o algoritmo da árvore de decisão não requer um alto esforço computacional para classificação, o que torna possível em aplicações futuras embarcar cada árvore individual na própria PMU.

Figura 7 – Abordagem descentralizada proposta para análise da estabilidade a pequenas perturbações



Fonte: Autoria própria.

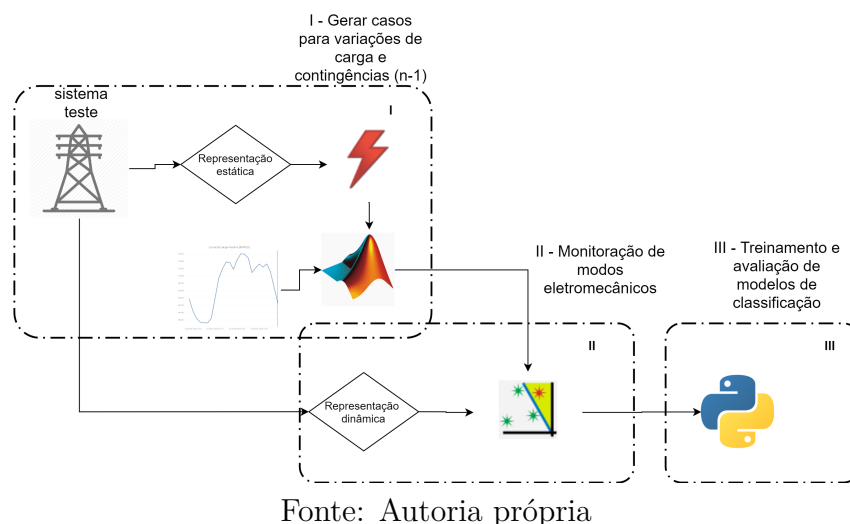
É importante mencionar que cada DT utiliza as medições referentes ao seu gerador como entrada para classificar a estabilidade angular da rede de transmissão. Assim, todas as classificações são enviadas ao operador do sistema e conseqüentemente, este pode avaliar a estabilidade a partir de qualquer uma das predições individuais.

A proposição de uma abordagem descentralizada é um diferencial dessa metodologia, uma vez que não é necessário que todas as árvores estejam disponíveis para que a análise de segurança possa ser realizada. Portanto, o operador do sistema de transmissão poderá avaliar o amortecimento das oscilações eletromecânicas, independente da perda de informação de algum gerador em específico ou de um *link* de comunicação, sendo essa uma característica altamente desejável para o SEP.

A primeira parte deste trabalho, consiste na elaboração e avaliação de classificadores de estabilidade angular do rotor a pequenas perturbações, a partir de árvores de decisão distribuídas pelos geradores do sistema teste e uma árvore centralizada recebendo dados de todos os geradores.

A Figura 8 apresenta um diagrama de blocos que representa a execução da metodologia proposta, desde a criação do banco, a partir da implementação do sistema teste nos programas do Centro de Pesquisa em Engenharia Elétrica (CEPEL), até os treinamentos e testes executados em Python.

Figura 8 – Panorama do método proposto para pequenas perturbações.



O procedimento é dividido em três etapas principais, em que as etapas I e II constituem a formação do banco de dados. Na primeira etapa é realizado o processamento e preparação de todos os casos, incluindo as variações randômicas de carga e uma lista de contingências. Já na segunda etapa, os modos referentes aos casos são monitorados e armazenados em um formato adequado para o treinamento do algoritmo, tal processo é realizado pelo *software* PacDyn. Na etapa final, as variáveis de entrada do classificador e saídas armazenadas são utilizadas para o treinamento do algoritmo, testes e obtenção da classificação do

amortecimento dos modos eletromecânicos.

4.2.1 Formação do banco de dados e Monitoramento do Modos Eletromecânicos

O primeiro estágio da metodologia proposta consiste na formação/preparação de um banco de dados que contenha a relação entre os pontos de operação de um sistema em estudo e as propriedades modais das oscilações eletromecânicas presentes nesta rede. É primordial que esse banco de dados contenha uma alta diversidade de casos, contemplando comportamentos distintos dos modos eletromecânicos no sistema em estudo.

Pelo histórico de operação de um determinado SEP, é possível definir uma curva diária de demanda de carga, com 24 pontos de operação, a qual permite identificar os períodos em que o sistema opera com carga máxima e fora da faixa de pico. Com base nesta curva diária, são estabelecidos 24 pontos de operação de base (POB) para o sistema teste.

A partir dos POB, novos cenários são definidos com intuito de representar as pequenas flutuações de demanda de carga que o sistema está sujeito em sua operação diária. Para cada ponto de operação base, a potência total consumida pela carga é alterada em relação ao POB por uma porcentagem sorteada no intervalo entre +1 % e -1 %, fornecido por uma distribuição gaussiana. Durante esse processo, é verificado se a variação de carga produzida já foi utilizada, em caso afirmativo, uma nova variação de carga é gerada até que se tenha um valor completamente novo. Dessa forma, garante-se que os casos gerados serão compostos por variações de carga únicas, em porcentagem.

Para cada k -ésimo novo ponto de operação, o resdepacho da potência ativa (P_{gi}) e reativa (Q_{gi}) injetada por cada i -ésimo gerador é realizado, conforme descrito nas equações abaixo:

$$P_{gi}(k) = P_{g0i}(k) + fp_i \Delta P_c(k), \quad (23)$$

$$Q_{gi}(k) = Q_{g0i}(k) + fp_i \Delta Q_c(k), \quad (24)$$

em que $\Delta P_c(k)$ e $\Delta Q_c(k)$ referem-se a variação da potência ativa e reativa da carga em relação ao ponto de operação base, respectivamente. P_{g0} e Q_{g0} são a potência ativa e reativa injetada pelo i -ésimo gerador em relação ao caso base. A participação de cada barra de geração na contribuição da potência ativa e reativa é definida pelo fator de participação dos geradores para o despacho ativo fp_i , o qual pode ser definido de acordo com o critério do TSO. Neste trabalho, tais fatores são calculados pela constante de inércia (H) de cada i -ésima máquina das ng do sistema, conforme descrito em (25).

$$fp_i = \frac{H_i}{\sum_{j=1}^{ng} H_j}, \quad (25)$$

Além do conjunto de dados da topologia base, um novo conjunto é gerado a partir da aplicação de uma lista de contingências sobre o sistema teste, estabelecidas de acordo com o critério $n-1$. Nesta lista, não são aplicadas faltas nas linhas que provocam ilhamento da unidade geradora. Para a formação desse conjunto, cada contingência é exposta ao mesmo processo de variação de carga aplicado à topologia base.

O fluxo de potência de cada um dos pontos de operação gerado é calculado no *software* ANAREDE. A partir dos relatórios de saída desse programa, os atributos (os quais irão compor as variáveis de entrada da árvore de decisão) são selecionados. No método proposto, são extraídos dos relatórios apenas os atributos que podem ser mensurados em campo por PMUs e que estão associados à barra de geração. Para cada ponto de operação, armazena-se as características (*features*) advindas do fluxo de potência que são constituídas pelo vetor,

$$[V_i \quad \theta_i \quad P_{gi} \quad Q_{gi} \quad P_{ci} \quad Q_{ci}], \quad (26)$$

em que V_i é a amplitude da tensão na barra monitorada, θ_i o ângulo da barra, P_{gi} a potência ativa de geração na barra i , Q_{gi} a potência reativa de geração na barra i , P_{ci} a potência ativa da carga na barra i , Q_{ci} a potência reativa de carga na barra. O alvo é apenas a classe estável ou instável, porém, para definir a classe é necessário que se armazene também características monitoradas sobre os modos correspondentes a cada caso. Com o auxílio da ferramenta *Monitoring of Oscillations* do *software* PacDyn, para cada caso são selecionados os modos interáreas que apresentam parte real próxima do eixo imaginário e, portanto, risco de transição para o semi-plano direito, resultando na instabilização do sistema. Os modos selecionados são monitorados para uma série de casos, sendo que a parte real e imaginária dos casos é registrada e, posteriormente, processada para classificar a estabilidade do modo.

A última etapa para a confecção da base de dados consiste no tratamento dos dados gerados pelo PacDyn e ANAREDE. Retira-se dos relatórios de fluxo de potência as variáveis apresentadas em (26) que serão usadas como entradas do algoritmo de aprendizagem e, a partir do relatório de acompanhamento de modos, são calculados os amortecimentos e, portanto, classificados os casos.

O amortecimento do modo é calculado a partir de sua componente real e imaginária pela equação (27), já apresentada no capítulo 2 em que σ_k representa a componente real e ω_k a componente imaginário do modo monitorado. A partir do amortecimento a estabilidade da amostra é classificada e armazenada, para, posteriormente, ser usada como saída na etapa de treinamento.

$$\zeta_k = \frac{-\sigma_k}{\sqrt{\sigma_k^2 + \omega_k^2}} \cdot 100 \quad (27)$$

Com o banco de dados devidamente processado, inicia-se a etapa de treinamento e testes das árvores de decisão.

4.3 Treinamento e testes com árvores de decisão

Neste trabalho, as árvores de decisão distribuídas são construídas de acordo com o algoritmo CART. Para o treinamento de cada árvore individual, o algoritmo 1 é utilizado.

Algoritmo 1 Treinamento de DTs distribuídas (Treinamento Offline).

```

Gerar casos (simulação ou medidas históricas de operação)
Avaliar e rotular cada caso como estável ou instável
for Barra_de_geração in Barras_de_geração do
    Extrair medições da Barra_de_geração para cada caso
    Treinar e validar DT relacionada a Barra_de_geração
end for

```

Depois desse processo, a estratégia com DTs distribuídas pode ser usada para avaliar a estabilidade a pequenas perturbações. Vale ressaltar que caso o sistema sofra mudanças significativas de topologia, um novo treinamento será necessário para atualizar os modelos.

As árvores são desenvolvidas em Python utilizando o pacote scikit-learn (GERON, 2017). Na etapa de treinamento, cada i -ésima DT recebe do *dataset* o vetor de *features* associado ao seu respectivo gerador. Para cada j -ésima amostra (caso), o vetor numérico é representado como $[V_i, \theta_i, P_{gi}, Q_{gi}, P_{li}, Q_{li}, status]$, onde *status* indica se o sistema é estável ou instável para o ponto de operação e topologia em análise. Lembrando que o status para cada caso é obtido através de análise modal e do cálculo de fator de amortecimento.

Nota-se que cada DT requer um número reduzido de entradas (6), uma vez que depende apenas de medidas locais de sua PMU. Então, a dimensão do vetor de entrada é independente do tamanho do sistema teste, o que é uma característica vantajosa. Nessa etapa, a quantidade reduzida de entradas de cada DT não justificam o uso de seleção de *features*, podendo-se relegar a seleção ao processo intrínseco das árvores (HUANG et al., 2016), que escolhem as *features* com maior relevância a partir do critério de impureza (Comumente, Gini ou entropia), durante o treinamento. Além disso, a escolha de um valor de profundidade baixo como critério de parada da DT, auxilia na interpretabilidade e, também mitiga a possibilidade de *overfitting*.

Com o objetivo de quantificar o desempenho do classificador de forma ampla e investigar possíveis melhorias, as métricas adotadas na execução deste trabalho são: precisão, *recall* e *f1-score*. O procedimento para cálculo de cada uma dessas métricas é descrito na subseção abaixo.

4.3.1 Métricas para análise de desempenho

Todas as métricas que serão utilizadas derivam e descrevem aspectos distintos da matriz de confusão, apresentada na Tabela 1. Nesta matriz, a coluna *atual* refere-se aos valores reais da amostra, já o termo *predito* corresponde aos valores de classe preditos pelo classificador.

Tabela 1 – Matriz de Confusão

Atual	Predito	
	Positivo	Negativo
Positivo	VP	FN
Negativo	FP	VN

Os verdadeiros positivos (VP) são as amostras que pertencem a classe positiva e que foram preditas como positivas pelo classificador. Os falsos negativos (FN) são as amostras que pertencem a classe positiva, mas foram classificadas como negativas. Os falsos positivos (FP) são amostras da classe negativa que foram classificadas como positivas e os verdadeiros negativos (VN) são amostras negativas que foram preditas como negativas. Vale ressaltar que, neste trabalho, a classe positiva será considerada estável e a classe negativa será atribuída aos casos instáveis. Essa configuração tem impacto direto na forma com que se aplicam as métricas propostas.

A precisão avalia a taxa de acertos do classificador, quantificando a quantidade de amostras que foram corretamente classificadas como positivas. Essa métrica é calculada pela seguinte expressão:

$$precisão = \frac{VP}{VP + FP} \quad (28)$$

O *recall* é a taxa de cobertura de classificação em relação às amostras reais da classe positiva e pode ser escrito como (29)

$$recall = \frac{VP}{VP + FN} \quad (29)$$

O índice *f1-score* corresponde à média harmônica entre as duas métricas anteriores: *recall* e precisão. Para o seu cálculo, a seguinte expressão é utilizada:

$$f1-score = 2 \cdot \frac{precisão \cdot recall}{precisão + recall} \quad (30)$$

Com o *f1-score* se avaliará o desempenho geral do classificador. Já com as métricas de precisão e *recall* se investigam possíveis problemas na classificação.

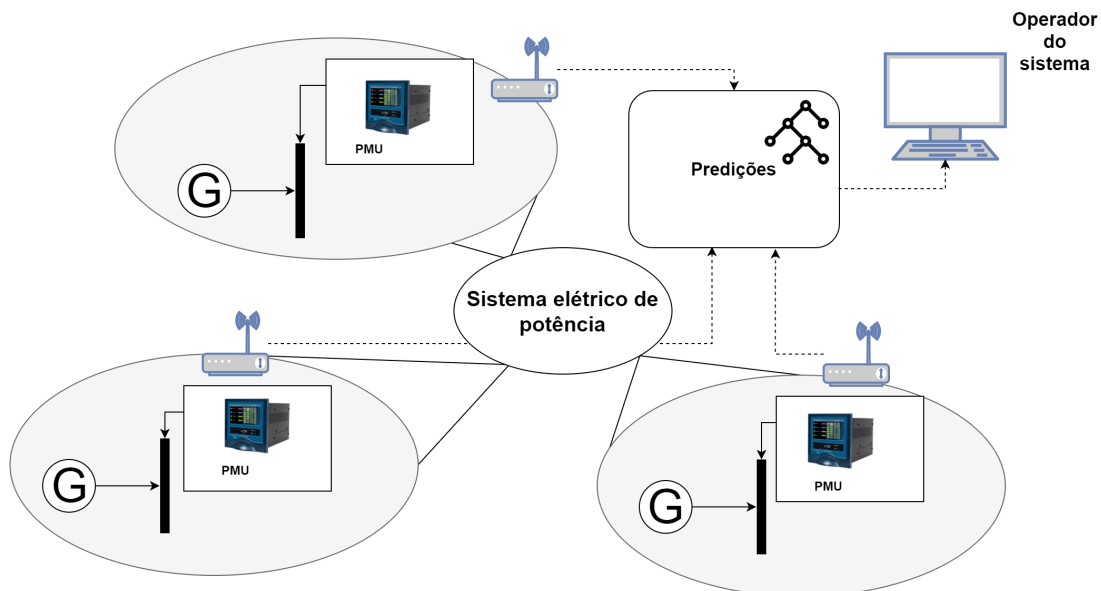
A acurácia é uma métrica mais genérica, quando comparada com as previamente citadas. Sua função é avaliar a taxa de acertos de todas as classes. No caso binário, a seguinte expressão é utilizada:

$$\text{acurácia} = \frac{VP + VN}{VP + FP + VN + FN} \quad (31)$$

4.4 Método 2: Classificação da Estabilidade Angular do Rotor a Grandes Perturbações

O segundo método proposto continuará fazendo uso das medições realizadas nas barras de geração, com o intuito de classificar a estabilidade. Para os casos em que o sistema é exposto a grandes perturbações, as árvores descentralizadas podem não ser capazes de realizar a análise e os dados medidos por suas PMUs devem ser enviados para uma árvore centralizada, responsável pela avaliação da estabilidade transitória. Seu papel, portanto, será complementar ao primeiro método, fornecendo informações sobre a estabilidade do sistema quando o mesmo é submetido a uma grande perturbação. Na Figura 9 pode-se observar o funcionamento da proposta, em que a árvore de decisão centralizada recebe medições de todas as barra de geração que compõe o SEP.

Figura 9 – Abordagem centralizada proposta para análise da estabilidade transitória



Fonte: Autoria própria

Os dados de entrada da DT são as medições advindas de todas as barras de geração do sistema teste. Diferente do método 1, as medições são acumuladas criando séries históricas dos períodos pós-falta. Isso permite que neste segundo método sejam exploradas combinações de dados das séries para compor o vetor de entrada do classificador, com

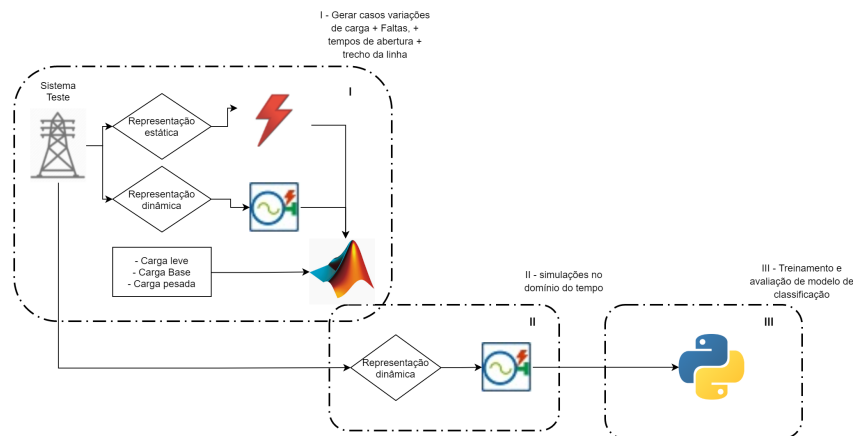
o objetivo de utilizar o mínimo de informação possível. Além disso, é investigada a importância dos dados mensurados nas barras para classificação do modelo, de modo que o conjunto de atributos essenciais ao classificador seja identificado.

A redução do vetor de entrada tem como objetivo reduzir a complexidade do modelo de classificação proposto, reduzindo assim o tempo de processamento do mesmo e a necessidade de grandes quantidades de amostras para a etapa de treinamento. Além disso, um vetor reduzido que utilize menos PMUs pode acarretar em economia na implementação da proposta, sem que se perca qualidade na predição.

A Figura 10 apresenta o fluxo do método 2. As etapas I e II representam a formação do banco de dados, onde se utiliza três patamares de carga: leve, base e pesada. Para cada patamar são utilizadas combinações de trechos da linhas e tempos de abertura para gerar as faltas que comporão as amostras. É válido ressaltar que não serão exploradas mudanças de topologia, ou seja, após o tempo de abertura considera-se que o sistema volta a operar com sua topologia original e sem perdas de linhas.

Utiliza-se, novamente o Matlab e o ANAREDE e, acrescenta-se o ANATEM para as simulações no domínio do tempo. Com o banco de dados montado a etapa III é muito semelhante a realizada no método anterior, utilizando *Python* para o treinamento e teste, porém foca-se em explorar diferenças na abordagem mantendo-se os mesmos cenários de teste, ao contrário do método 1.

Figura 10 – Panorama do método proposto para estabilidade transitória.



Fonte: Autoria própria

A seguir, explica-se com maior detalhamento o fluxo apresentado, iniciando-se pela formação do banco de dados.

4.4.1 Formação do Banco de Dados

Para a formação do banco de dados, assim como na investigação da estabilidade a pequenas perturbações, teve-se como objetivo produzir um conjunto de dados rico em

diversidade e com alta quantidade de amostras de todas as classes, para que o classificador consiga adquirir um alto poder de generalização. Dessa forma, gerou-se casos com variações de carga, bem como aplicações de falta com diferentes tempos de abertura, explorando a resposta dinâmica do sistema. Para tanto, foi utilizado um procedimento semelhante ao adotado no método 1.

A variação de carga abrange 3 casos distintos que incluem carga leve, base e pesada. Além da componente estática dos cenários, o Matlab também é utilizado para produzir faltas (curto-circuito) a partir das combinações de tempos de abertura, trechos da linha afetada e linhas de transmissão em que a falta é aplicada. Todos os casos tem uma falta aplicada em 1 segundo após o início da simulação.

Com os cartões do ANAREDE e ANATEM preparados pelo passo anterior, faz-se uso da simulação em lote do ANATEM para a análise dos casos no domínio do tempo. O ANATEM é responsável por armazenar as medições de ângulo e magnitude de tensão das barras geradoras, bem como os ângulos dos rotores durante toda a operação. Com os relatórios produzidos pelo ANATEM, novamente utiliza-se o Matlab para separar os dados de entrada e calcular-se o índice de estabilidade.

As séries referentes às medições de PMU são separadas por ciclo de medição, os quais são coletados no período pós-falta da resposta do sistema a uma perturbação. Um período de amostragem de 20 ms é utilizada para realizar as medições, portanto, uma taxa compatível com o de uma PMU¹. Tomando como base o estudo realizado em (YU et al., 2018), oito ciclos de medição no período pós-falta são realizadas.

Para ilustrar o processo de medição, na Figura 11 observa-se a medição da magnitude de tensão de uma barra de geração para um caso específico, quando uma falta trifásica é aplicada ao sistema teste, com duração de 400 ms (tempo de abertura de 1,4 s). As cruzes representam os pontos em que houve medição, com um período de amostragem de 20 ms. Os oito ciclos de medição, que compõe a série temporal da variável são armazenados.

Cada um dos 8 ciclos de medição é armazenado em uma planilha distinta. Dessa forma, tem-se total controle de a qual ciclo as entradas pertencem, bem como tempo de abertura utilizado e variável de entrada armazenada, magnitude e ângulo da tensão.

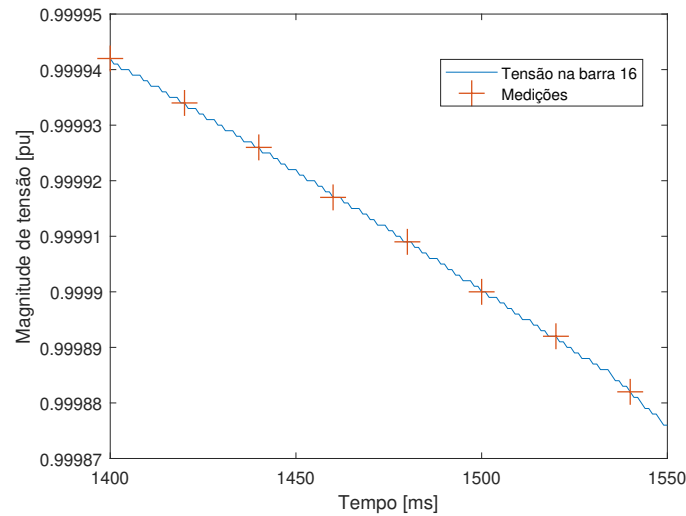
Para a definição de estabilidade utiliza-se o índice de estabilidade transitória baseado em diferença angular do rotor - TSI, o qual já foi apresentado no capítulo 2. Os ângulos dos rotores advindos dos relatórios do ANATEM são utilizados para o cálculo dos índices.

No TSI avalia-se a diferença máxima entre quaisquer dois geradores do sistema, segundo a fórmula (32):

$$TSI = \frac{360 - \delta_{max}}{360 + \delta_{max}} \quad (32)$$

¹ As PMUs possuem período de amostragem, entre 8,33 e 100 ms (ZHOU et al., 2015), aproximadamente, sendo que em grande parte as PMUs instaladas tem uma taxa de amostragem de 60 ou 50 Hz.

Figura 11 – Ciclos de medição do pós falta



Fonte: Autoria própria

em que δ_{max} é a diferença máxima entre quaisquer dois geradores do sistema, medido no período pós-falta. Quando $TSI > 0$, o sistema é estável, caso contrário, o sistema é instável. Em outros termos, se $\delta_{max} > 360$, o sistema é considerado instável.

Para que fique mais clara a forma de cálculo a partir desse índice e a influência da quantidade de ciclos de amostras para o seu cálculo, considere um sistema composto por 16 geradores síncronos, no qual aplica-se uma falta trifásica no instante $t = 1s$ com duração de 550 ms. A resposta do ângulo do rotor dos geradores conectado no período em falta e pós-falta são exibidas na Figura 12 bem como do valor do índice TSI em cada instante de tempo. Nota-se que no primeiro ciclo de medição o TSI ainda apresenta um valor de 0,14, enquanto que no oitavo ciclo, um TSI de -0,05 é calculado, permitindo afirmar com maior segurança que há instabilidade no sistema. Portanto, quanto maior o intervalo de tempo após a remoção da falta, mais evidente é o comportamento da rede na simulação no domínio do tempo e a partir do valor do TSI.

A partir dos índices de estabilidade se definem as classes do problema sendo estável se o índice é maior que 0 e instável em caso contrário. Por fim, os dados de saída desejada são armazenados em planilhas distintas por tempo de abertura. Os ângulos dos rotores advindos dos relatórios do ANATEM são utilizados para o cálculo dos índices.

4.4.2 Seleção de *features*

Antes de iniciar o treinamento da árvore de decisão, são realizadas duas investigações para avaliar a importância dos atributos que compõem o vetor de entrada, a saber: avaliação da importância de *features* individuais e de *features* combinadas.

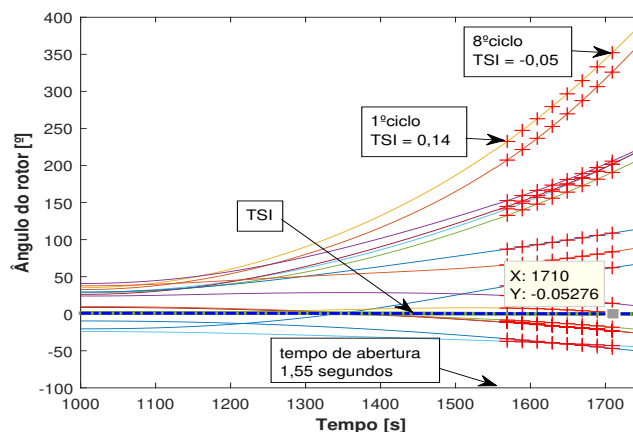
Na análise dos atributos individuais, se avalia se há um determinado tipo de variável do

vetor de entrada que apresenta um impacto mais significativo na classificação. Portanto, no método proposto, essa análise permite investigar se as medidas da magnitude da tensão nas barras são mais significativas que as do ângulo para realizar a classificação, ou vice-versa. Já a análise de importância combinada busca definir grupos de variáveis com maior potencial para classificação. Neste trabalho, essa análise permite identificar entre as barras de geração quais são as mais significativas para o modelo de predição, reduzindo a quantidade de PMUs necessárias para o monitoramento de informação.

Para definição do grau de relevância na análise individual, se utiliza o teste estatístico F (ELSSIED; IBRAHIM; OSMAN, 2014). Neste teste, se a amostra (variável) apresenta variância semelhante em duas classes, tem-se como resultado um valor baixo para F , caso contrário esse valor é alto. Na análise individual de uma *feature* a partir dessa análise, admiti-se que caso a variância da variável em análise seja semelhante nas classes estáveis e instáveis, a entrada não é significativa para a classificação. Em contrapartida, quando uma *feature* é relevante, suas respectivas variâncias entre as classes se distanciam, resultando em um alto valor para F . Além da estatística F , também é necessário que se avalie um valor, denominado valor p da distribuição (ELSSIED; IBRAHIM; OSMAN, 2014). Quanto menor o valor p , maior a confiança para se rejeitar a hipótese nula, a qual afirma que a variância das amostras são semelhantes nas duas classes.

Na análise de *features* combinadas, utiliza-se um tipo de árvore de decisão, a *Random Forest* para explorar quais as entradas mais relevantes para a tarefa de classificação. Nesse sentido, emprega-se o processo de seleção de *features* realizado pela *Random Forest*, intrínseco às árvores de decisão para avaliar as *features* mais relevantes. No método proposto, essa análise é realizada por ciclos de medição, o que permite elaborar uma matriz que quantifique a importância de cada *feature*, magnitude e ângulo da tensão da barra de geração, pela quantidade de ciclos de medição que são considerados no vetor de entrada do modelo de predição.

Figura 12 – Ângulos dos rotores no pós-falta



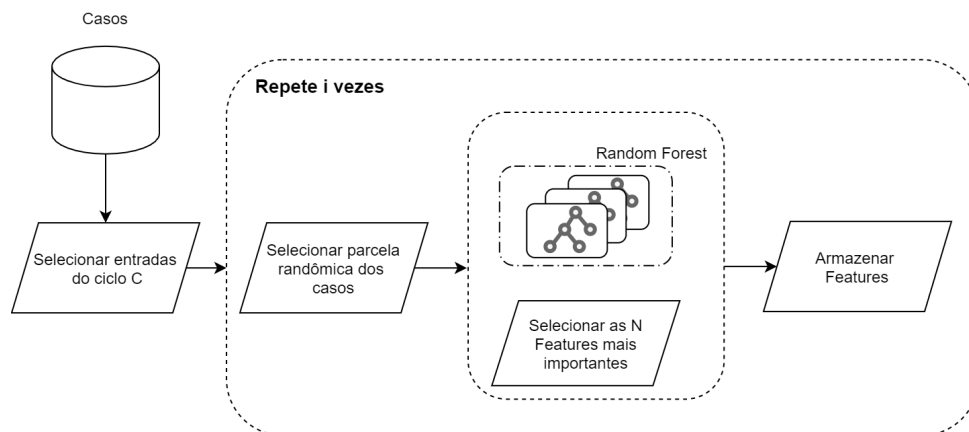
Fonte: Autoria própria

A Figura 13 ilustra o procedimento de seleção de *features* para a análise combinada de atributos. Para cada ciclo C são realizados i treinamentos, em que realiza-se um sorteio formando um novo grupo de amostras a cada iteração. Ao final de cada treinamento, extrai-se do *rank* interno do modelo as 10 *features* mais relevantes, armazenando essa informação. Findadas as iterações, tem-se um vetor de frequências de cada *feature* entre as 10 mais relevantes do ciclo C . De tal forma, que quando se executa o processo inteiro, a combinação de cada um desses vetores produz uma matriz, apresentada na Tabela 2, em que cada linha apresenta as frequências das *features* e cada coluna representa um ciclo.

Tabela 2 – Matriz de frequência de *features* mais importantes por ciclo

	ciclo 1	ciclo 2	...	ciclo n
$feature_1$	c_{1_1}	c_{1_2}	...	c_{1_n}
$feature_2$	c_{2_1}	c_{2_2}	...	c_{2_n}
\vdots	\vdots	\vdots	\vdots	\vdots
$feature_n$	c_{n_1}	c_{n_2}	...	c_{n_n}

Figura 13 – Fluxograma de seleção de features



Fonte: autoria própria

A partir da matriz de importâncias são selecionados subconjuntos do vetor de medições, removendo-se *features* menos importantes, uma a uma, para se analisar o impacto da redução de medições no desempenho do modelo. Dessa forma, busca-se identificar um conjunto mínimo de PMUs, a partir do qual não se observa ganho significativo de desempenho caso se acrescente mais medições.

4.4.3 Treinamento e Validação da Árvore de decisão

Com o banco de dados montado, contendo as medições que são utilizadas para compor as entradas, e a classe de estabilidade baseada no índice de instabilidade, é executado o treinamento da árvore de decisão. Essa etapa é semelhante à executada na análise

da estabilidade a pequenas perturbações. Porém, diferente do método 1 que fixa uma abordagem e avalia diferentes cenários, no método 2 o foco é dado a diferentes abordagens para o mesmo cenário.

Inicialmente, é realizada uma etapa de seleção de *features*, conforme descrito na subseção 4.4.2, com o objetivo de definir conjuntos distintos de vetores de entrada para investigar o comportamento do classificador. Em sequência, um pré-processamento importante é realizado, antes de se iniciar os treinamentos e teste, o balanceamento de classes do *dataset*.

O balanceamento de classes busca mitigar um problema comum em árvores de decisão que reside na sua sensibilidade a variações ou viés do *dataset*. Em especial, um conjunto de amostras com alto grau de desbalanceamento em relação às classes utilizadas no treinamento de um classificador baseado em árvore de decisão pode ocasionar uma classificação que tenda, na maioria das vezes, à classe dominante presente no treinamento. Visando mitigar essa potencial dificuldade, emprega-se o método *undersampling*. O *undersampling* consiste na remoção aleatória de amostras da classe dominante do conjunto de dados de treinamento, de forma que o *dataset* final apresente maior balanceamento entre as classes.

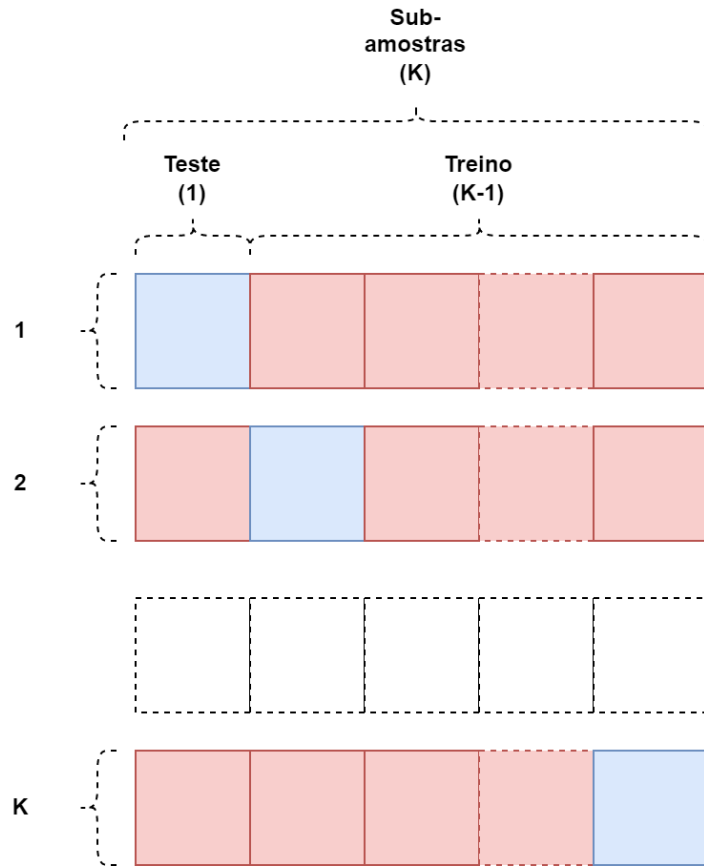
Com o banco devidamente balanceado, para verificar o desempenho do classificador, são utilizadas, novamente, a acurácia e o *f1-score* como métricas de avaliação (definidas na subseção 4.3.1), porém também é empregado o método *k-folds* para promover a avaliação dos modelos em todo o *dataset*.

O método *k-folds* é uma técnica de validação cruzada, cujo objetivo consiste em utilizar todo o *dataset* disponível como teste e, assim, avaliar a capacidade de generalização do classificador. Nessa técnica, como representado na Figura 14, as amostras são divididas em *k* subconjuntos podendo haver escolha aleatória dos elementos ou mantendo-se a ordem das amostras. Durante *k* iterações, um subconjunto é reservado para teste, enquanto os outros *k-1* subconjuntos são utilizados para o treinamento do modelo. Ao final do processo, se obtém o desempenho médio do classificador em todas as amostras disponíveis.

Por fim, um último passo antes dos treinos serem iniciados, é a definição do hiperparâmetro profundidade da DT. Realiza-se uma breve análise avaliando o desempenho da árvore para diferentes profundidades, com o objetivo de definir uma profundidade mínima, que é empregada nos demais testes.

Como são armazenados 8 ciclos de medição das variáveis de magnitude e ângulo da tensão das barras de geração, é possível avaliar como combinações distintas do vetor de entrada podem interferir na acurácia atingida pela árvore centralizada proposta. Com intuito, de verificar esse comportamento, as subseções abaixo detalham cada um dos testes realizados.

Figura 14 – Método K-folds



Fonte: autoria própria

4.4.4 Relação entre desempenho e a quantidade de ciclos de medição

O principal objetivo do segundo método reside na avaliação da capacidade de predição do classificador quando se reduz a quantidade de ciclos de medição do pós-falta. Essa análise possibilita avaliar o ganho do classificador com o aumento de ciclos de medição, definindo um *tradeoff* entre qualidade de classificação e velocidade.

Dessa forma, inicia-se as investigações com uma árvore central, a qual é alimentada com um tipo de variável de entrada, a mais significativa entre magnitude ou ângulo da tensão. Para essa definição, a seleção de *features* individuais também é aplicada.

Uma vez definida a *feature* a ser avaliada, varia-se a quantidade de ciclos de medição que compõe o vetor de entrada do modelo. Supondo que a *feature* mais significativa seja magnitude da tensão das barras de geração, o formato do vetor de entrada para cada ciclo avaliado é dado por:

$$[V_{1_1}, V_{2_1}, V_{3_1}, \dots, V_{nb_nc}] \quad (33)$$

Nesse vetor V_{i_j} representa a medida da magnitude de tensão na barra de geração i e

no ciclo de medição j , para cada uma nb barras geração do sistema teste; nc representa os ciclos de medição.

Para comparar o desempenho da classificação fornecida pela árvore de decisão com o de outras técnicas de aprendizagem de máquina, também aplica-se a técnica SVM, a qual é treinada seguindo os mesmos critérios estabelecidos para árvore centralizada.

A partir da curva obtida nessa etapa, o operador do sistema poderá optar por uma quantidade de ciclos mínima que lhe forneça uma acurácia aceitável, agilizando o processo.

4.4.5 Classificadores distribuídos

Busca-se também avaliar a viabilidade de uma abordagem análoga à realizada para pequenas perturbações, com classificadores distribuídos pelas barras geradores e recebendo medições únicas de suas barras.

Inicia-se os testes, com o vetor de entrada de cada árvore distribuída sendo composto pelas medidas de magnitude e ângulo da tensão da sua respectiva barra de geração, conforme exibido:

$$[V_{i_1}, A_{i_1}, V_{i_2}, \dots, A_{i_nc}] \quad (34)$$

Nesse vetor V_{i_j} permanece a magnitude de tensão e A_{i_j} representa o ângulo de tensão, ambos da barra i no ciclo j . Dessa forma, cada classificador poderá receber a magnitude e ângulo da tensão corresponde da sua barra e diferentes quantidades de ciclos de medição.

Caso os resultados sejam promissores avaliar-se-á apenas o uso da magnitude de tensão.

4.4.6 Identificação do conjunto mínimo de medidores necessários

Para esse teste, repete-se o processo de seleção de *features* com *Random Forest* por ciclo de medição, com o objetivo de identificar barras são mais significativas para classificação e, assim, reduzir a quantidade de PMUs necessárias no sistema em estudo.

Após os grupos de barras selecionadas, são feitos testes em que se avalia o efeito da redução do número de barras medidas na performance do classificador. Análogo ao realizado na etapa de avaliação de ciclos mínimos, o objetivo é identificar uma quantidade mínima de barras que seja capaz de produzir uma acurácia aceitável.

Capítulo 5

Resultados - Classificação de Estabilidade Angular do Rotor a Pequenas Perturbações a partir de Árvores de Decisão Distribuídas

Neste capítulo são apresentados os testes realizados e os resultados obtidos com a aplicação do método 1, ou seja, para a classificação da estabilidade a partir da análise do fator de amortecimento dos modos eletromecânicos.

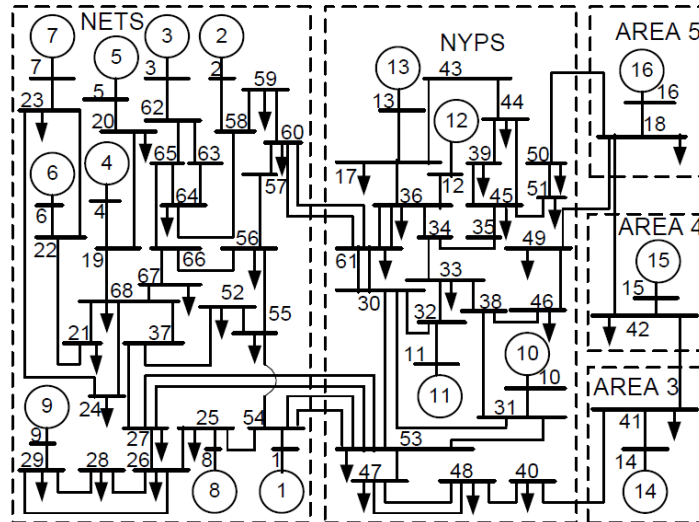
Para avaliar a eficiência do método proposto testes são conduzidos no sistema IEEE 68 barras (CANIZARES et al., 2017), o qual é amplamente utilizado na literatura para avaliação de estabilidade angular, como evidenciado no Capítulo 3. As principais características do sistema teste e do banco de dados gerado, considerando quatro cenários distintos são descritos a seguir. Além disso, os resultados obtidos para cada cenário são analisados e discutidos individualmente.

5.1 Sistema Teste

O sistema 68 barras, ilustrado na Figura 15, o qual é um modelo equivalente de ordem reduzida do NETS (em inglês, *New England Test System*) e do NYPS (em inglês, *New York Power System*). O sistema teste é composto por 5 regiões geográficas, sendo que as duas primeiras são equivalentes reduzidos do sistema *New England* (NETS) e *New York* (NYPS). As regiões 3, 4 e 5 são equivalentes de área. Os geradores 1 a 12 são equipados com regulador de tensão e estabilizador de potência, cujo modelo e parâmetros

são detalhados em (CANIZARES et al., 2017). Os geradores 14 e 16 também estão equipados com regulador de tensão, para que cenários instáveis sejam observados a partir da variação de carga.

Figura 15 – 68 barras NETS-NYPS



Fonte: (CANIZARES et al., 2017)

No caso base, o sistema tem três modos interárea, que serão rastreados, descritos na Tabela 3. Os três modos interárea descritos são monitorados durante a simulação pelo PacDyn. É importante ressaltar que o modo 1, de frequência 0,53 Hz é o que apresenta o menor fator de amortecimento, de apenas 2,38%, sendo assim esse modo será monitorado nas simulações. Esse modo caracteriza a oscilação da área 5 (gerador 16) contra as áreas 4 (gerador 15) e 3 (gerador 14).

Tabela 3 – Modos eletromecânicos do sistema teste.

Modo	$\sigma \pm j\omega$	Freq. (Hz)	$\zeta(\%)$
1	$-0,079 \pm j3,331$	0,5302	2,380
2	$-0,145 \pm j4,901$	0,7800	2,956
3	$-0,335 \pm j3,680$	0,5857	9,070

5.2 Banco de dados do sistema teste de 68 Barras

A eficácia do modelo será avaliada em 4 cenários distintos. O primeiro cenário é composto pelas variações de carga diária obtidas através da interpolação dos pontos de operação base, apresentados no Capítulo 4, as variações na parcela ativa e reativa da carga são idênticas e não há aplicação de faltas. O segundo cenário busca avaliar o modelo em

casos mais extremos, acrescentando ao primeiro cenário perdas de linhas (contingência $N-1$), desta forma avalia-se a capacidade de generalização do modelo quando há mudanças de topologia da rede. No terceiro cenário, busca-se avaliar a influência de variações na direção de crescimento das cargas no desempenho do classificador. Finalmente, o quarto cenário utiliza a mesma premissa de curva de carga com fatores de potência distintos e adiciona contingências $N-1$. Cada um dos cenários é apresentado nas subseções seguintes. Vale ressaltar que a curva de carga base só será apresentada no primeiro cenário, por se tratar da mesma para os demais.

5.2.1 Cenário 1

A partir de uma curva de carga diária, obtida através do site da ONS, 24 pontos de operação base foram estabelecidos para o sistema. A curva base é normalizada em torno da média, para que a mesma possa representar variações de carga em porcentagem.

Nesse primeiro cenário, a partir da interpolação dos 24 pontos de operação base, são gerados 5000 pontos de operação sobre a curva de carga para o sistema teste. Além disso, para cada ponto gerado acrescenta-se uma variação de potência ativa e reativa de $\pm 1\%$. Para cada novo ponto de operação gerado a partir da variação de potência ativa e reativa de cargas, o redespacho de potência dos geradores também é realizado.

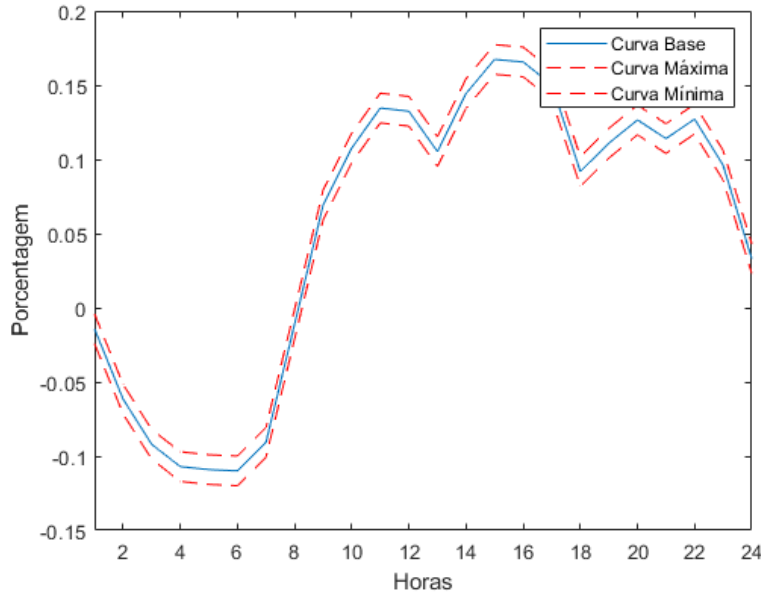
Para definição do carregamento máximo utiliza-se fluxo de potência continuado. Dessa forma, encontra-se 17,11% de variação de carga em relação à base como ponto máximo, enquanto que o carregamento mínimo é definido em $-11,46\%$. Na Figura 16 a curva base normalizada é apresentada em azul, os envelopes que definem sua variação máxima de $\pm 1\%$ são apresentados em vermelho.

Dos 5000 casos, 4498 apresentaram convergência e, portanto são válidos para a utilização. Deste total, em 966 casos o sistema é classificado como instável, enquanto o restante é definido como estável. A Figura 17 mostra o fator de amortecimento e a frequência do modo interárea 1, que é o único modo desse sistema que apresenta amortecimento negativo e, portanto será o modo rastreado nas simulações. Nota-se que o modo descreve uma trajetória bastante comportada, e seria possível estimar uma relação direta entre a frequência e o fator de amortecimento.

5.2.2 Cenário 2

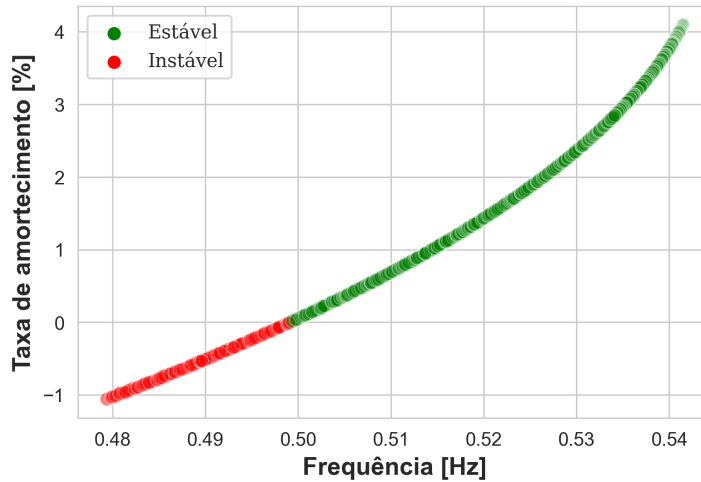
No segundo cenário avalia-se a aplicação de contingências $N-1$. Semelhante ao realizado em (MANSOUR; ALBERTO; RAMOS, 2016), contingências que causem ilhamento são desconsideradas na análise. Além disso, combinações entre faltas e variações de carga em que não há convergência do fluxo de potência também são descartadas. Dadas as considerações, este cenário é composto por uma lista 29 faltas possíveis. Aplica-se essa lista para valores distintos de carregamento do sistema com relação ao seu caso base,

Figura 16 – Variações de curva diária



Fonte: Autoria própria.

Figura 17 – Distribuição de casos - Cenário 1



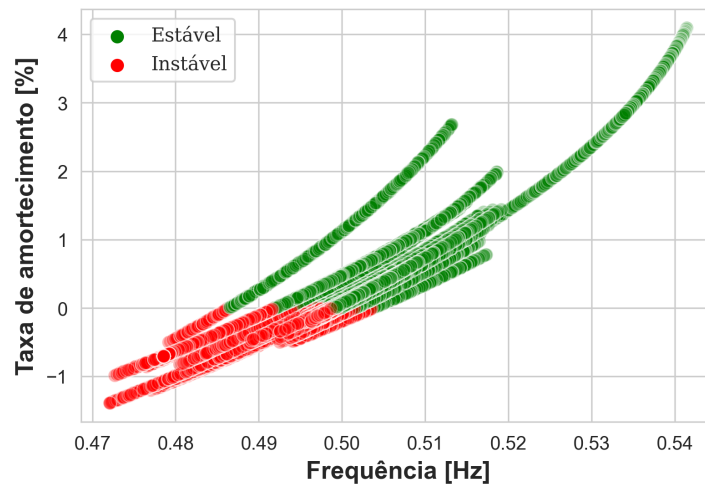
Fonte: Autoria própria

conforme apresentado na seção 4.2.1. Apesar da curva base ser a mesma, as variações de carga de $\pm 1\%$ (ruído randômico) geram pontos diferentes dos utilizados no cenário 1. Dessa forma, cada topologia apresenta diversos casos de variação de carga distintos.

No total, a distribuição de casos em cada classe é de 15.262 estáveis e 12.363 instáveis. Novamente, a trajetória do modo 1 para o semi-plano direito do plano complexo ocasiona casos de instabilidade, como representado na Figura 18. Porém, a proporção de casos instáveis nesse *dataset* é maior que no cenário 1. Além disso, como pode-se observar, o modo interárea 1 desloca-se sobre novos valores de fator de amortecimento em um mesmo

intervalo de frequência observado no cenário 1. Esse segundo *dataset* representa uma tarefa não trivial para os classificadores, uma vez que o comportamento do sistema teste e, especialmente, do modo interárea 1 se altera de forma significativa de acordo com cada contingência.

Figura 18 – Distribuição de casos - Cenário 2



Fonte: Autoria própria.

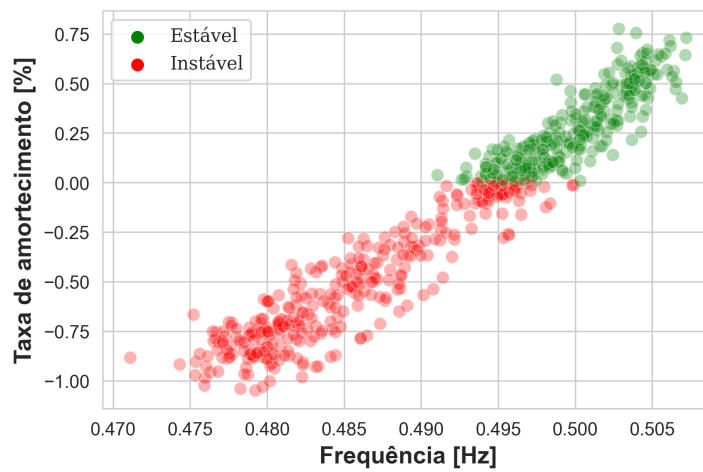
5.2.3 Cenário 3

Com o objetivo de investigar o impacto de diferentes direções de crescimento da carga nas propriedades modais da oscilação interárea, o terceiro, bem como o quarto cenário, são preparados considerando variações multidirecionais de crescimento de carga.

Nesses dois cenários, o valor mínimo de fator de potência é de 0,8 e o máximo é de 1,0. Além disso, visando obter casos na região de transição, onde o sistema teste opera próximo do limiar de mudança entre estabilidade e instabilidade, novos pontos de operação são gerados entre carregamento de 10,04% e 17,23%. O cenário 3 é composto de 330 casos instáveis e 269 casos estáveis.

Como pode-se observar na Figura 19 a relação entre frequência e fator de amortecimento não é tão nítida como no cenário 1 e a distribuição se torna mais dispersa. Além disso, as diferentes direções de vetores de crescimento de carga parecem contribuir com diferentes trajetórias do modo interárea analisado. Novamente, ressalta-se que o trecho percorrido da curva base é menor para este cenário, o que ocasiona menores variações de fator de amortecimento e frequência, porém tem-se casos concentrados em torno da região de transição e maior balanceamento entre o número de situações estáveis e instáveis.

Figura 19 – Distribuição de casos - cenário 3



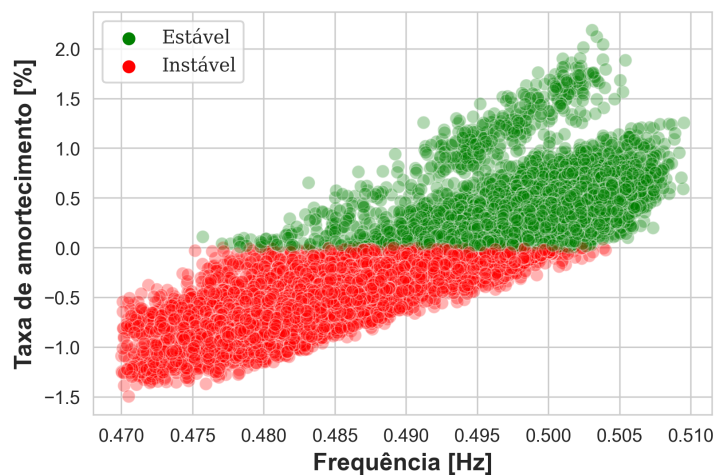
Fonte: Autoria própria.

5.2.4 Cenário 4

Finalmente, o cenário 4 é bastante parecido com o cenário 3. O procedimento de variação de cargas é o mesmo, porém acrescentam-se as mesmas 29 contingências abordadas no cenário 2. Para esse cenário produziu-se 17.971 casos válidos, em torno de 600 casos por contingência, sendo 8.108 estáveis e 9.863 instáveis.

Na Figura 20, novamente observa-se maior dispersão em relação aos cenários 1 e 2. O que sugere alta sensibilidade do modo à direção do vetor de crescimento de carga. O último cenário, como se observa pela distribuição, é o mais difícil para a classificação de estabilidade, não apresentando trajetórias tão claras como nos casos anteriores.

Figura 20 – Distribuição de casos - Cenário 4



Fonte: Autoria própria.

5.3 Desempenho dos classificadores

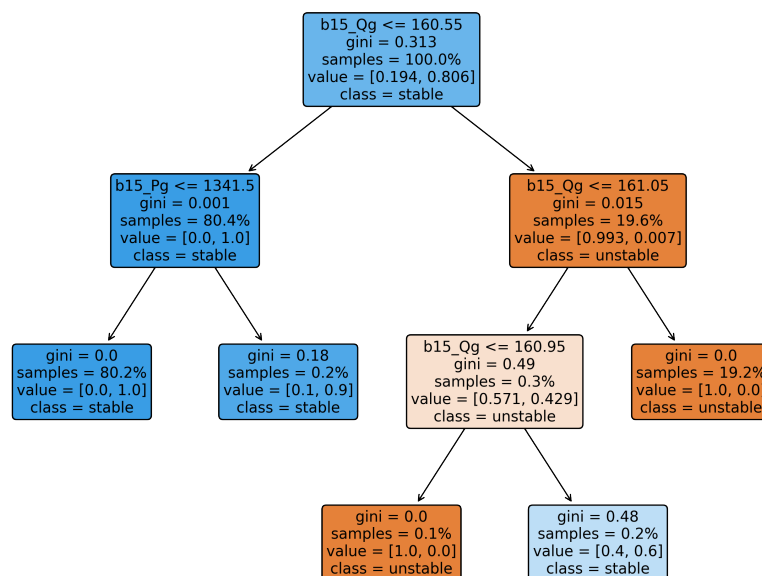
Para cada cenário 300 casos aleatórios foram separados para teste com uma composição de 1/5 dos casos próximos à região de transição e 4/5 distantes. Com esses conjuntos de dados analisou-se as performances dos classificadores distribuídos e do classificador centralizado, a partir da acurácia e do f1-score. É interessante notar que em todos os cenários foi realizado o balanceamento de classes nos *datasets* de teste, possivelmente, contribuindo para o f1-score e acurácia próximos, como ficará claro nas subseções a seguir.

5.3.1 Cenário 1 - Variações diária de carga e vetor de crescimento unidirecional

Neste cenário, foi usado o primeiro *dataset*, onde a eficiência do modelo proposto é avaliada para variações diárias de carga. Como mostrado na Figura 17, a taxa de amortecimento mínima alcançada pelo modo interárea 1 está próxima de $-1,0\%$, enquanto que a máxima está próxima de $4,0\%$.

Para ilustrar os resultados das DTS, a Figura 21 mostra os níveis da DT relacionada a barra geradora de número 15, com profundidade máxima de 4. Como pode-se observar, a potência ativa (P_g) e a potência reativa (Q_g) injetadas pelo gerador são *features* importantes para a classificação da estabilidade do sistema, aparecendo no nó raiz e outros nós internos. Além disso, pode-se notar que o índice de Gini ponderado dos nós filhos diminui quando comparado com seus nós pais, como esperado.

Figura 21 – Árvore de decisão da barra 15.



Fonte: Autoria própria.

O desempenho de cada DT distribuída comparado com uma DT centralizada é mostrado na Tabela 4, onde o desempenho médio das árvores distribuídas também é calculado. É possível notar os altos valores obtidos nas métricas de classificação (distribuídas e centralizada), onde a menor acurácia e f1-score obtidos são de 0,983. Sendo assim, usando apenas informação local, qualquer DT distribuída poderá prever a estabilidade sem perda significativa de desempenho, quando comparada à DT centralizada.

Tabela 4 – Performance no primeiro cenário

Árvore de decisão	Acurácia	F1-Score
Barra 01	0,990	0,990
Barra 02	0,990	0,990
Barra 03	0,993	0,993
Barra 04	0,990	0,990
Barra 05	0,983	0,983
Barra 06	1,000	1,000
Barra 07	0,990	0,990
Barra 08	0,990	0,990
Barra 09	0,990	0,990
Barra 10	0,990	0,990
Barra 11	0,990	0,990
Barra 12	0,990	0,990
Barra 13	0,990	0,990
Barra 14	0,996	0,996
Barra 15	0,996	0,996
Barra 16	1,000	1,000
Média DTs distribuídas	0,991	0,991
DT centralizada	1,000	1,000

5.3.2 Cenário 2 - Variações de carga, vetor de crescimento unidirecional e contingências

Apesar da dificuldade de avaliar esse segundo cenário, em que a classificação da estabilidade do sistema muda significativamente a depender da topologia, um f1-score acima de 0,95 foi obtido por todas as DTs distribuídas. O desempenho médio das DTs distribuídas também é muito próximo do alcançado pela árvore centralizada, em que são necessários dados de todos os geradores. Nota-se que, mesmo se houver perda de informação de um gerador específico, a classificação provida por qualquer das DTs distribuída remanescentes terá um alto desempenho.

5.3.3 Cenário 3 - Variações de carga e vetor de crescimento multidirecional

Neste terceiro cenário, o impacto da direção do crescimento de carga na eficiência da abordagem proposta é avaliado. Como mostra-se na Figura 19, esse *dataset* é composto

Tabela 5 – Performance no segundo cenário.

Árvore de decisão	Acurácia	F1-Score
Barra 01	0,976	0,977
Barra 02	0,956	0,958
Barra 03	0,970	0,970
Barra 04	0,963	0,964
Barra 05	0,976	0,977
Barra 06	0,960	0,961
Barra 07	0,960	0,961
Barra 08	0,976	0,977
Barra 09	0,960	0,961
Barra 10	0,966	0,967
Barra 11	0,976	0,977
Barra 12	0,960	0,961
Barra 13	0,976	0,977
Barra 14	0,960	0,961
Barra 15	0,950	0,952
Barra 16	0,960	0,961
Média DTs distribuídas	0,965	0,966
DT Centralizada	0,983	0,983

por casos com menor intervalo de taxa de amortecimento quando comparado aos cenários anteriores, com valores entre $-1,05\%$ e $0,77\%$. Dessa forma, o sistema opera próximo da região de transição, que é a fronteira entre estabilidade e instabilidade.

Apesar das variações multidirecionais na variação de carga, os resultados dos classificadores não são significativamente afetados, como pode-se observar na Tabela 6. A acurácia e f1-score de cada DT distribuída está acima de 0,96. Além disso, a acurácia média das DTs distribuídas é de 0,976, o que representa um valor próximo do obtido no primeiro cenário (0,991), onde se avaliou o crescimento de carga unidirecional.

Quando comparadas à DT centralizada, a acurácia média das DTs distribuídas apresenta uma diferença menor que 1%. Ainda nesse cenário, alta acurácia é alcançada pela proposta distribuída.

5.3.4 Cenário 4 - Variações de carga, vetor de crescimento multidirecional e contingências

Nesse cenário de maior dificuldade, além do sistema operar com variações de carga multidirecionais, também são consideradas contingências $N-1$. Mesmo nesse caso, as DTs distribuídas apresentam alto desempenho, alcançando uma acurácia média de 0,947 e f1-score de 0,95 acima dos obtidos pela DT centralizada (0,943 e 0,946, respectivamente). Apenas de medições locais, todas as DTs individuais alcançaram altos valores nas métricas de classificação, onde a menor acurácia obtida foi de 0,933 pela DT da barra 11.

Esses resultados mostram a eficácia da abordagem descentralizada proposta, que provê informação suficiente para a análise de segurança dinâmica para pequenas perturbações

Tabela 6 – Performance no terceiro cenário.

Árvore de decisão	Acurácia	F1-Score
Barra 01	0,980	0,980
Barra 02	0,980	0,980
Barra 03	0,986	0,986
Barra 04	0,983	0,983
Barra 05	0,976	0,976
Barra 06	0,983	0,983
Barra 07	0,980	0,980
Barra 08	0,966	0,967
Barra 09	0,980	0,980
Barra 10	0,976	0,977
Barra 11	0,973	0,974
Barra 12	0,966	0,967
Barra 13	0,976	0,977
Barra 14	0,970	0,970
Barra 15	0,980	0,980
Barra 16	0,970	0,970
Média DTs distribuídas	0,976	0,977
DT Centralizada	0,983	0,983

(de qualquer predição das DTs individuais) e auxilia na decisão do operador do sistema de transmissão.

Tabela 7 – Performance no quarto cenário.

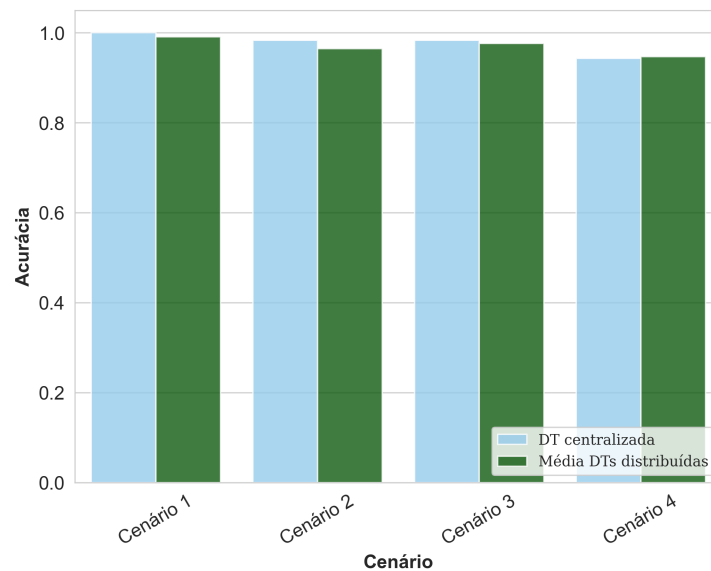
Árvore de decisão	Acurácia	F1-Score
Barra 01	0,943	0,946
Barra 02	0,956	0,958
Barra 03	0,956	0,958
Barra 04	0,943	0,946
Barra 05	0,943	0,946
Barra 06	0,943	0,946
Barra 07	0,963	0,964
Barra 08	0,943	0,946
Barra 09	0,960	0,961
Barra 10	0,936	0,940
Barra 11	0,933	0,935
Barra 12	0,953	0,955
Barra 13	0,946	0,949
Barra 14	0,946	0,949
Barra 15	0,943	0,946
Barra 16	0,953	0,955
Média DTs distribuídas	0,947	0,950
DT Centralizada	0,943	0,946

5.4 Discussões

Uma comparação visual entre as acurácias médias obtidas pela abordagem proposta e pela abordagem centralizada pode ser observada na Figura 22. Nota-se que não há diferença significativa entre as acurácias alcançadas em diferentes cenários. Além disso,

quando se observa a degradação do desempenho entre o cenário 1 e 4, nota-se que a abordagem descentralizada apresenta perda entre 4,1% e 4,4%, enquanto que a mesma análise revela perda entre 5,4% e 5,7%, de f1-score e acurácia, respectivamente, para a abordagem centralizada. Dessa forma, a abordagem centralizada apresenta um comportamento menos estável entre os diferentes cenários.

Figura 22 – Acurácia em diferentes cenários



Fonte: Autoria própria.

Por fim, para a medição do tempo de processamento necessário para que as árvores façam a predição de uma amostra, utilizou-se 300 casos e retirou-se a média do tempo de predição. As árvores distribuídas apresentam tempo médio de predição por amostra de 0,8 ms, enquanto que a árvore de decisão centralizada apresenta tempo médio de 0.87 ms. Dessa forma, ambas as abordagens seriam capazes de classificar uma amostra muito mais rápido do que os métodos convencionais.

Capítulo 6

Resultados - Classificação de Estabilidade Angular do Rotor Transitória a partir de Árvore de Decisão Centralizada

Neste capítulo, são apresentados os resultados obtidos na análise de estabilidade transitória a partir de uma árvore de decisão centralizada, analisando, principalmente os impactos da redução da quantidade de ciclos de medição e seleção de *features* no desempenho do classificador, com o intuito de agilizar a identificação da instabilidade no pós-falta e reduzir a quantidade de medidores necessários.

O sistema teste permanece o IEEE 68 barras, porém os cenários de operação são gerados de forma que a resposta temporal do sistema no período pós-falta seja distinta. Para verificar a eficácia do classificador proposto e mensurar a sua capacidade de identificação da instabilidade transitória a partir de uma quantidade reduzida de ciclos, três tipos distintos de testes são realizados:

1. Robustez da DT quando comparada com outro tipo de classificador tradicionalmente empregado na literatura;
2. Análise de viabilidade da abordagem distribuída;
3. Impactos de redução de *features* no desempenho do classificador: quantidade de ciclos de medição e do número de medidas.

Os resultados obtidos da realização de cada teste, bem como uma análise sobre a profundidade da árvore são discutidos com maior detalhamento no decorrer deste capítulo.

6.1 Composição do Banco de Dados

Inicialmente, é necessário preparar o banco de dados de modo que cenários sejam variados e contribuam para o poder de generalização do modelo. Para isso, são aplicadas faltas trifásicas no sistema teste, em que quatro parâmetros foram alterados: nível de carregamento da rede, linha em que a falta é aplicada, posição da falta na linha e tempo de abertura. A partir do carregamento base, foram definidos 3 níveis distintos de carregamento (com base no fluxo de carga continuado, conforme definido no capítulo 5):

- ❑ Carga leve: -17% do carregamento base;
- ❑ Carga base;
- ❑ Carga pesada: $+17\%$ do carregamento base.

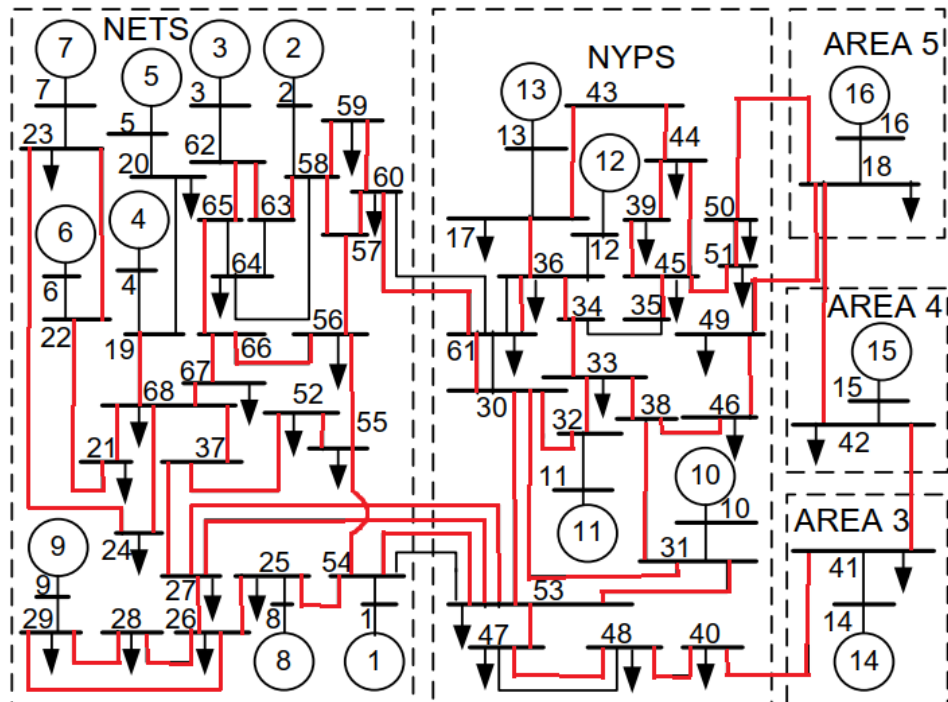
Para cada nível de carregamento, foram aplicadas faltas trifásicas, no instante $t = 1s$ de cada simulação. Ao todo foram aplicadas faltas em 63 linhas, marcadas em vermelho na Figura 23, variando-se os tempos de abertura entre 1,5, 1,55, 1,6 e 1,65 segundos. Além disso, também variou-se o os trechos de abertura na linha entre 20%, 40%, 60% e 80%. Todas essas combinações, geraram cartões distintos do ANAREDE e ANATEM, os quais foram elaborados com auxílio de um *script* desenvolvido no Matlab.

Com os cartões prontos, as simulações no domínio do tempo são realizadas, monitorando-se e armazenando-se as magnitudes e ângulos de tensão nas barras de geração, além dos ângulos dos rotores. A partir do ângulos dos rotores armazenados, calcula-se o TSI para cada um dos ciclos de medição, para que se tenha visão ampla da simulação e possa se definir rotular as saídas (como estáveis ou instáveis) com maior segurança. Nessa etapa, considera-se amostras com TSI superior a 0, em todos os ciclos, como estáveis, enquanto que qualquer ciclo com TSI igual ou menor a 0 caracteriza a amostra como instável.

Na Figura 24 é possível observar as quantidades de casos estáveis e instáveis para os diferentes tempos de abertura avaliados. Nota-se que quando se aumenta o tempo de abertura, aumenta-se a proporção de casos instáveis, como era esperado. Como citado no Capítulo 4 o processo de balanceamento de classes é importante para o treino e validação¹. Para esse *dataset*, é possível se observar pelo histograma da Figura 24 que um corte em 300 casos proporciona um balanceamento adequado entre os tempos de abertura, sem que se perca uma grande quantidade de amostras. Dessa forma, todos os testes são realizados com a seleção aleatória de 300 casos estáveis e instáveis para cada tempo de eliminação.

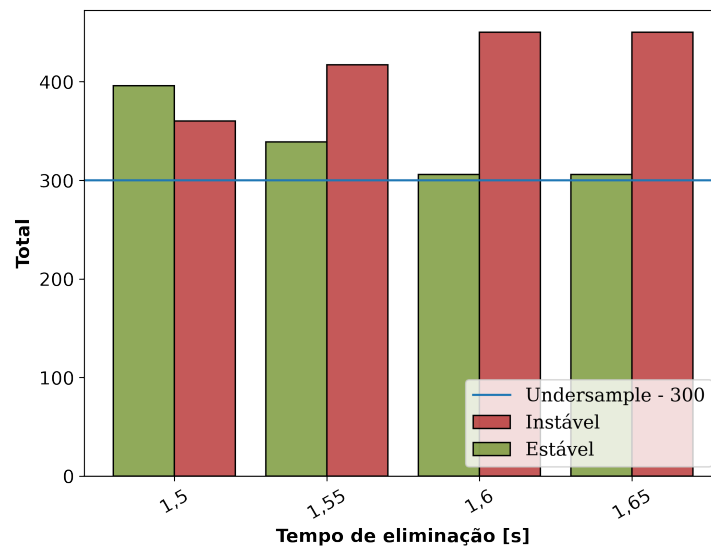
¹ No método 1 apenas se realizou o balanço nos testes, porque, com exceção do primeiro cenário, os demais *datasets* apresentavam proporção entre classes muito próxima de 1.

Figura 23 – Falhas trifásicas



Fonte: Autoria própria

Figura 24 – Distribuição de casos x tempo de eliminação



Fonte: Autoria própria

Com o banco de dados montado, a estabilidade das amostras rotulada e o processo de *undersample* realizado é possível dar continuidade a construção do classificador, estabelecendo o modelo de predição. Nas seções seguintes são apresentados os critérios utilizados nos testes propostos e os resultados obtidos.

6.2 Critério de simulação dos testes propostos

Todos os testes realizados fazem uso dos mesmos dados, tornando possível analisar como diferentes abordagens se comportam diante do mesmo problema e, explorar suas possíveis limitações. Para que se tenha uma avaliação estatística do potencial do modelo também usa-se *k-Folds*, com $k = 10$. Dessa forma, avalia-se diferentes conjuntos disjuntos de amostras e detecta-se possíveis inconsistências no desempenho dos classificadores.

Antes de iniciar os testes, aplica-se a seleção de *features* individuais para identificar a importância do tipo de variável do vetor de entrada na acurácia da classificação. Após ser estabelecido a *feature* de maior relevância, define-se o parâmetro de profundidade da DT. Nessa etapa, avalia-se a variação de diferentes métricas à variação da profundidade.

6.2.1 Seleção de *Features* - tipo de Variável

Nesta etapa, com o intuito de identificar se há maior importância em um dos tipos de variável (magnitude e ângulo) para a classificação e, assim, promover a redução do vetor de entradas, aplica-se o Teste F .

Na Figura 25 observa-se os valores F para as amostras medidas no primeiro ciclo de acordo com a barra de geração, quando o sistema está sujeito a faltas trifásicas com 500 ms de duração (tempo de abertura de 1,5 s). Note que as medições de tensão apresentam valores significativamente superiores aos ângulos. Além dos valores F , o teste também retorna valores p nulos para a maioria das barras, de forma que, como apresentado no Capítulo 4, pode se considerar que os valores F são indicativos de *features* com maior relevância na separação das classes.

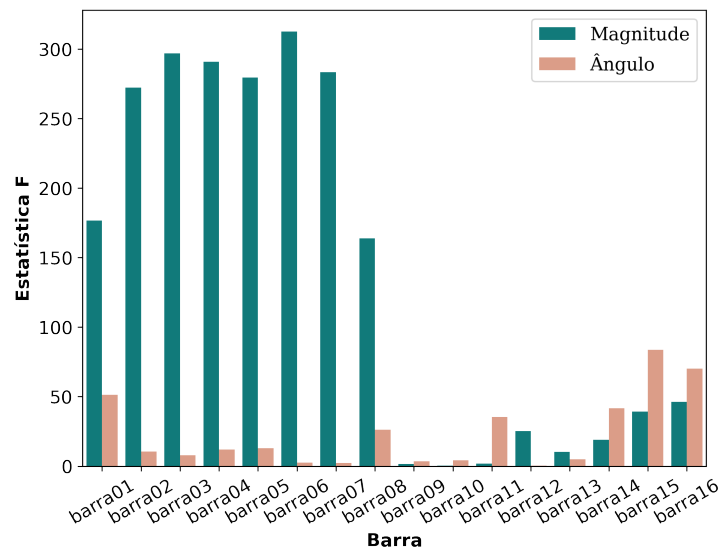
Esse mesmo padrão se repete nos demais ciclos e em outros tempos de eliminação, de forma que se opta por usar apenas o grupo de medições de magnitude de tensão nos demais testes, com exceção do teste com classificadores distribuídos que já apresenta um vetor de entradas extremamente reduzido e poderia ser prejudicado por essa seleção.

6.2.2 Escolha da profundidade

Para a escolha da profundidade, o banco de dados foi separado em 40% de amostras para teste e 60% para treino. Após definido as amostras que compõe o banco de treino e de teste, executou-se o treinamento e verificou-se a classificação atingida pela árvore centralizada quando profundidades distintas foram avaliadas, com níveis de 1 a 30, a partir dos dados do banco de teste.

Nesse estudo, utilizou-se a magnitude de tensão de cada uma das barras de geração do sistema teste, com 8 ciclos de medição de cada variável. Na Figura 26 observa-se que com uma profundidade próxima de 10 as métricas tendem a saturar e não há ganho na escolha de um valor superior, pelo contrário, o aumento da profundidade tornaria a DT mais

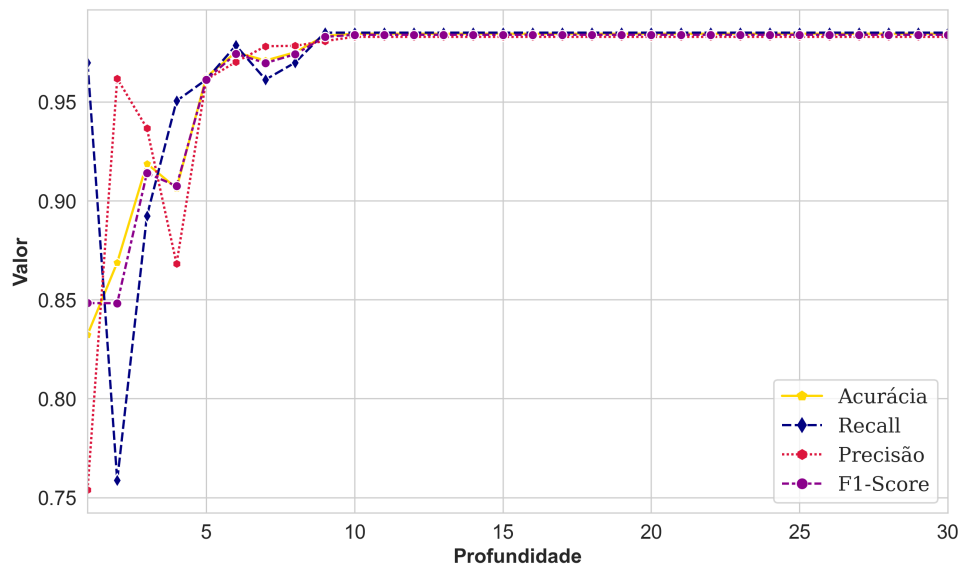
Figura 25 – Teste F - Tempo de eliminação de 1,5 segundos



Fonte: Autoria própria

complexa e poderia levá-la ao *overfitting*. Dessa forma, para os demais testes executados fixou-se a profundidade em 10.

Figura 26 – Desempenho da DT para diferentes profundidades



Fonte: Autoria própria

6.3 Resultados

Os resultados obtidos para composições distintas do vetor de entrada são apresentados em sequência.

6.3.1 Robustez à redução de ciclos de medição

Neste primeiro teste avalia-se o desempenho da DT centralizada quando a quantidade de ciclos de entrada é reduzido. Neste caso o vetor de entrada do classificador é composto pela magnitude da tensão nas barras de geração, definido como $[V_{1_1}, V_{2_1}, \dots, V_{N_{barras_N_{ciclos}}}]$, lembrando que o primeiro subíndice indica a barra de medição e o segundo indica o ciclo de medição. A dimensão do vetor de entrada se altera de acordo com o número de ciclos avaliados, uma vez que se deseja identificar o desempenho da DT centralizada quando a quantidade de ciclos de entrada é reduzido.

Pela Equação (35) é possível obter a dimensão do vetor de entrada (dim) a partir dos valores de número de geradores ($n_geradores$), quantidade de *features* por gerador ($n_features_por_gerador$) e quantidade de ciclos de medição (n_ciclos). Considerando que o sistema teste possui 16 geradores e que apenas a magnitude da tensão foi utilizada, para 1 ciclo de medição o vetor de entrada teria dimensão 16, enquanto que para 8 ciclos o vetor teria dimensão 128.

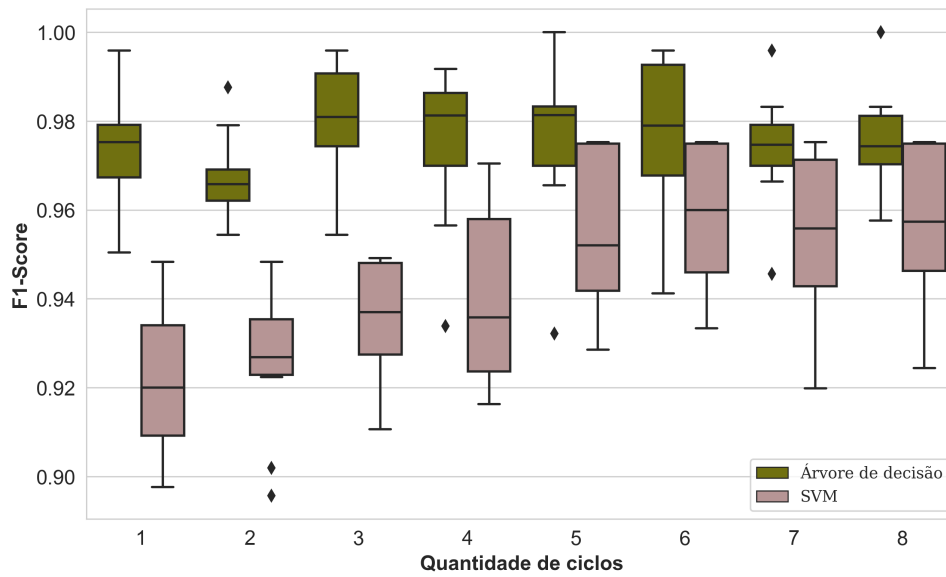
$$dim = n_geradores \cdot n_features_por_gerador \cdot n_ciclos \quad (35)$$

A partir desse mesmo vetor de entrada, um segundo classificador é treinado utilizando a técnica SVM, a título de comparação. Para as duas técnicas, aplica-se a técnica de *k-folds*, com $k = 10$. Na Figura 27 é possível observar o *boxplot* com relação ao *f1-score* obtido pela DT centralizada e pela SVM, em que cada *box* é resultado dos 10 testes realizados (10-folds).

Como pode-se notar, ambos os classificadores obtêm bom desempenho próximos da quantidade máxima de ciclos, porém quando se reduz essa quantidade ocorre uma degradação mais acentuada no SVM, enquanto que a DT permanece com um nível de classificação estável no intervalo de ciclos analisado. A DT obtêm *f1-score* de 97,6% para 8 ciclos e cai para 97,3% para 1 ciclo, enquanto que o SVM parte de 95,7% para 8 ciclos e atinge 92,2% para 1 ciclo. Nota-se, portanto, que a DT centralizada é capaz de realizar a classificação com precisão mesmo quando a quantidade de ciclos é reduzida.

Quanto ao tempo de processamento, quando se utiliza os 8 ciclos a árvore de decisão apresenta um tempo médio de predição de 0,93 ms por amostra, enquanto que para 1 ciclo esse tempo é reduzido para 0,85 ms. Para a obtenção dos tempos de predição, as árvores treinadas foram usadas na classificação de 440 amostras, sendo que ao final do processo foi extraída a média dos tempos por amostra. Fica evidente que mesmo sem a redução de ciclos ou número de PMUs, o tempo necessário para a realização de uma predição é ínfimo quando comparado aos métodos tradicionais.

Figura 27 – Desempenho de DT e SVM para diferentes quantidades de ciclos de medição



Fonte: Autoria própria

6.3.2 Abordagem distribuída

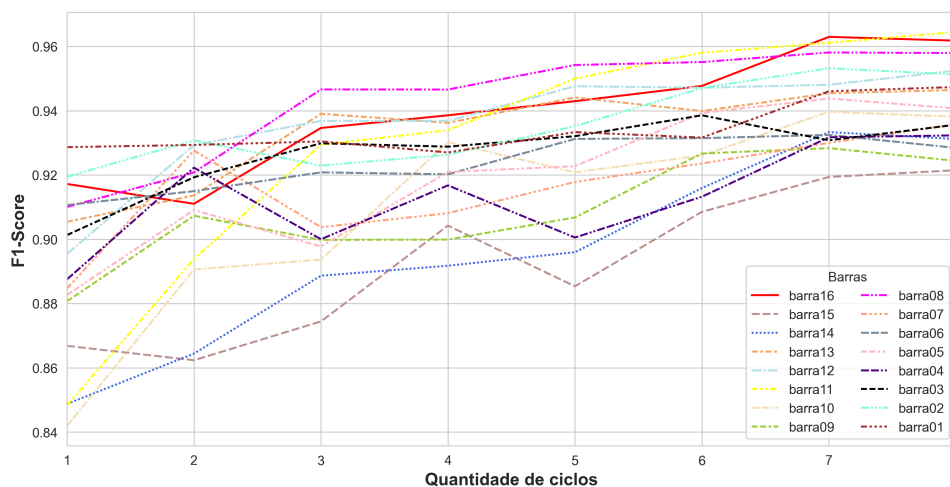
Neste teste, procurou-se analisar a viabilidade de utilizar uma abordagem distribuída para a classificação de estabilidade transitória, de forma análoga ao que foi feito para pequenas perturbações, em que foram empregadas DTs distribuídas na tarefa de identificação da estabilidade. Para tanto, além da magnitude da tensão também se usou o ângulo das barras geradoras como entrada, sendo que cada classificador faz uso de duas *features*. Novamente, ressalta-se que para os classificadores distribuídos a dimensão do vetor de entrada já é bastante reduzida e, a remoção das medições de ângulo poderia prejudicar o desempenho dos modelos.

Realizou-se um primeiro teste dividindo o *dataset* em 40% teste e 60% treino para se analisar o *f1-score* em todas as barras. Após esse teste, selecionou-se uma barra que obteve resultados promissores para um teste aprofundado usando *k-folds* com as mesmas configurações do teste anterior.

Como pode se observar na Figura 28 todos os classificadores perdem desempenho de forma significativa quando se reduz a quantidade de ciclos de medição. Com 8 ciclos o melhor classificador, relacionado à barra 11, obtêm um resultado de 96,4%, enquanto que o pior classificador, relacionado à barra 15 atinge 92,1% de *f1-score*.

Para uma análise usando todo o *dataset*, fazendo-se uso de *k-folds*, a barra 16 é selecionada para o teste. Essa barra apresenta 96,2% de *f1-score* para 8 ciclos, o que representa o segundo melhor desempenho, atrás apenas da barra 11, mas quando se considera 1 ciclo de medição a barra 11 sofre uma perda de desempenho aguda atingindo 84,8%, estando entre os 3 piores classificadores juntamente com as barras 10 e 14, enquanto que a barra 16 permanece entre os 3 melhores classificadores com 91,7% juntamente com a barra 1 e

Figura 28 – Desempenho das DTs distribuídas para diferentes quantidades de ciclos de medição



Fonte: Autoria própria

barra 2.

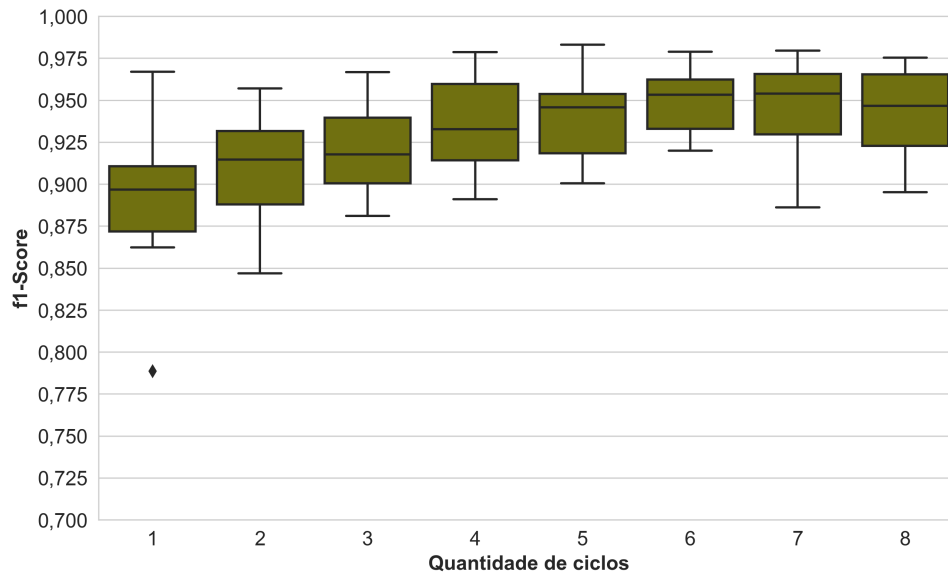
No entanto, os resultados obtidos com a árvore de decisão na barra 16, mostram que o classificador distribuído não é robusto a redução de ciclos. A mediana do $f1 - score$ é de 94,3% para 8 ciclos, reduzindo para 89,2% quando utiliza-se apenas um ciclo. Também é importante ressaltar que com o *boxplot* e o *k-folds* torna-se possível uma análise mais confiável do potencial dos classificadores. Quando se observa a Figura 28, onde não ocorre uma avaliação do classificador para todo o *dataset*, o desempenho do classificador da barra 16 alcança 96,2% de $f1 - score$ em 8 ciclos de medição. No entanto, quando se aplica *k-folds* observa-se uma mediana de 94,3%, ressaltando que mesmo com 8 ciclos o classificador não apresenta desempenho tão próximo da DT centralizada.

6.3.3 Seleção de *features* combinadas - Avaliação das barras de medição

Nesta etapa, novamente, se faz uso de seleção de *features* para reduzir ainda o vetor de entrada. Desta vez, o objetivo é reduzir o número de PMUs necessárias para o classificação do modelo e avaliar os impactos no seu desempenho. Para tanto, utilizou-se *Random Forest* como seletor, por sua capacidade intrínseca de quantificar a importância das entradas combinadas. Vale ressaltar que qualquer outro modelo de aprendizagem de máquina com ranqueamento interno de *features* também poderia ser utilizado.

Como apresentado no Capítulo 4, o procedimento consiste em treinar a *Random Forest* considerando conjuntos distintos de ciclo de medição (de 1 à 8), com 255 iterações por ciclo. Em cada iteração, em que um novo ciclo de medição é incorporado ao vetor de entrada do classificador, um novo conjunto de amostras é selecionado do *dataset* de forma

Figura 29 – Desempenho da DT 16 para diferentes quantidades de ciclos de medição



Fonte: Autoria própria

aleatória.

Para cada combinação, ciclo e iteração, as 10 *features* com maior importância são armazenadas. Ao final do processo, é possível construir uma matriz contendo a frequência de aparição das *features* entre as 10 entradas mais significativas por ciclo, apresentada na Tabela 8.

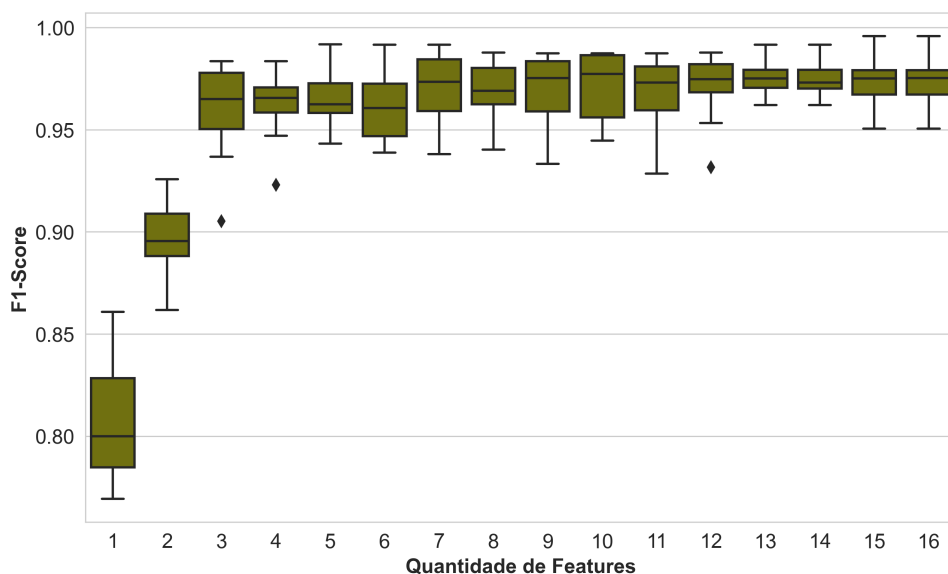
Tabela 8 – Importância de *features* por ciclo - Seleção por *Random forest*

	ciclo 1	ciclo 2	ciclo 3	ciclo 4	ciclo 5	ciclo 6	ciclo 7	ciclo 8
B01	1.0	0.97	0.81	0.43	0.26	0.29	0.36	1.0
B02	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
B03	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
B04	0.18	0.15	0.33	0.36	0.53	0.45	0.35	0.18
B05	0.13	0.11	0.23	0.36	0.42	0.67	0.59	0.13
B06	0.95	0.91	0.58	0.53	0.55	0.26	0.42	0.95
B07	0.65	0.32	0.43	0.59	0.62	0.71	0.64	0.65
B08	1.0	1.0	1.0	0.95	0.83	0.84	0.89	1.0
B09	0.14	0.36	0.57	0.62	0.47	0.36	0.28	0.14
B10	0.84	0.66	0.36	0.36	0.68	0.93	0.98	0.84
B11	1.0	0.98	0.99	1.0	1.0	1.0	1.0	1.0
B12	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
B13	1.0	1.0	1.0	0.99	0.93	0.87	0.81	1.0
B14	0.01	0.0	0.01	0.01	0.0	0.02	0.06	0.01
B15	0.0	0.0	0.0	0.0	0.0	0.08	0.49	0.0
B16	0.13	0.53	0.69	0.82	0.71	0.52	0.13	0.13

Nota-se, na Tabela 8 que algumas barras aparecem entre as *features* mais relevantes em diferentes ciclos, como a barra 1, 2, 3, 11, 12 e 13, por exemplo. O que sugere a possibilidade de se utilizar um grupo reduzido de entradas para a classificação. Para testar essa hipótese, nessa última etapa, a DT centralizada é testada por etapas, acrescentando-se *features*, uma a uma, desde a mais importante até a menos importante. Dessa forma, busca-se um limiar de quantidade de entradas a partir da qual o acréscimo de mais PMUs não resultaria em ganho de desempenho significativo na classificação. A seguir são apresentados e discutidos os resultados da variação de quantidade de entradas, por ciclo de medição.

Pode-se observar na Figura 30 que a árvore alcança um patamar de saturação com 3 das *features* mais significativas, apresentando uma mediana de 0,964. O valor máximo de mediana alcançado com 1 ciclo de medição foi de 0,977, referente à 10 *features*, ou seja, não parece haver ganho significativo de desempenho por parte do classificador quando se utiliza mais de 3 PMUs para a construção do vetor de entrada.

Figura 30 – Desempenho da DT centralizada para o ciclo 1 em função da quantidade de *features*

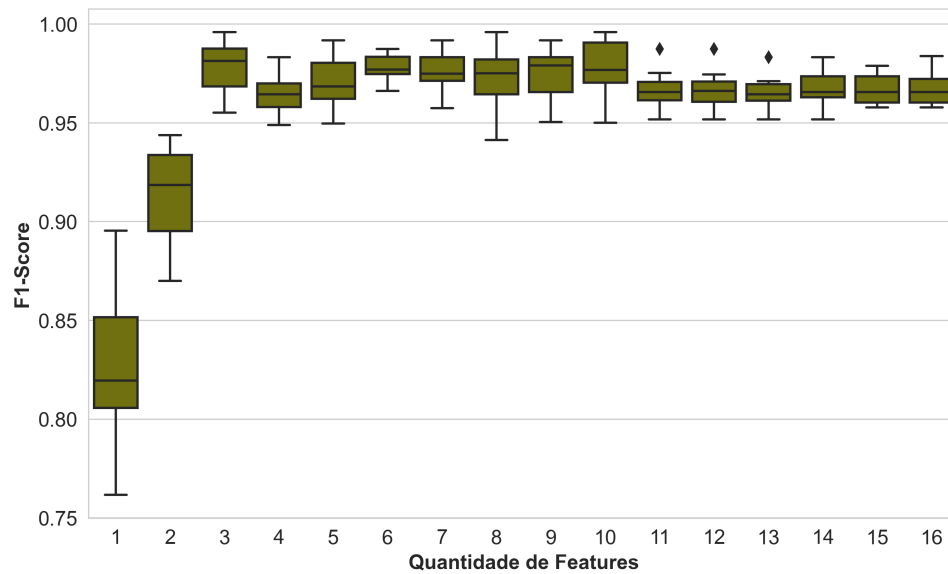


Fonte: Autoria própria

Na Figura 31, observa-se um comportamento semelhante ao primeiro ciclo, em que o classificador alcança um patamar estável de desempenho com 3 *features*. Neste caso, a mediana máxima de 0,981 é obtida com 3 *features* e, com o acréscimo de mais PMUs o modelo permaneceria com um f1-score em torno de 0,97. Novamente, não há ganho significativo além das 3 *features* mais importantes.

Para o terceiro ciclo de medição, a Figura 32 mostra que os resultados, quando estabilizados, permanecem em torno de 0,97. Com 3 *features*, já se alcança 0,966 de mediana. Resultado que reforça o obtido para os dois primeiros ciclos e sugere não haver diferença

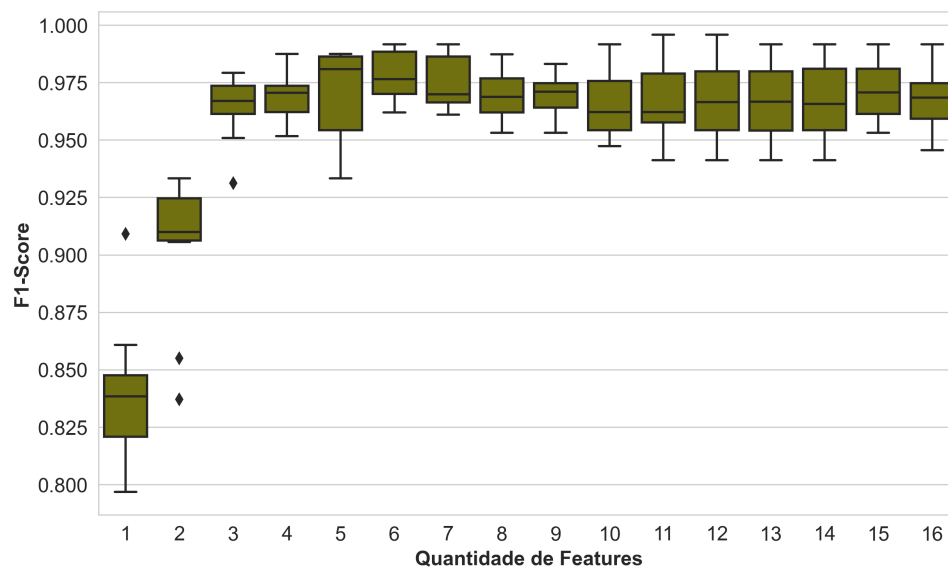
Figura 31 – Desempenho da DT centralizada para o ciclo 2 em função da quantidade de *features*



Fonte: Autoria própria

significativa de desempenho, mesmo entre ciclos.

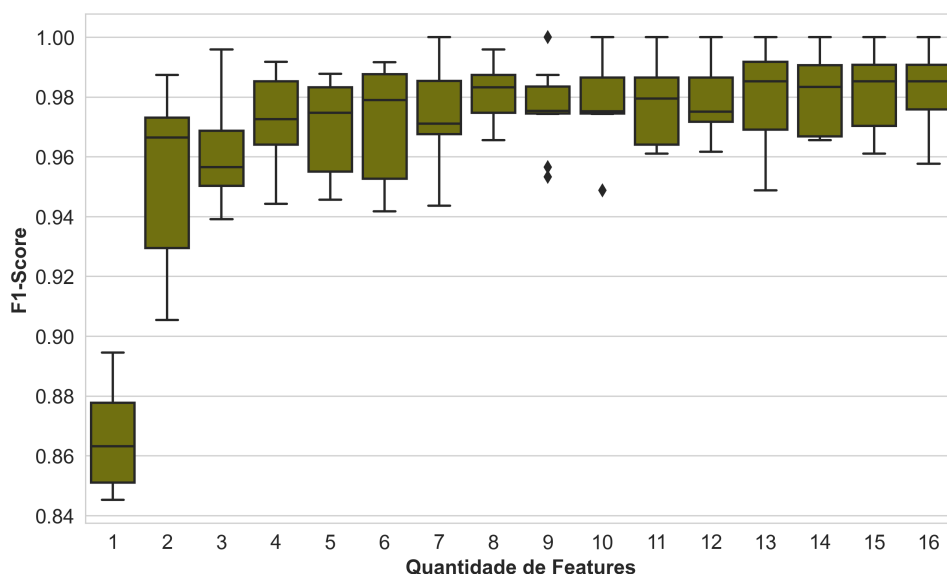
Figura 32 – Desempenho da DT centralizada para o ciclo 3 em função da quantidade de *features*



Fonte: Autoria própria

Para o quarto ciclo, representado na Figura 33, a quantidade de 4 *features* parece mais próxima da saturação, apresentando uma mediana de 0,972. O máximo observado de 0,9852 é obtido com 13 entradas. Para este ciclo, os valores de mediana estabilizam próximos de 0,98.

Figura 33 – Desempenho da DT centralizada para o ciclo 4 em função da quantidade de *features*



Fonte: Autoria própria

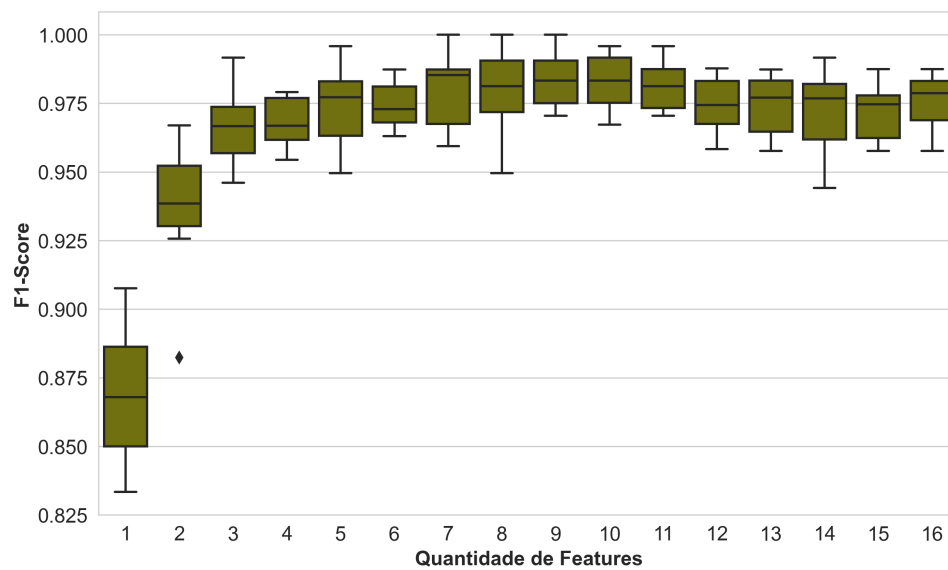
Para o quinto ciclo a quantidade de entradas mínima se desloca para 5 entradas, com uma mediana de 0,977, como se observa na Figura 34. Neste caso, a curva satura em torno de 0,98, atingindo seu máximo de 0,985 para 7 entradas. Observa-se que mesmo com um número um pouco maior de entradas, quando comparado aos resultados anteriores, ainda é possível realizar uma classificação com alto f1-score com, aproximadamente, um terço do total de PMUs disponíveis.

Na Figura 35, observa-se um deslocamento da saturação para 6 *features*, obtendo uma mediana de 0,977. Com o acréscimo de medidores o modelo permanece próximo de 0,98.

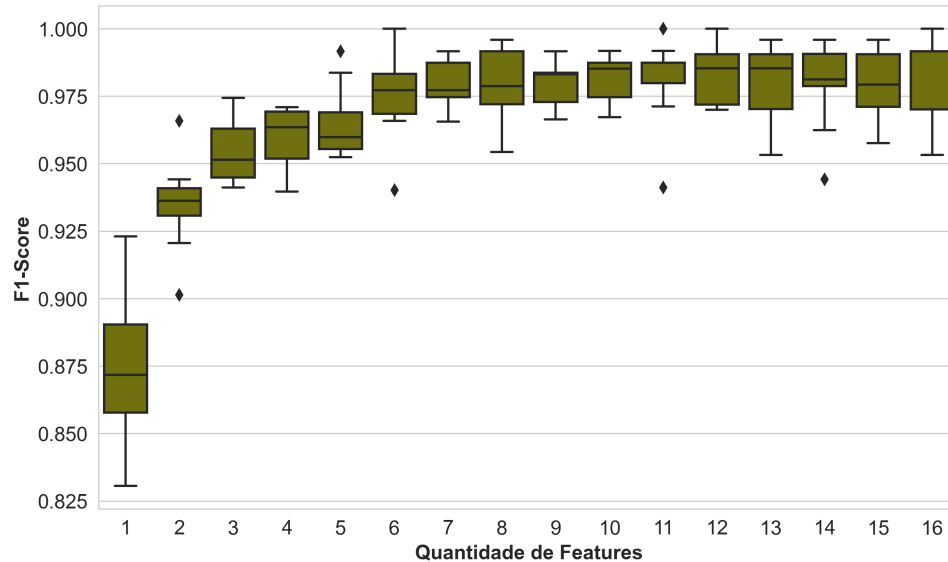
Na Figura 36, para o sétimo ciclo de medição, o desempenho se estabiliza com 6 entradas, alcançando uma mediana de 0,975, e permanecendo em torno de 0,97 de f1-score mesmo com a adição de PMUs. O máximo de 0,985 é alcançado com 7 *features*.

Finalmente, para o oitavo ciclo, a Figura 37 apresenta desempenho próximo de 0,97, sendo que com 3 *features* já se alcança uma mediana de 0,964. O valor máximo obtido no último ciclo é de 0,977, para 10 *features*. Novamente, não há ganho significativo de desempenho quando se utiliza mais de 3 PMUs para a classificação.

De forma geral, todos os casos avaliados apresentaram um comportamento semelhante, em que se atingiu a saturação do desempenho com poucas *features*, evidenciando a viabilidade da redução da quantidade de PMUs para a classificação, o que gera economia na implementação do modelo. Além disso, esse resultado também é consistente com a análise de impacto de redução da quantidade de ciclos no desempenho da árvore de decisão. Como observado no início do capítulo, a árvore já é capaz de classificar com alto desempenho desde o primeiro ciclo, não havendo diferença significativa com o aumento

Figura 34 – Desempenho da DT centralizada para o ciclo 5 em função da quantidade de *features*

Fonte: Autoria própria

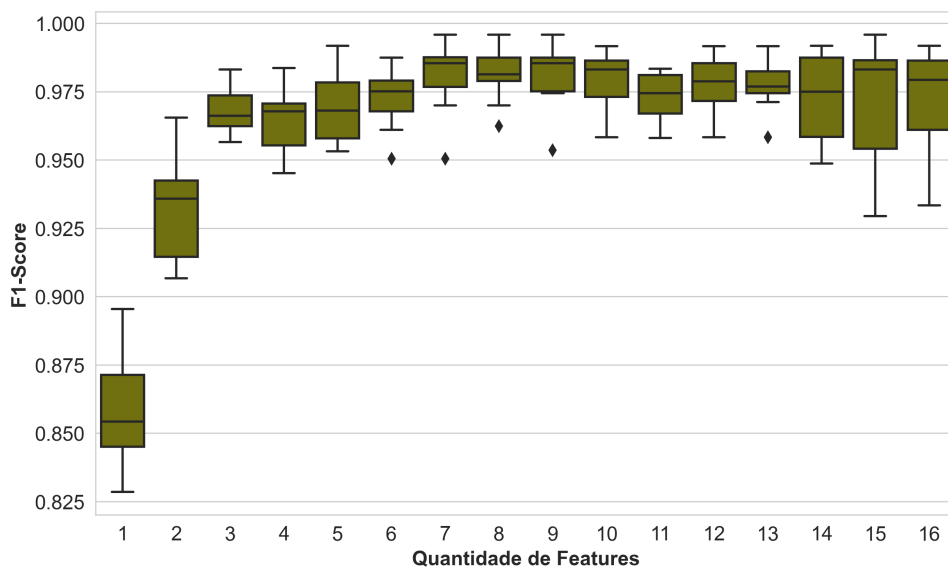
Figura 35 – Desempenho da DT centralizada para o ciclo 6 em função da quantidade de *features*

Fonte: Autoria própria

de ciclos.

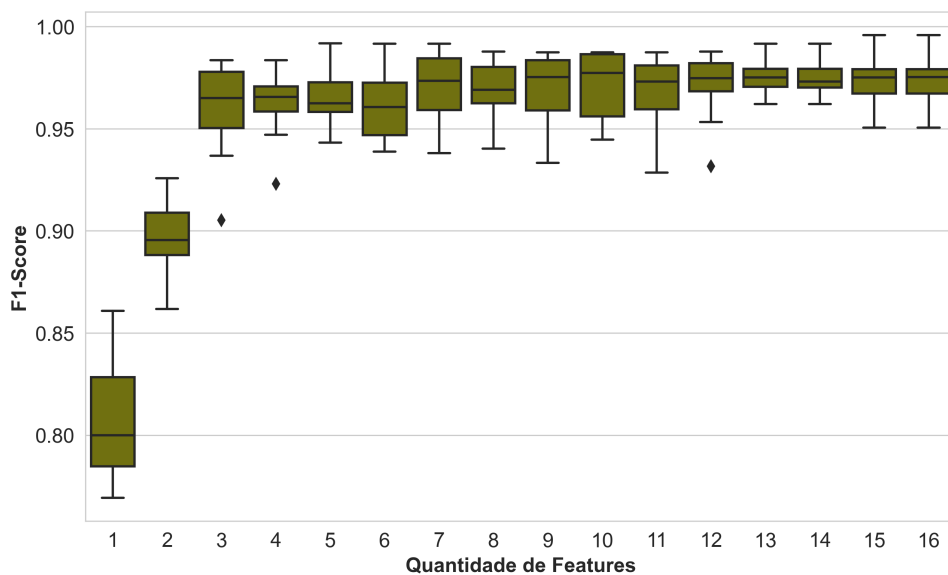
Novamente, como feito no teste referente à redução do número de ciclos de medição, executou-se predições com a árvore de decisão treinada com as 3 principais *features*, com o objetivo de obter o tempo de predição por amostra. Foram usadas 440 amostras e o tempo médio de predição obtido foi de 0,86 ms. A redução de número de PMUs, nesse caso, não apresentou ganho na velocidade da predição da árvore, que já era de 0,85 ms

Figura 36 – Desempenho da DT centralizada para o ciclo 7 em função da quantidade de features



Fonte: Autoria própria

Figura 37 – Desempenho da DT centralizada para o ciclo 8 em função da quantidade de features



Fonte: Autoria própria

no teste realizado com todas as PMUs. Portanto, ressalta-se que o tempo necessário para o uso do classificador durante o monitoramento do SEP é extremamente baixo.

Capítulo 7

Conclusões

Como exposto durante os capítulos anteriores, este trabalho tem como proposta aproveitar a disponibilidade de dados do SEP fornecida pelas PMUs e avaliar a viabilidade do uso de árvores de decisão na classificação da estabilidade angular do rotor, tanto a pequenas perturbações quanto a grandes perturbações.

Para o desenvolvimento dos testes, é fundamental a criação dos bancos de dados que simulem amostras de operação do SEP em diversos cenários, permitindo que o algoritmo de aprendizagem ganhe poder de generalização e possa classificar casos distintos dos disponíveis nas bases de dados. Para a criação do banco de dados utilizado no treinamento do classificador de estabilidade a pequenas perturbações foram utilizadas variações de carga diárias e contingências do tipo $N-1$, a partir dos *softwares* Matlab, ANAREDE e PacDyn. Para a criação do banco de dados para a análise de estabilidade transitória foram armazenadas séries temporais das respostas do sistema diante de perturbações e, para tanto, foi utilizado o *software* ANATEM para realização das simulações no domínio do tempo.

Os resultados obtidos para a análise de estabilidade a pequenas perturbações mostram uma alto desempenho das árvores de decisão distribuídas e da árvore centralizada, mesmo quando expostas a diferentes topologias advindas das contingências $N-1$. O desempenho das árvores distribuídas corrobora com a proposta de utilizar as medições das PMUs associadas às barras geradoras de forma individualizada, aumentando a redundância e, por consequência, a robustez da avaliação frente a perdas de medição.

Foram analisados quatro cenários distintos. Um primeiro com variações de carga ativa e reativa de mesma intensidade apresentou os melhores resultados, com um *f1-score* médio de 99,1% contra 100% da DT centralizada, é importante ressaltar que a curva de deslocamento do modo interárea analisado, neste caso, apresentou-se bastante

comportada e os bons resultados na classificação eram esperados. Esse cenário apresentou o maior desbalanceamento entre classes e, é possível que as DTs distribuídas obtivessem desempenhos ainda superiores se houvesse um processo de balanceamento.

No segundo cenário, manteve-se variações de carga com a mesma direção de crescimento do cenário anterior e, incluiu-se contingências $N-1$, dessa forma observou-se maior dispersão entre as trajetórias do modo analisado e os classificadores distribuídos obtiveram uma média de 96,6% de *f1-score* contra 98,3% da DT centralizada. Neste cenário, a distribuição de classes estava próxima de 1 e, dessa forma, reduziu-se as chances de vies nos modelos e garantiu-se um bom resultado.

O terceiro cenário é semelhante ao primeiro, manteve-se a topologia base, porém as variações de carga ativa e reativa foram independentes, ocasionando vetores de variação de carga com direções distintas. Nesse contexto, observou-se maior dispersão da trajetória do modo quando comparada ao cenário 1 o que dificulta o processo de classificação. Ainda assim, as árvores distribuídas mantiveram um bom desempenho de 97,7% de *f1-score* contra 98,3% da árvore centralizada. Quanto ao balanceamento, no cenário 3 foram gerados dados em torno da região de transição e, dessa forma, o banco de dados já apresentava um balanceamento próximo de 1, sem a necessidade de posteriores intervenções.

Finalmente, o quarto cenário é composto por diferentes direções de crescimento do vetor de variação de carga e contingências $N-1$. Nesse cenário, as DTs apresentam um desempenho inferior aos outros casos e as trajetórias do modo de interesse aparecem com maior dispersão. As DTs distribuídas obtêm um *f1-score* médio de 95% contra 94,6% da árvore centralizada. Novamente, foram gerados casos próximos da região de transição entre estável e instável e, dessa forma, o balanceamento das classes esteve muito próximo de 1.

Como pode-se observar em todos os cenários, as DTs distribuídas obtiveram métricas elevadas, sendo seu pior desempenho no quarto cenário com 95% de *f1-score*. É importante ressaltar que o principal objetivo das DTs distribuídas é fornecer a classificação dos casos, com possibilidade de perda de medições, sem que haja perda significativa de desempenho quando comparadas a uma DT centralizada. Em todos os cenários é possível observar resultados muito próximos entre a abordagem distribuída e centralizada. Dessa forma, conclui-se que não há perda significativa na adoção da abordagem distribuída e, principalmente confere-se ao sistema robustez a perdas de medição.

Na avaliação a grandes perturbações, utilizou-se os *softwares* Matlab, ANAREDE e ANATEM, sendo o último responsável por executar as simulações no domínio do tempo e armazenar todas as medições de barras geradores e ângulos dos rotores. Para a montagem do banco de dados considerou-se variações de carga, faltas em diferentes linhas de transmissão, trechos das linhas e diferentes tempos de abertura. Vale ainda ressaltar que após o tempo de abertura não se considerou cenários com perdas de linhas e, portanto mudança de topologia. Foram executados testes com o intuito de avaliar diferentes abordagens e

comportamentos.

Se obteve resultados promissores com uma árvore de decisão utilizando todas as magnitudes de tensão de barras geradoras, mesmo quando se reduziu a quantidade de ciclos de medição, o que demonstrou robustez do algoritmo e um potencial para identificação de casos de instabilidade com antecedência, logo no primeiro ciclo de medição. Esse comportamento é interessante, pois possibilita maior tempo para atuação e mitigação de danos. Como critério de comparação treinou-se um SVM com os mesmos dados de entrada e notou-se que para 8 ciclos os desempenhos dos dois algoritmos não estavam tão distante, porém quando se reduziu o número de ciclos notou-se uma degradação significativa no SVM, acentuando a diferença com a DT. Enquanto isso, a DT permaneceu com desempenhos semelhantes independente da quantidade de ciclos.

Em um segundo momento, testou-se uma abordagem distribuída análoga à implementada para pequenas perturbações. Nesses testes foram usados ângulo e magnitude das barras gerados no treinamento de DTs individuais conectadas à cada barra geradora. Ao contrário do observado a pequenas perturbações, notou-se que mesmo com a quantidade máxima de ciclos nem todas as DTs apresentaram bom desempenho e, na análise de redução de ciclos todas as árvores mostraram degradação significativa. Para uma análise aprofundada, repetiu-se os testes com uma das DTs mais promissoras utilizando-se *k-folds* para abranger o *dataset* completo. Esse último teste mostrou um desempenho ainda inferior ao observado no teste anterior, em que a mediana do *boxplot* gerado encontrou valores ainda menores dos previamente obtidos.

Finalmente, como último teste analisou-se o efeito de seleção de *features* no desempenho dos classificadores. A preparação dessa etapa fez uso do ranqueamento implícito de *features* realizado por algoritmos baseados em DTs, em específico, utilizou-se *Random forest* que é um modelo comum para essa finalidade. Gerou-se uma matriz representando a frequência de aparecimento de *features* entre as mais importantes para diferentes quantidades de ciclos de medição. A partir dessa matriz realizou-se testes por ciclo, avaliando-se o impacto de diferentes quantidade de *features* no desempenho da árvore, partindo de *features* menos significativas para mais significativas. Os resultados obtidos mostram o potencial para classificação com o uso de apenas 3 medidores, em quase todos os ciclos analisados, sendo que não se observa ganho de desempenho significativo com o acréscimo de mais medidores. Dessa forma, este último resultado reafirma a capacidade da árvore de decisão de obter alto desempenho com apenas um ciclo de medição e, ainda apresenta a possibilidade de realização da classificação com menos de um quarto (3 medidores) da quantidade total de PMUs, oferecendo maior economia na implementação da solução.

Como pesquisas futuras, para pequenas perturbações as trajetórias de diferentes topologias sugerem a possibilidade de clusterização, dessa forma, pode-se investigar o treinamento de classificadores específicos para grupos de topologias com dinâmicas semelhantes e, assim, melhorar o desempenho de classificação. Ainda a pequenas perturbações, sobre

a formação de banco de dados, o processo de rotulação das amostras depende do modelo matemático do sistema e de análise modal. Para que o treinamento de novos classificadores seja independente do modelo matemático do SEP, seria interessante investigar o uso de técnicas de estimação modal para a definição de estabilidade offline.

Referências

ALBERTO, L. F. C.; BRETAS, N. G. **Estabilidade transitória em sistemas eletroenergéticos**. São Carlos, Brasil: EESC-USP, 2000.

ALIMI, O. A.; OUAHADA, K.; ABU-MAHFOUZ, A. M. A Review of Machine Learning Approaches to Power System Security and Stability. **IEEE Access**, v. 8, p. 113512–113531, 2020. ISSN 21693536.

AMERONGEN, R. van; MEETEREN, H. P. van. A generalised ward equivalent for security analysis. **IEEE Transactions on Power Apparatus and Systems**, PAS-101, n. 6, p. 1519–1526, 1982.

AMRAEE, T.; RANJBAR, S. Transient instability prediction using decision tree technique. **IEEE Transactions on Power Systems**, IEEE, v. 28, n. 3, p. 3028–3037, 2013. ISSN 08858950.

AN, J. et al. A Data-driven Method for Transient Stability Margin Prediction Based on Security Region. **Journal of Modern Power Systems and Clean Energy**, v. 8, n. 6, p. 1060–1069, 2020. ISSN 21965420.

ANDERSON, P. M.; FOUAD, A. A. **Power System Control and Stability**. [S.l.]: IEEE Press, 1993.

ANGEL, A. D. D.; GLAVIC, M.; WEHENKEL, L. Using Artificial Neural Networks to Estimate Rotor Angles and Speeds from Phasor Measurements. **Isap-2003**, p. 1–6, 2003.

ASVAPOOSITKUL, S. **Data-Driven Small-Signal Stability Assessment and Preventive Control in Low Inertia Power Systems**. Tese (Doutorado) — University of Manchester, 2021.

ASVAPOOSITKUL, S.; PREECE, R. Decision tree-based prediction model for small signal stability and generation-rescheduling preventive control. **Electric Power Systems Research**, v. 196, p. 107200, 2021. ISSN 0378-7796.

Babakmehr, M. et al. Compressive sensing-based topology identification for smart grids. **IEEE Transactions on Industrial Informatics**, v. 12, n. 2, p. 532–543, 2016.

Balu, N. et al. On-line power system security analysis. **Proceedings of the IEEE**, v. 80, n. 2, p. 262–282, 1992.

- CANIZARES, C. et al. Benchmark models for the analysis and control of small-signal oscillatory dynamics in power systems. **IEEE Transactions on Power Systems**, v. 32, n. 1, p. 715–722, Jan 2017. ISSN 0885-8950.
- CAO, J.; FAN, Z. Deep learning-based online small signal stability assessment of power systems with renewable generation. In: **2018 IEEE Smart World, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation**. [S.l.: s.n.], 2018. p. 216–221.
- CHUNG, T. S.; YING, F. An ann-based network equivalent approach for power system on-line voltage security assessment. p. 1504–1507, 1998.
- CZAJKOWSKI, M.; JURCZUK, K.; KRETOWSKI, M. Tree based advanced relative expression analysis. In: _____. [S.l.: s.n.], 2020. p. 496–510. ISBN 978-3-030-50419-9.
- DEL ANGEL, A. et al. Estimation of rotor angles of synchronous machines using artificial neural networks and local PMU-based quantities. **Neurocomputing**, v. 70, n. 16-18, p. 2668–2678, 2007. ISSN 09252312.
- ELSSIED, N.; IBRAHIM, A. P. D. O.; OSMAN, A. H. A novel feature selection based on one-way anova f-test for e-mail spam classification. **Research Journal of Applied Sciences, Engineering and Technology**, v. 7, p. 625–638, 01 2014.
- FERNANDES, T. C. d. C. **Estimação de Velocidade Angular de Geradores Síncronos para Estudo da Estabilidade a Pequenas Perturbações em Sistemas de Potência**. Tese (Tese (Doutorado em Sistemas Elétricos de Potência)) — Universidade de São Paulo -USP, São Carlos, SP, 2017.
- GAVGANI, M. H.; EFTEKHARNEJAD, S. A graph model for enhancing situational awareness in power systems. **2017 19th International Conference on Intelligent System Application to Power Systems, ISAP 2017**, 2017.
- GERON, A. **Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. United States: O’reilly., 2017.
- GHAHREMANI, E.; KAMWA, I. Online state estimation of a synchronous generator using unscented kalman filter from phasor measurements units. **IEEE Transactions on Energy Conversion**, v. 26, n. 4, p. 1099–1108, Dec 2011. ISSN 0885-8969.
- GOMEZ-EXPOSITO, A.; CONEJO, A. J.; CANIZARES, C. **Sistemas de Energia Elétrica: Análise e Operação**. 1. ed. Rio de Janeiro: LTC, 2011. 572 p. ISBN 978-85-216-1802-7.
- GUO, T.; MILANOVI, J. V. Online identification of power system dynamic signature using pmu measurements and data mining. **IEEE Transactions on Power Systems**, v. 31, p. 1760–1768, Maio 2016.
- Hatziargyriou, N. et al. Definition and classification of power system stability revisited extended. **IEEE Transactions on Power Systems**, p. 1–1, 2020.

- HENRIQUES, R. M. **Utilização de Autovalores e Autovetores no Problema de Fluxo de Potência para Determinação de Áreas de Controle de Tensão**. Tese (Doutorado), 2009.
- HUANG, N. et al. Power quality disturbances feature selection and recognition using optimal multi-resolution fast s-transform and cart algorithm. **Energies**, Multidisciplinary Digital Publishing Institute, v. 9, n. 11, p. 927, 2016.
- KALYANI, S.; SWARUP, K. S. Binary SVM approach for security assessment and classification in power systems. **Proceedings of INDICON 2009 - An IEEE India Council Conference**, IEEE, p. 11–14, 2009.
- KRISHNAN, V. et al. Efficient database generation for decision tree based power system security assessment. **IEEE Transactions on Power Systems**, v. 26, n. 4, p. 2319–2327, 2011.
- KUNDUR, P. **Power System Stability and Control**. [S.l.]: McGraw-Hill, 1994.
- _____. **Power System Stability and Control**. New York, NY: McGraw-Hill, 1994.
- KUNDUR, P. et al. Definition and classification of power system stability iee/cigre joint task force on stability terms and definitions. **Power Systems, IEEE Transactions on**, v. 19, n. 3, p. 1387 – 1401, aug. 2004. ISSN 0885-8950.
- KUNDUR, P.; ROGERS, G.; KLEIN, M. A fundamental study of inter-area oscillations in power systems. **IEEE Transactions on Power Systems**, IEEE, p. 914–920, 1991.
- LEANDRO, R. B. et al. Identification of the oscillation modes of a large power system using ambient data. v. 26, n. 4, p. 441–453, 2015. ISSN 2195-3899. Disponível em: <<https://doi.org/10.1007/s40313-015-0187-1>>.
- LIU, C. et al. A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees. **IEEE Transactions on Power Systems**, v. 29, n. 2, p. 717–730, 2014.
- LIU, S. et al. An Integrated Scheme for Online Dynamic Security Assessment Based on Partial Mutual Information and Iterated Random Forest. **IEEE Transactions on Smart Grid**, v. 11, n. 4, p. 3606–3619, 2020. ISSN 19493061.
- MANSOUR, M. R.; ALBERTO, L. F. C.; RAMOS, R. A. Preventive control design for voltage stability considering multiple critical contingencies. **IEEE Transactions on Power Systems**, v. 31, n. 2, p. 1517–1525, 2016.
- MARCO, F. J. D.; MARTINS, N.; FERRAZ, J. C. R. An automatic method for power system stabilizers phase compensation design. **IEEE Transactions on Power Systems**, v. 28, n. 2, p. 997–1007, May 2013. ISSN 0885-8950.
- MARTINS, N. Efficient eigenvalue and frequency response methods applied to power system small-signal stability studies. **Power Engineering Review, IEEE**, PER-6, n. 2, p. 47–47, 1986. ISSN 0272-1724.
- MASSERAN, T.; GOMES, S.; AMARAL, T. Análise de contingências e monitoração de oscilações no pacdyn para avaliação de segurança dinâmica de sistemas de potência. In: **XXIV SNPTEE**. curitiba, Brasil: [s.n.], 2017. p. 1–9.

MCNABB, P.; WILSON, D.; BIALEK, J. Classification of model damping and amplitude in power systems using synchrophasor measurements and classification trees. **IEEE Transactions on Power Systems**, v. 28, p. 1988–1996, Maio 2013.

Morison, K.; Lei Wang; Kundur, P. Power system security assessment. **IEEE Power and Energy Magazine**, v. 2, n. 5, p. 30–39, 2004.

MUKHERJEE, R.; DE, A. Development of an Ensemble Decision Tree-Based Power System Dynamic Security State Predictor. **IEEE Systems Journal**, IEEE, v. 14, n. 3, p. 3836–3843, 2020. ISSN 19379234.

OLIVEIRA, W. D. et al. Power system security assessment for multiple contingencies using multiway decision tree. **Electric Power Systems Research**, v. 148, p. 264–272, 2017. ISSN 03787796.

PATIDAR, P.; TIWARI, A. Handling Missing Value in Decision Tree Algorithm. **International Journal of Computer Applications**, v. 70, n. 13, p. 31–36, 2013.

QUINLAN, J. R. **J. Ross Quinlan_C4.5_ Programs for Machine Learning.pdf**. 1993. 302 p.

SAJADI, A.; PREECE, R.; MILANOVIĆ, J. V. Evaluation of Suitability of Different Transient Stability Indices for Identification of Critical System States. **arXiv**, 2020.

SENE SOULIN, F.; HONGESOMBUT, K.; DECHANUPAPRITTHA, S. Estimation of Dominant Power Oscillation Mode using LSTM-RNN based on Synchrophasor Data. **Proceedings - 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe, IEEEIC/I and CPS Europe 2019**, 2019.

SUBRAMANI, C. et al. Artificial neural network based voltage stability analysis in power system. **Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2016**, IEEE, p. 16–19, 2016.

SULLA, F. et al. Statistical analysis and forecasting of damping in the nordic power system. **IEEE Transactions on Power Systems**, v. 30, n. 1, p. 306–315, Jan 2015. ISSN 0885-8950.

TEEUWSEN, S.; ERLICH, I.; EL-SHARKAWI, M. Small-signal stability assessment based on advanced neural network methods. In: **2003 IEEE Power Engineering Society General Meeting**. [S.l.: s.n.], 2003. v. 4, p. 2349–2649.

TELEGINA, E. Impact of Rotational Inertia Changes on Power System Stability. p. 1–104, 2015. Disponível em: <<https://www.ethz.ch/content/dam/ethz/special-interest/itet/institute-eeh/power-systems-dam/documents/SAMA/2015/Telegina-MA-2015.pdf>>.

THEODORO, E. A. R. **Contribuição à análise de estabilidade transitória, em duas escalas de tempo, de sistemas elétricos de potência via métodos diretos**. Tese (Tese (Doutorado em Sistemas Elétricos de Potência)) — Universidade de São Paulo -USP, São Carlos, SP, 2017.

- WANG, T. et al. Decision tree based online stability assessment scheme for power systems with renewable generations. **CSEE Journal of Power and Energy Systems**, v. 1, n. 2, p. 53–61, 2015.
- WERBESTON, D. O. **Avaliação da Segurança de Sistemas de Potência para Múltiplas Contingências Usando Árvore de Decisão**. Tese (Doutorado), 2017.
- WILCOX, R. R. **Understanding and Applying Basic Statistical Methods Using R**. [S.l.]: Wiley, 2016.
- YU, J. J. et al. Intelligent time-adaptive transient stability assessment system. **IEEE Transactions on Power Systems**, IEEE, v. 33, n. 1, p. 1049–1058, 2018. ISSN 08858950.
- Zhang, J. et al. Instantaneous electromechanical dynamics monitoring in smart transmission grid. **IEEE Transactions on Industrial Informatics**, v. 12, n. 2, p. 844–852, April 2016. ISSN 1941-0050.
- ZHOU, N. et al. Electromechanical mode shape estimation based on transfer function identification using PMU measurements. **2009 IEEE Power and Energy Society General Meeting, PES '09**, IEEE, p. 1–7, 2009.
- _____. Dynamic state estimation of a synchronous machine using pmu data: A comparative study. **IEEE Transactions on Smart Grid**, v. 6, n. 1, p. 450–460, Jan 2015. ISSN 1949-3053.