



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ELETRÔNICA E SISTEMAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

DOUGLAS CONTENTE PIMENTEL BARBOSA

**DIAGNÓSTICO DE FALHAS ESTRUTURAIS EM HASTES DE ÂNCORA POR
MEIO DA ANÁLISE DE SINAIS ELETROMAGNÉTICOS POR FERRAMENTAS DE
INTELIGÊNCIA ARTIFICIAL**

Recife

2020

DOUGLAS CONTENTE PIMENTEL BARBOSA

**DIAGNÓSTICO DE FALHAS ESTRUTURAIS EM HASTES DE ÂNCORA POR
MEIO DA ANÁLISE DE SINAIS ELETROMAGNÉTICOS POR FERRAMENTAS DE
INTELIGÊNCIA ARTIFICIAL**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Comunicações.

Orientador: Prof^o. Dr. Marcos Tavares de Melo.

Recife

2020

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

B238p Barbosa, Douglas Contente Pimentel.
Diagnóstico de falhas estruturais em hastes de âncora por meio da análise de sinais eletromagnéticos por ferramentas de inteligência artificial / Douglas Contente Pimentel Barbosa. - 2020.
141 folhas, il., gráfs., tabs.

Orientador: Prof. Dr. Marcos Tavares de Melo.
Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2020.
Inclui Referências.

1. Engenharia Elétrica. 2. Detecção de falhas. 3. Inteligência artificial. 4. Aprendizado de máquina. 5. Parâmetros eletromagnéticos. 6. Linhas de transmissão de energia elétrica. I. Melo, Marcos Tavares de (Orientador). II. Título

UFPE

621.3 CDD (22. ed.) BCTG/2020-226

DOUGLAS CONTENTE PIMENTEL BARBOSA

**DIAGNÓSTICO DE FALHAS ESTRUTURAIS EM HASTES DE ÂNCORA POR
MEIO DA ANÁLISE DE SINAIS ELETROMAGNÉTICOS POR FERRAMENTAS DE
INTELIGÊNCIA ARTIFICIAL**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Engenharia Elétrica.

Aprovada em: 14 de agosto de 2020.

BANCA EXAMINADORA

Prof^o. Dr. Marcos Tavares de Melo (Orientador)
Universidade Federal de Pernambuco

Prof^o. Dr. Luiz Henrique Alves de Medeiros (Examinador Externo)
Universidade Federal de Pernambuco

Prof^o. Dr. Lauro Rodrigo Gomes da Silva Lourenço Novo (Examinador Externo)
Universidade Federal de Pernambuco

Prof^o. Dr. Bruno Gomes Moura de Oliveira (Examinador Externo)
Instituto Federal de Pernambuco

Prof^a. Dr^a. Tatiana Escovedo (Examinadora Externa)
Petrobras / Pontifícia Universidade Católica do Rio de Janeiro

Dedico este trabalho ao meu filho Caetano. Busque todos os dias a verdade e o saber. Que sua vida seja repleta de descobrimentos e que você tenha sempre a certeza de que só com conhecimento nós podemos ser verdadeiramente livres.

AGRADECIMENTOS

Há um provérbio africano que diz “Se quiser ir rápido, vá sozinho. Se quiser ir longe, vá em grupo”. Os seres humanos são criaturas capazes de grandes realizações quando agem juntos. Por esse motivo, não considero este um resultado individual, mas sim o fruto das contribuições de diversas pessoas incríveis com as quais tive a oportunidade e o privilégio de conviver, e pelas quais sou muito grato. Assim, não poderia deixar de demonstrar explicitamente minha gratidão a todos aqueles que, das mais diversas formas, contribuíram, direta ou indiretamente, para a realização deste trabalho.

Agradeço à minha família pelos fundamentos de tudo que eu sou. Aos meus pais, Ramos e Roseneide pelo amor, apoio, estímulo, exemplo e por todos os valores que me foram passados, dentre eles o de que a educação transforma. Espero ser tão competente quanto vocês na formação de meus filhos. Agradeço ao meu irmão Diogo e a Mariana, uma irmã que a vida me presenteou, por, mesmo estando geograficamente distantes, se manterem sempre presentes.

À minha esposa, Yllen. Seria impossível expressar em palavras o quanto eu devo a você. Este é apenas o mais recente passo de uma longa jornada na qual você sempre esteve ao meu lado, e que não aconteceria sem o apoio da incrível companheira que você é. Obrigado pelas contribuições, por entender minha ausência nesse período e por se mostrar, a cada instante, a melhor escolha que fiz em minha vida, sobretudo nos momentos mais difíceis. Obrigado por ser uma mãe extraordinária e por me dar a segurança de saber que, nesses tempos tão turbulentos, nosso filho foi cuidado e amado por alguém tão único e maravilhoso. Amo você!

Ao meu filho, Caetano, que mesmo com tão pouca idade já me ensinou tantas coisas e me apresentou a um novo mundo desde sua chegada em nossas vidas. Obrigado por me permitir descobrir e sentir o mais profundo amor. Sempre saiba que o papai te ama muito.

À minha família herdada, Laietes (*in memoriam*), Ivone, Livia, Cristhiane, Rodrigo, Maria Clara e Rodriguinho. Obrigado por terem me acolhido com todo amor e carinho e por sempre terem me apoiado e torcido pelo meu crescimento. Muito obrigado.

Agradeço a todos os amigos que a vida me concedeu, e pelos quais tenho enorme apreço e carinho. Considero que sou uma pessoa privilegiada por sempre ter tido pessoas tão especiais ao meu redor, para juntos dividir os pesares e comemorar as vitórias. Forte abraço em todos. Agradeço ao meu orientador, professor Marcos Tavares, pela confiança e disponibilidade em me orientar, e pelas produtivas discussões dos mais diversos assuntos no decorrer dessa jornada. Obrigado por conseguir extrair o melhor de mim. Meus sinceros agradecimentos ao amigo e

professor Lauro Novo pelo convite para integrar esse projeto. Essa oportunidade transformou completamente a minha vida. Agradeço ao professor Luiz Henrique por todo o apoio recebido, pelos conselhos, pela troca de conhecimentos, pela confiança e pelo respeito. Ao professor e amigo Bruno Oliveira, e à professora e colega Tatiana Escovedo pelos comentários, sugestões e contribuições valiosas. Muito obrigado a todos vocês.

Aos colegas de pesquisa, em especial a Marcelo Coutinho, Marcelo Alves, Renan, Vinicius, Bruno, Jorge, Pedro, Hawson e José Mário, pela parceria, pela troca de conhecimentos e pelos momentos de descontração nas salas de aula, laboratórios, congressos, viagens ou durante as refeições de procedência duvidosa. À Cinthia, pela amizade e dedicação ao nosso trabalho. Muito obrigado por tornarem essa luta mais leve e por me fazer lembrar o enorme prazer que é ser estudante novamente. Levarei sempre todos vocês em meu coração.

Agradeço a Petrobras pela compreensão e flexibilidade que me foi concedida para que eu pudesse me dedicar ao doutorado enquanto realizava minhas atividades profissionais. Agradeço ao programa de pesquisa e desenvolvimento da Agência Nacional de Energia Elétrica - ANEEL, a Sistema de Transmissão Nordeste - STN S/A e ao Instituto Avançado de Tecnologia e Inovação - IATI pelo apoio financeiro e logístico dados a esse projeto, em colaboração com a Universidade Federal de Pernambuco - UFPE. Agradeço aos professores e funcionários do Programa de Pós-Graduação em Engenharia Elétrica - PPGE, pelo conhecimento adquirido, e por todo o apoio dado para a realização e divulgação dos resultados deste trabalho. Meus sinceros agradecimentos.

Oriundo de avós semianalfabetos e pais que decidiram romper o ciclo da pobreza por meio do estudo, eu agora posso me orgulhar por ter continuado o caminho desbravado por eles e conquistado no intervalo de três gerações, com muito esforço e humildade, o mais alto grau acadêmico. Assim, pude observar na prática, em minha vida, a máxima do pensador pernambucano Paulo Freire: “A educação não muda o mundo. A Educação muda as pessoas, e as pessoas transformam o mundo”. No entanto, digo sem hesitar que esse caminho de transformação não poderia ter sido percorrido por mim sem a contribuição de cada um de vocês.

Muito obrigado a todos!

Douglas Contente Pimentel Barbosa

“We can only see a short distance ahead, but we can see plenty there that needs to be done” (Alan Turing, 1912-1954).

“Nós só podemos ver um pouco do futuro, mas o suficiente para perceber que há muito a fazer” (Alan Turing, 1912-1954).

RESUMO

As hastes de âncora são elementos estruturais que garantem a estabilidade das torres estaiadas de linhas de transmissão de energia elétrica. No entanto, falhas durante a instalação dessas hastes, associadas à má qualidade do solo, podem levar à sua corrosão. Os danos gerados na estrutura dessas hastes por tais processos corrosivos são um problema crítico que, em casos extremos, pode ser capaz de provocar a queda da torre. Por esse motivo, a ação de monitorar e manter a integridade estrutural das hastes de âncora é um fator chave para evitar acidentes e a interrupção de serviço de fornecimento de energia elétrica. De um lado, o método tradicional de verificar a condição dessas estruturas é complexo, perigoso, dispendioso e demorado, pois requer a escavação das âncoras para realização de inspeção visual nas hastes. Por outro lado, é sabido que parâmetros eletromagnéticos são capazes de transportar informação sobre a constituição de um meio no qual uma onda de alta frequência se propaga. Neste trabalho, essa característica é explorada com o objetivo de propor um sistema inteligente para detecção de falhas estruturais em hastes de âncora. Uma abordagem por aprendizado de máquina é apresentada para extração de características subjacentes dos dados e estabelecer relações entre esses parâmetros e a presença de falhas nas hastes, sem a necessidade de modelos físicos. Um banco de dados híbrido foi construído com amostras medidas e simuladas dos parâmetros perda de retorno, impedância de entrada e coeficiente de onda estacionária de tensão para vários exemplos de hastes normais e defeituosas para treinamento dos modelos. Análises comparativas foram realizadas para avaliação da configuração mais adequada dos classificadores baseados em aprendizado de máquina propostos para detecção das falhas nessas hastes. Os resultados experimentais obtidos mostram que a estratégia inovadora proposta neste trabalho, de utilizar algoritmos de aprendizado de máquina na análise de parâmetros eletromagnéticos para o diagnóstico de falhas estruturais em hastes de âncora, supera o desempenho obtido pelas metodologias similares conhecidas e possibilita o desenvolvimento de um sistema de detecção viável e de alta confiabilidade para uso em campo.

Palavras-chave: Detecção de falhas. Inteligência artificial. Aprendizado de máquina. Parâmetros eletromagnéticos. Linhas de transmissão de energia elétrica.

ABSTRACT

Anchor rods are structural solutions to provide stability for guyed power transmission lines towers. However, fails during installation and poor quality of the soil can lead to the corrosion of such structures. The damage caused by corrosion processes in anchor rods is a critical problem that, in extreme cases, may provoke the fall of the tower. For this reason, monitoring and keeping the structural integrity of such anchors is a key factor in avoiding accidents and the interruption of the power delivery service. On one hand, the traditional method to assess the anchor condition is complex, risky, expensive and time-consuming, because it requires excavation and visual inspection of the rods. On the other hand, is known that electromagnetic parameters are capable of conveying information about the constitution of a medium in which a high-frequency wave propagates. In this work, such characteristic is exploited in order to propose an intelligent system to detect structural faults on the anchor rods. A machine learning approach is presented to perform the automatic capture of underlying characteristics from the data and to establish relationships between such parameters and the presence of faults in the rods, without the need for physical models. A compound database is built with both simulated and measured samples of the parameters return loss, input impedance, and voltage standing wave ratio, for several normal and faulty rods examples, in order to train the models. A series of comparative analysis is performed to evaluate the best setup for the machine learning classifiers used to detect the faults on the anchor rods. The experimental results showed that the innovative strategy proposed in this work, of using machine learning algorithms to process electromagnetic parameters in order to detect structural faults in anchor rods, overcomes the accuracies achieved by other known methodologies and allows the development of a feasible and reliable fault detection system.

Keywords: Fault detection. Artificial intelligence. Machine learning. Electromagnetic parameters. Power transmission lines.

LISTA DE FIGURAS

Figura 1 – Tipos de torres mais comumente utilizadas em linhas de transmissão. Autoportante (a). Estaiada em V (b).	21
Figura 2 – Esquema de instalação da haste de âncora. Em rocha (a). Em solo (b).	21
Figura 3 – Esquema simplificado do processo de corrosão em hastes de âncora.	22
Figura 4 – Queda de torre estaiada de linha de transmissão (a). Procedimento tradicional de inspeção visual por escavação da haste de âncora (b).	23
Figura 5 – Exemplos de linhas de transmissão de altas frequências. Par de fios paralelos (a). Linha de microfita (b). Cabo coaxial (c).	33
Figura 6 – Representação de um trecho de uma linha de transmissão por um par de fios paralelos.	33
Figura 7 – Modelagem de um trecho infinitesimal (Δz) de uma linha de transmissão.	34
Figura 8 – Linha de transmissão terminada em uma carga Z_L	36
Figura 9 – Exemplos de modelos com boa generalização e com ocorrência de subajuste e sobreajuste para os casos de classificação e regressão.	42
Figura 10 – Exemplo de construção da hiperesfera do algoritmo K-NN para $K=5$	46
Figura 11 – Esquema de classificação de um novo ponto por árvore de decisão.	47
Figura 12 – Conceitos de hiperplano de separação, margem e vetores de suporte do SVM.	51
Figura 13 – Função sigmóide ou logística.	56
Figura 14 – Esquema simplificado da estrutura de um neurônio biológico.	58
Figura 15 – Modelo esquemático de um neurônio artificial.	59
Figura 16 – Exemplos de funções de ativação.	60
Figura 17 – Estrutura básica de uma Rede Neural Artificial tipo MLP.	61
Figura 18 – Fluxo de informação no sentido direto da rede FF-MLP.	63
Figura 19 – Exemplo de funcionamento do método do Gradiente Descendente.	64
Figura 20 – Fluxo de informação no sentido reverso para ajuste dos pesos sinápticos da rede FF-MLP no algoritmo de <i>backpropagation</i>	65
Figura 21 – Esquema do processo de redução de dimensionalidade com PCA.	69
Figura 22 – Esquema simplificado do comportamento da onda eletromagnética para os casos da haste normal (a) e defeituosa (b).	71
Figura 23 – Detalhe da conexão entre a haste de âncora e haste de referência.	73
Figura 24 – Conjunto de hastes de teste confeccionadas para medição.	74

Figura 25 – Modelo computacional de simulação das hastes experimentais utilizadas.....	77
Figura 26 - Formato da base de dados de parâmetros eletromagnéticos utilizada para treinamento e teste dos algoritmos de aprendizado de máquina.	79
Figura 27 – Critério utilizado para classificar a haste como normal ou defeituosa.....	80
Figura 28 – Características de composição da base de dados utilizada.....	80
Figura 29 – Visão geral dos componentes do sistema de diagnóstico de falhas em hastes de âncora.	82
Figura 30 – Conjunto de hastes do Campo Experimental da UFPE.....	83
Figura 31 – Conectores em desenvolvimento para realização de medições em campo.	84
Figura 32 – Arquitetura do sistema proposto para avaliação de parâmetros eletromagnéticos.....	85
Figura 33 – Estrutura geral das RNAs no estágio de classificação.	89
Figura 34 – Esquema de teste utilizado para análise comparativa entre algoritmos.	91
Figura 35 – Comparação entre os sinais medido e simulado da magnitude de S_{11} para uma haste normal (H1N) na faixa de frequências de 2 MHz a 1 GHz.....	94
Figura 36 – Detalhe do parâmetro Magnitude de S_{11} medido para as hastes H1N, H1D15, H1D25 e H1D65 nas vizinhanças dos 04 principais pontos de ressonância.....	96
Figura 37 – Distribuição das frequências de ressonância.....	96
Figura 38 – Comparação entre os sinais medido e simulado das partes real e imaginária do Z_{in} (normalizado por Z_0) para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15).....	98
Figura 39 – Comparação entre os sinais medido e simulado da magnitude e fase de S_{11} para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15).....	99
Figura 40 – Comparação entre os sinais medido e simulado do VSWR para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15), com detalhe em torno dos pontos de ressonância.....	100
Figura 41 – Distribuição dos componentes principais da magnitude de S_{11}	101
Figura 42 – Distribuição dos componentes principais da fase de S_{11}	102
Figura 43 – Distribuição dos componentes principais do VSWR.	102
Figura 44 – Distribuição dos componentes principais da parte real de Z_{in}	102
Figura 45 – Distribuição dos componentes principais da parte imaginária de Z_{in}	103
Figura 46 – Acurácia do classificador multiparâmetro em função do índice de combinação.....	104

Figura 47 – Análise comparativa do desempenho do modelo proposto com referências encontradas na literatura em função do algoritmo de aprendizado de máquina utilizado.	107
Figura 48 – Importância relativa das componentes de frequência no processo de detecção das falhas estruturais nas hastes de âncora pelo algoritmo <i>random forests</i>	110
Figura 49 – Detalhamento do grau de importância das bandas de frequência de classificação nas vizinhanças pontos de ressonância.....	111
Figura 50 – Resposta simulada da $\text{Mag}\{S_{11}\}$ para diferentes comprimentos de hastes sem falhas.	112
Figura 51 – Esquema de avaliação da arquitetura de múltiplos parâmetros utilizando o modelo <i>random forests</i>	113
Figura 52 – Acurácia obtida pela arquitetura de múltiplos parâmetros utilizando o modelo <i>random forests</i>	115
Figura 53 – Efeito da redução de dimensionalidade na acurácia e tempo de processamento observados pelo modelo <i>random forests</i>	115
Figura 54 – Arquitetura proposta para o sistema de diagnóstico.	118
Figura 55 – Medições de campo realizadas em Linha de Transmissão no município de Sobral – CE.	121
Figura 56 – Protótipo conceitual do sistema embarcado de diagnóstico com a disposição futura dos equipamentos na maleta compacta de transporte (fora de escala). ..	122
Figura 57 – Adaptação da metodologia proposta para detecção de falhas em parafusos olhais de postes de concreto. Medição em bancada (a) e (b). Parafusos olhais em serviço (c).	123
Figura 58 – Visões do modelo em desenvolvimento para aplicação do método proposto na detecção de corrosão em dutos de transporte de óleo e gás.	123
Figura 59 – Diagramas de radiação de arranjo de antenas Yagi-Uda obtidos com otimização por PSO. 0° – 12.98 dBi (a). 30° – 10.89 dBi (b). 60° – 12.98 dBi (c). 90° – 13.26 dBi (d).....	124
Figura 60 – Protótipo de sistema para detecção rápida de vírus por absorção eletromagnética. Visão explodida do sensor (a). Esquema básico de funcionamento (b).....	125

LISTA DE TABELAS

Tabela 1 – Quatro categorias de definições de inteligência artificial.....	40
Tabela 2 – Configuração das hastes de âncora de teste utilizadas.....	75
Tabela 3 – Características dos modelos computacionais simulados.	77
Tabela 4 – Relações entre os índices de combinação e os parâmetros eletromagnéticos.	87
Tabela 5 – Hiperparâmetros selecionados para as RNAs do classificador multiparâmetro.	90
Tabela 6 – Hiperparâmetros escolhidos para cada um dos algoritmos.....	91
Tabela 7 – Indicadores utilizados para avaliação do desempenho dos algoritmos.	92
Tabela 8 – Correlação entre os sinais medidos e simulados da haste normal H1N.	95
Tabela 9 – Características do Sinal de Magnitude do S_{11} para a haste normal (H1N).....	97
Tabela 10 – Características do Sinal de Magnitude do S_{11} para a haste defeituosa (H1D15)..	97
Tabela 11 – Parâmetros de distribuição das frequências de ressonância.	100
Tabela 12 – Componentes principais requeridos de cada parâmetro de modo a reter 99% da variância.	101
Tabela 13 – Análise comparativa do desempenho obtidos pelas combinações de parâmetros.	105
Tabela 14 – Análise comparativa do desempenho dos algoritmos de aprendizado de máquina.	106
Tabela 15 – Desempenhos obtidos pelas diferentes combinações de parâmetros eletromagnéticos de entrada utilizando o modelo <i>random forests</i>	114

LISTA DE ABREVIATURAS E SIGLAS

ADAM	Adaptive Moment Estimation
ARV	Analisador de Redes Vetorial
BP	Banda Primária
BS	Banda Secundária
CHAAF	Conector Haste-Analisador para Altas Frequências
CNN	Convolutional Neural Network
COV-2	Coronavirus 2
COVID-19	Coronavirus 2019
CSV	Comma Separated Values
CTG	Centro de Tecnologia e Geociências
DBF	Deep Belief Networks
DL	Deep Learning
DS	Differential Search
ELM	Extreme Learning Machines
FF-MLP	Feed-Forward Multi-Layered Perceptron
FSS	Frequency Selective Surface
GAN	Generative Adversarial Network
HDD	Hard Disk Drive
HFSS	High Frequency Structure Simulator
IA	Inteligência Artificial
IOT	Internet of Things
K-NN	K-Nearest Neighbors
MDSC	Microwave Device of Supporting and Connection
MFL	Magnetic Flux Leakage
MLP	Multi-Layered Perceptron
OPGW	Optical Ground Wire
PCA	Principal Component Analysis
PH	Potencial de Hidrogênio
PSO	Particle Swarm Optimization
RELU	Rectified Linear Unit
RF	Radiofrequência
RNA	Rede Neural Artificial

RNN	Recurrent Neural Network
SARS	Severe Acute Respiratory Syndrome
SVM	Support Vector Machines
TEM	Transverse Electromagnetic
VSWR	Voltage Standing Wave Ratio

LISTA DE SÍMBOLOS

α	Constante de amortecimento
β	Número de onda
ε	Permissividade elétrica em determinado meio
ε_r	Constante dielétrica
ε_0	Permissividade elétrica no vácuo
δ_s	Profundidade de penetração
λ	Comprimento de onda referente à frequência
μ	Permeabilidade magnética em determinado meio
μ_0	Permeabilidade magnética no vácuo
ω	Frequência angular
c	Velocidade da luz no vácuo
γ	Constante de propagação complexa
σ	Condutividade
σ^2	Variância
η	Taxa de aprendizado
Γ	Coefficiente de reflexão
Σ	Matriz de covariância
Λ	Matriz diagonal de autovalores
C	Capacitância por unidade de comprimento
\bar{E}	Campo Elétrico
f	frequência
f_n	False negative (Falso negativo)
f_p	False positive (Falso positivo)
G	Condutância por unidade de comprimento
\bar{H}	Campo Magnético
L	Indutância por unidade de comprimento
R	Resistência por unidade de comprimento
S_{ij}	Parâmetros de espalhamento de uma onda se propagando da porta j para a i
t_n	True negative (Verdadeiro negativo)
t_p	True positive (Verdadeiro positivo)
v_p	Velocidade de fase

Z_0	Impedância característica de uma linha de transmissão
Z_{in}	Impedância de entrada de uma linha de transmissão
Z_L	Impedância de carga de uma linha de transmissão

SUMÁRIO

1	INTRODUÇÃO.....	20
1.1	CARACTERIZAÇÃO E RELEVÂNCIA DO PROBLEMA	20
1.2	REVISÃO DA LITERATURA E ESTADO DA ARTE.....	24
1.3	MOTIVAÇÃO.....	28
1.4	OBJETIVOS	30
1.5	ESTRUTURA DA TESE	31
2	PARÂMETROS ELETROMAGNÉTICOS DE ALTAS FREQUÊNCIAS	32
2.1	LINHAS DE TRANSMISSÃO DE ALTAS FREQUÊNCIAS	32
2.2	IMPEDÂNCIA DE ENTRADA.....	36
2.3	COEFICIENTE DE ONDA ESTACIONÁRIA DE TENSÃO.....	37
2.4	PARÂMETROS S	38
3	INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	39
3.1	K-NEAREST NEIGHBORS	44
3.2	RANDOM FORESTS	46
3.3	SUPPORT VECTOR MACHINES.....	50
3.4	REGRESSÃO LOGÍSTICA.....	55
3.5	REDES NEURAIS ARTIFICIAIS.....	57
3.6	ANÁLISE DE COMPONENTES PRINCIPAIS.....	67
4	DESENVOLVIMENTO DO SISTEMA DE DIAGNÓSTICO	71
4.1	MODELOS COMPUTACIONAIS SIMULADOS.....	76
4.2	ESTRUTURA DA BASE DE DADOS.....	77
4.3	DESENVOLVIMENTO DAS APLICAÇÕES	81
4.4	MODO DE OPERAÇÃO DO SISTEMA	82
4.5	INFRAESTRUTURA DISPONÍVEL PARA O PROJETO.....	82
5	EXPERIMENTOS.....	85
5.1	DESENVOLVIMENTO DO CLASSIFICADOR MULTIPARÂMETRO	85
5.2	AVALIAÇÃO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA.....	90
6	RESULTADOS E DISCUSSÕES.....	94
6.1	ANÁLISE DOS SINAIS MEDIDOS E SIMULADOS DA BASE DE DADOS	98
6.2	AVALIAÇÃO DO CLASSIFICADOR MULTIPARÂMETRO	104
6.3	ANÁLISE COMPARATIVA DOS ALGORITMOS DE CLASSIFICAÇÃO	105

7	CONCLUSÕES.....	116
7.1	TRABALHOS FUTUROS: EVOLUÇÃO DO MÉTODO	120
7.2	TRABALHOS FUTUROS: EXTENSÕES	122
7.3	CONTRIBUIÇÕES E PRODUÇÃO CIENTÍFICA.....	125
7.3.1	Trabalhos publicados em periódicos.....	125
7.3.2	Trabalhos publicados em anais de conferências	125
7.3.3	Trabalhos aceitos em conferências aguardando publicação.....	126
	REFERÊNCIAS.....	129

1 INTRODUÇÃO

A corrosão em hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica é um problema que põe em risco o abastecimento de eletricidade e que expõe as operadoras dessas linhas a penalidades e sanções financeiras por parte das agências reguladoras no caso de ocorrência de eventuais sinistros. Por esse motivo, tais empresas devem manter um monitoramento constante do estado dessas âncoras, o que tem sido feito tradicionalmente por meio da escavação seguida da inspeção visual das hastes.

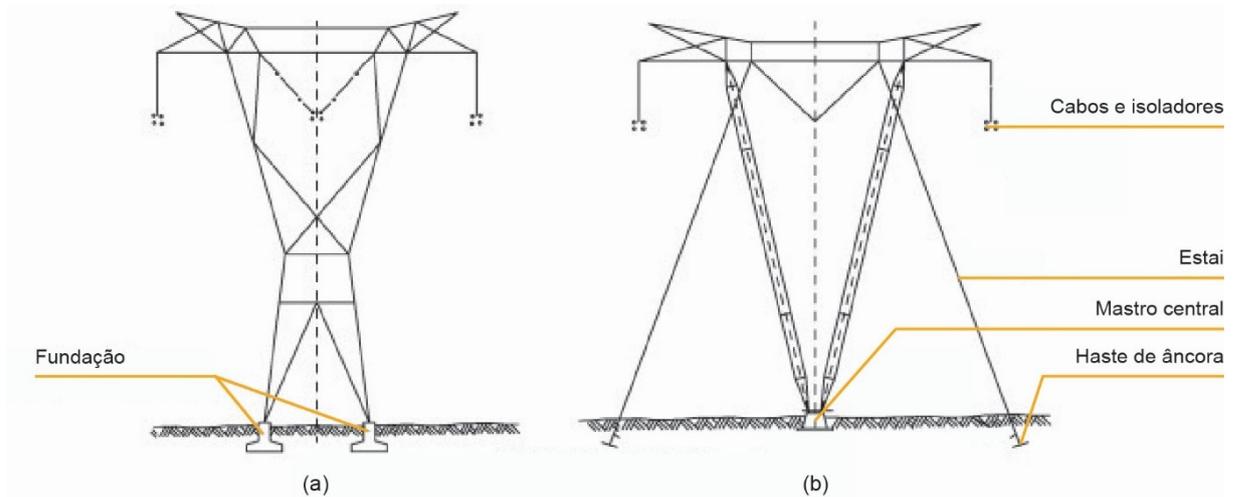
O procedimento de inspeção visual das hastes por escavação possui custo elevado pois requer a solução de desafios técnicos e logísticos como a mobilização de maquinário e de pessoal em locais de difícil acesso, construção de ancoragens temporárias durante a inspeção e a posterior recomposição da ancoragem inspecionada. Além disso, esse procedimento expõe os executantes ao risco de quedas, soterramento e impactos no caso de rompimento da haste ou desprendimento do cabo de aço que liga a âncora à torre durante o processo de inspeção.

Diante dos problemas apresentados, esta tese propõe uma nova solução para diagnóstico inteligente de falhas estruturais em hastes de âncora operacionais, que possa ser adotada como alternativa viável ao método tradicional de análise por escavação seguida de inspeção visual.

1.1 CARACTERIZAÇÃO E RELEVÂNCIA DO PROBLEMA

As linhas de transmissão de energia elétrica de alta tensão são sistemas compostos por diversos equipamentos que possuem a função de conduzir a energia produzida nos centros geradores ou transformada em uma subestação até um ponto de distribuição ou um consumidor. Essas linhas são geralmente compostas por torres metálicas responsáveis por suportar os cabos condutores de alta potência, os cabos para-raios OPGW (*Optical Ground Wire*) e as cadeias de isoladores, podendo ser classificadas nos dois tipos principais mostrados na Figura 1, quanto à forma de transferência para o solo dos esforços a que são submetidas. Enquanto nas torres autoportantes a transferência de carga para o solo é feita diretamente através de fundações, nas torres estaiadas a transferência se dá através de um conjunto de quatro cabos de aço, também chamados de estais, firmemente conectados a um sistema de âncoras instaladas sob o solo (DYWIDAG-SYSTEMS-INTERNATIONAL, 2018).

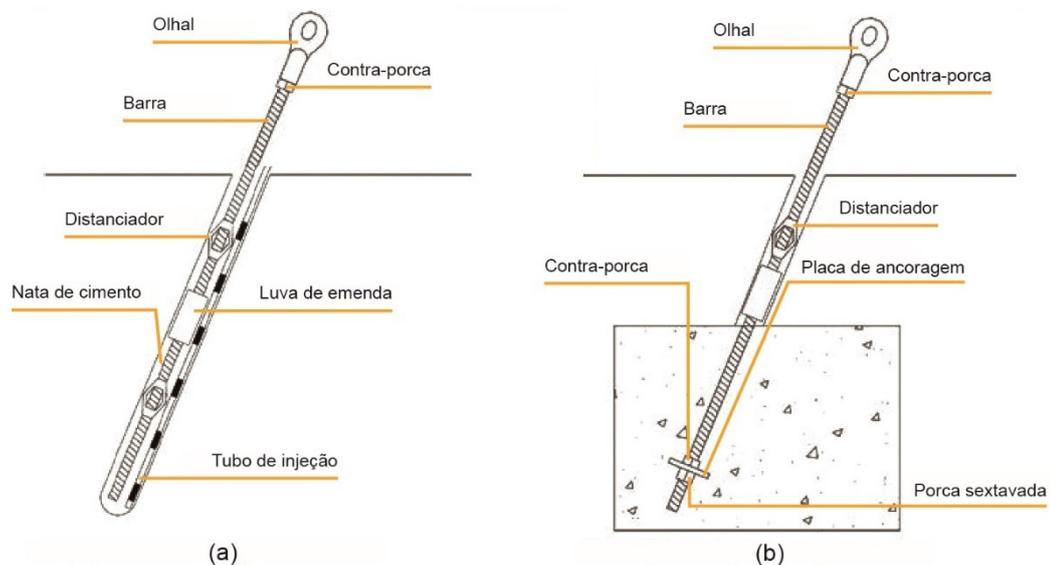
Figura 1 – Tipos de torres mais comumente utilizadas em linhas de transmissão. Autoportante (a).
Estaiada em V (b).



Fonte: Adaptado de Pacheco, Danziger e Ferreira Pinto, 2008.

Uma haste de âncora consiste basicamente em uma estrutura cilíndrica de aço reforçado, com um olhal em uma das extremidades e ranhuras ao longo de seu comprimento para aumentar sua aderência, como mostrado na Figura 2. Uma das extremidades da haste está tipicamente inserida na própria rocha ou em uma base de concreto construída no solo, enquanto a outra possui um olhal para a conexão do cabo, o que é geralmente feito por meio de um conjunto de fixação de estais do tipo VARI-GRIP ou similar (LOURENÇO NOVO, 2015).

Figura 2 – Esquema de instalação da haste de âncora. Em rocha (a). Em solo (b).

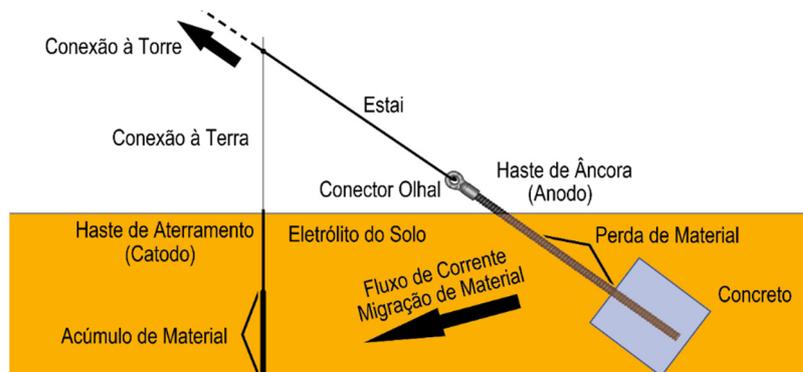


Fonte: Adaptado de Dywidag, 2018.

Por serem mais leves e utilizarem menor quantidade de estrutura metálica, o custo de instalação de uma torre estaiada é aproximadamente 15% a 20% menor do que o de uma torre autoportante, para um serviço equivalente (KUMARAVEL e TAMILVANAN, 2016). No entanto, apesar de seu baixo custo, a confiabilidade de uma torre estaiada é menor, o que se reflete na frequência de acidentes registrados para esse tipo de estrutura, sendo diversos os relatos descritos na literatura (ZAMPIRON, 2008), (BENTES, 2013). A queda de uma torre em uma linha de transmissão geralmente causa a interrupção do fornecimento de energia elétrica em uma vasta região, além do risco de impactos diretos a pessoas e ao meio ambiente. Tais problemas geram danos financeiros e de reputação significativos às companhias envolvidas, decorrentes dos custos de manutenção emergencial e das multas e sanções aplicadas pelas agências reguladoras (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, 2019).

Nesse cenário, a corrosão das hastes de âncora é uma das principais causas naturais capazes de provocar a queda de uma torre estaiada em uma linha de transmissão. Esta corrosão deve-se majoritariamente a falhas no procedimento de instalação da âncora e à baixa qualidade do terreno. Conjuntamente, esses fatores propiciam a ocorrência de um processo de corrosão galvânica, no qual a estrutura metálica da haste de âncora perde material através do eletrólito formado pelo solo, conforme mostrado na Figura 3. Esse tipo de corrosão é um problema altamente não-linear, sendo influenciado por uma combinação de características do solo e da âncora, tais como: resistividade, pH, umidade, nível de oxigênio, textura, temperatura, contaminação por sais solúveis e pela presença de micro-organismos (TAHERI, MANSOURI, *et al.*, 2017), (MATOS, PORTELLA, *et al.*, 2018).

Figura 3 – Esquema simplificado do processo de corrosão em hastes de âncora.



Fonte: O autor, 2020.

Os custos diretos totais gerados pela corrosão são estimados, em uma escala global, a cerca de 2,5 trilhões de dólares por ano, sem incluir os custos indiretos com indenizações, meio ambiente e pessoas afetadas (KOCH, 2017). Apenas no setor de transmissão e distribuição de energia elétrica, os custos diretos anuais atribuídos à corrosão totalizam 0,6 bilhões de dólares por ano, apenas nos Estados Unidos (KOCH, BRONGERS, *et al.*, 2002). No entanto, através de práticas de monitoramento e controle da corrosão, avalia-se que entre 15% e 35% desse valor pode ser economizado (KOCH, 2017). A ordem de magnitude desse custo ressalta que, por meio de investimentos adequados em sistemas de gerenciamento da corrosão, um percentual significativo das receitas de empresas e indústrias poderia deixar de ser gasto, considerando-se toda a vida útil de seus ativos.

Devido às graves consequências decorrentes da queda de uma torre, monitorar e manter a integridade estrutural de sua ancoragem são fatores-chave para evitar possíveis acidentes e seus custos associados. O método tradicional de verificação das hastes de âncora é feito através da escavação e quebra da estrutura de ancoragem para permitir a inspeção visual da haste, porém, apesar de confiável, esse método é destrutivo, demorado, de alto custo e arriscado para seus executantes. A Figura 4(a) apresenta a imagem da queda de um conjunto de torres da LT 500 kV Teresina II / Sobral III ocorridas em 2006, próximo à cidade de Piriipipi – PI. A energização do circuito de emergência após essa queda demandou 10 dias de trabalho, 100 pessoas na execução, 40 pessoas no apoio, além de uma infraestrutura de helicópteros, máquinas de elevação, tratores, veículos de carga, ambulância e veículos leves (REGIS JR, EDUARDO, *et al.*, 2007). A Figura 4(b) apresenta uma haste sendo visualmente inspecionada no campo por um trabalhador, após a sua escavação.

Figura 4 – Queda de torre estaiada de linha de transmissão (a). Procedimento tradicional de inspeção visual por escavação da haste de âncora (b).



Fonte: Adaptado de Lourenço Novo, 2015 e Regis Jr *et al.*, 2007.

A inspeção visual das hastes por meio de escavação possui uma série de inconveniências. Em primeiro lugar, a escavação das hastes é um método logisticamente complexo pois requer, em muitas situações, a mobilização de pessoal e equipamentos em locais distantes dos grandes centros urbanos e de difícil acesso, onde está instalada uma grande parcela das torres das linhas de transmissão. Nesses casos, a atividade de escavação só pode ser executada por mão de obra humana, sem o apoio de máquinas, o que torna o serviço ainda mais lento e caro. Em segundo lugar, a âncora perde a sua função de sustentação da torre enquanto está sendo escavada. Logo, antes da escavação, é necessária a construção de uma ancoragem provisória, o que aumenta ainda mais a quantidade de homens-hora necessários para a realização do trabalho e, conseqüentemente, o custo da atividade. Em terceiro lugar, o método tradicional por inspeção visual é destrutivo de um ponto de vista prático pois, na maioria das intervenções, independentemente do estado diagnosticado, a âncora inspecionada acaba sendo substituída após a escavação, por questões de confiabilidade e custo. Por fim, por ser um procedimento invasivo envolvendo escavação e movimentação de cargas majoritariamente realizados com trabalho braçal humano, há o risco constante de ocorrência de acidentes aos trabalhadores e de queda da torre, causados por eventuais falhas na execução dos serviços.

1.2 REVISÃO DA LITERATURA E ESTADO DA ARTE

A corrosão é uma das principais causas da depreciação e da redução do tempo de vida de equipamentos e instalações em todo o mundo. Especificamente na área de transmissão e distribuição, observa-se que antigas e novas linhas de transmissão de energia elétrica, em muitos casos torres estaiadas, têm sido alvos de quedas catastróficas absolutamente previsíveis, motivadas pela corrosão (SENS e SEBRÃO, 2012). Por esse motivo, o monitoramento da corrosão tem sido um desafio tecnológico constante para a indústria e um terreno fértil para a pesquisa, sendo diversos os trabalhos acadêmicos voltados a esse tema.

Algumas abordagens apresentam métodos não-destrutivos para detecção das falhas em estruturas metálicas a partir da medição das reflexões de uma onda mecânica aplicada ao material. Isso é possível pois variações na impedância acústica causada pelas falhas são capazes de refletir parte da energia da onda incidente, possibilitando, assim, a detecção. Ondas mecânicas guiadas foram utilizadas para medir, de forma não destrutiva, o comprimento de parafusos de ancoragem (LEI, LI, *et al.*, 2018), avaliar o estado de corrosão em parafusos de ancoragem (MEDFOUNI, KODJO, *et al.*, 2017), monitorar o nível de corrosão em concreto

armado (LIU, QIN e HUANG, 2018), detectar falhas e avaliar o risco de corrosão em hastes de âncora enterradas no solo (SHOJI e HIRATA, 2016), (REZAEI, CHAR, *et al.*, 2019), bem como para detectar pequenas avarias em estruturas tais como dutos, trilhos e barras metálicas (ZHANG, HAO, *et al.*, 2018). Adicionalmente, correntes de Foucault foram utilizadas para medir de forma não-destrutiva a espessura de paredes internas de estruturas industriais, como vasos, tanques e dutos sujeitos a corrosão (GUILIZZONI, FINCH e HARMON, 2019). No entanto, o processamento dos dados e a obtenção do diagnóstico pelos métodos apresentados nesses trabalhos não são realizados automaticamente, mas dependem da análise das informações por um especialista humano.

O desenvolvimento do campo da inteligência artificial disponibilizou à academia e ao mercado uma série de algoritmos e soluções para processamento e modelagem de sinais e sistemas. Aplicações de tais conhecimentos à área de eletromagnetismo têm causado uma evolução significativa nos métodos de projeto, modelagem, otimização e medição de dispositivos, componentes e sistemas que operam em altas frequências (FAN, ZHOU e CHANG, 2015), (BURRASCANO, FIORI e MONGIARDO, 1999).

Modelos de aprendizado de máquina foram utilizados em substituição aos métodos tradicionais de computação numérica na solução de diversos problemas eletromagnéticos nos domínios do tempo e frequência (AMBASANA, MUTNURY e GOPE, 2015), (DUMIN, KHMARA e SHYROKORAD, 2017), (YAO e JIANG, 2018), (KHAN, e LOWTHER, 2019), (BARMADA, FONTANA, *et al.*, 2020) e na modelagem e otimização de motores elétricos (DOI, SASAKI e IGARASHI, 2019), (SASAKI e IGARASHI, 2019), (IBRAHIM, SILVA, *et al.*, 2020), (ASANUMA, DOI e IGARASHI, 2020). Tais modelos também foram utilizados na modelagem e projeto de dispositivos (LONGFEI, ZHIHONG e VENKATESH, 2017), (JAMSHIDI, LALBAKHSB, *et al.*, 2019), na caracterização dos efeitos da vegetação em enlaces sem fio em ambientes externos (*outdoor*) (GÓMEZ-PÉREZ, CALDEIRINHA, *et al.*, 2015) e na avaliação da propagação de sinais de radiofrequência (RF) em ambientes internos (*indoor*) para aplicações de internet das coisas (*IoT – Internet of Things*) (ALHAJRI, ALI e SHUBAIR, 2018).

Redes neurais artificiais foram utilizadas na modelagem (KAPETANAKIS, VARDIAMBASIS, *et al.*, 2018), projeto (GOSAL, ALMAJALI, *et al.*, 2015) e otimização (TAK, , *et al.*, 2018) de antenas de diversas topologias, bem como para estimar a diretividade (MISHRA, YADAV e SINGH, 2015), definir a direção de chegada da onda (HARKOUSS, SHRAIM e HUSSEIN, 2018) e ajustar os coeficientes dos elementos de arranjos de antenas

inteligentes, permitindo a modificação do padrão de irradiação global e o direcionamento do feixe para múltiplos pontos do campo próximo (AYESTARÁN, 2018). No campo de sensores, algoritmos de aprendizado de máquina foram aplicados na reconstrução de imagens 3-D em sistemas de tomografia por impedância elétrica (MARTIN e CHOI, 2017), na determinação da posição de sensores magnéticos (RIBEIRO, NETO e CARDOSO, 2018), e no reconhecimento de objetos, quer seja por meio dos padrões de difração e reflexão de uma onda eletromagnética (LYASOTA, MOROZOV e MAGRO, 2015), quer seja utilizando sensores táteis magnetostrictivos (ZHANG, WANG, *et al.*, 2018), (ZHENG, WANG, *et al.*, 2019). Diferentes arquiteturas de redes neurais foram aplicadas para identificação dos movimentos da cabeça e pescoço de seres humanos a partir de mudanças no coeficiente de transmissão de um conjunto de antenas instaladas no corpo (BRESNAHAN, LI, *et al.*, 2017), bem como em detecção automática de canais magnéticos para recuperação de dados gravados em discos rígidos (HDD – *Hard Disk Drive*) (QIN e ZHU, 2020).

Abordagens que utilizam inteligência artificial para diagnosticar automaticamente falhas nos mais variados tipos de estruturas e sistemas são abundantes na literatura. Algoritmos como ELM (*Extreme Learning Machines*), SVM (*Support Vector Machines*), K-NN (*K-Nearest Neighbors*) e Árvores de Decisão (*Decision Trees*) e foram utilizadas com sucesso para localização de falhas em cabos subterrâneos (RAY, MISHRA e BUDUMURU, 2016) e de faltas em linhas de transmissão de energia elétrica (UNAL e EKICI, 2017), (CHEN, FINK e SANSAVINI, 2018), (SOWAH, DZABENG, *et al.*, 2018). Algoritmos de aprendizado de máquina e, especialmente, modelos de redes neurais artificiais, têm sido amplamente utilizados também no processamento de dados relacionados à corrosão, resultando em diversas soluções, metodologias e aplicações.

Na área industrial, redes neurais artificiais tem sido utilizadas para desenvolver sistemas de monitoramento da corrosão em unidades de destilação de petróleo cru (JAZDZEWSKA, DAROWICKI, *et al.*, 2018), para prever corrosão em equipamentos petroquímicos (LI e LIANG, 2015), na modelagem da corrosão em dutos de caldeiras submetidos a altas temperaturas e pressões (KHADOM, KARIM e FARHAN, 2018), e na avaliação de corrosão em dutos submarinos com base em características físico-químicas do processo (DE MASI, GENTILE, *et al.*, 2015).

A flexibilidade dos algoritmos de aprendizado de máquina pode ainda ser verificada pela diversidade de sinais de entrada de naturezas distintas que eles são capazes de processar. Redes neurais artificiais foram aplicadas para analisar parâmetros extraídos de ondas acústicas

(FAHAD, KAMAL, *et al.*, 2017), sinais de ultrassom (WANG e SANIIE, 2017) ou do método dos fluxos magnéticos de fuga (*magnetic flux leakage - MFL*) (GONZÁLEZ, ARIZMENDI, *et al.*, 2018), visando detectar a corrosão em tubulações industriais e detectar falhas em tubos longos ou inacessíveis (CAU, FANNI, *et al.*, 2006). Adicionalmente, esses algoritmos ainda se mostraram capazes de reconstruir de forma rápida o perfil dos defeitos medidos (FENG, LI, *et al.*, 2016) e também de estimar a extensão das falhas observadas (LU, FENG, *et al.*, 2018).

Algoritmos de aprendizado de máquina também têm sido amplamente aplicados para analisar a corrosão em sistemas e equipamentos industriais. Redes neurais artificiais foram utilizadas conjuntamente com algoritmos de árvores de decisão para estimar a taxa de corrosão do aço carbono em sistemas de águas ácidas (WU, REN, *et al.*, 2016). SVMs demonstraram ser capazes de avaliar e quantificar o risco de corrosão em uma embarcação ou planta industrial baseadas nas características e nas condições de processo dessas instalações (TAN, GOH, *et al.*, 2018). Um sistema combinado de diferentes modelos de inteligência artificial, utilizando conjuntamente redes neurais artificiais, SVM, árvores de decisão, e regressão linear foi proposto para modelar o risco de corrosão no concreto armado e a taxa de corrosão marinha do aço carbono (CHOU, NGO e CHONG, 2017).

Na área de modelagem, também são vários os exemplos de trabalhos que fazem uso de alguma estratégia de aprendizado de máquina para mapear relações entre sinais de entrada e de saída em sistemas relacionados à corrosão ou à detecção de falhas, sem a necessidade de obtenção de expressões matemáticas fechadas ou modelos físicos complexos. Redes neurais artificiais foram utilizadas para estimar a taxa de corrosão atmosférica no cobre em zonas industriais (LO, CHIU e LIN, 2017) e para prever a taxa de corrosão intragranular em ligas de alumínio processadas por fricção (HARIKESHA, SRINIVASAN, *et al.*, 2017). Redes neurais artificiais foram aplicadas ainda para detectar falhas (WOOTTON, DAY e HAYCOCK, 2018), modelar (VERMA, SENGUPTA, *et al.*, 2018) e avaliar (GUO-DONG, 2015) o grau de corrosão no aço utilizado em estruturas de concreto armado, e, com o auxílio de lógica difusa (*fuzzy logic*), avaliar a durabilidade de pontes construídas com esse material (LIN, 2017). Um sistema de detecção de danos estruturais em construções à partir da análise dos períodos e amplitudes das respostas de vibração livre dessas estruturas foi proposto com o uso de redes neurais artificiais (KAO e HUNG, 2003), (LIU, JU, *et al.*, 2011).

Por meio do processamento de imagens por redes neurais artificiais, sistemas especialistas foram desenvolvidos para classificar o tipo e severidade da corrosão baseados em imagens digitais (HATHOUT, CALLERY, *et al.*, 2017) e para realizar automaticamente a

inspeção e detecção da corrosão e de avarias no revestimento de instalações marítimas (*off-shore*) com o auxílio de drones (LIU, TAN, *et al.*, 2018).

Aplicações específicas relacionadas à modelagem, monitoramento ou detecção de falhas em elementos estruturais diversos também são abundantes. RNAs foram aplicadas para avaliar o nível de corrosão das fundações das torres de linhas de transmissão de energia elétrica a partir de parâmetros do solo (URUCHURTU-CHAVARIN, MALO-TAMAYO e HERNÁNDEZ-PEREZ, 2012). Modelos utilizando os algoritmos *K-Means* e *Elman Neural Networks* foram propostos para detecção de falhas de instalação em parafusos de ancoragem a partir do processamento de parâmetros extraídos da onda de stress gerada após um impacto recebido pela estrutura (SUN, XING, *et al.*, 2016), (SUN, GONG, *et al.*, 2019). Arquiteturas baseadas em SVM e DBF (*Deep Belief Networks*) foram propostas para detectar falhas estruturais em hastes de âncora a partir dos sinais de vibração induzidos por impacto mecânico (ZHENG, ZHANG e SUN, 2017), (ZHENG, YANG, *et al.*, 2018).

A grande quantidade de publicações recentes abordando o uso de soluções de inteligência artificial para detecção de falhas causadas por corrosão em estruturas metálicas indica que esse tópico é pertinente e atual. No entanto, apesar da extensa lista de trabalhos sobre o assunto, apenas um pequeno conjunto aborda a detecção da corrosão em hastes de âncora de torres estaiadas, um número ainda mais restrito faz uso de alguma técnica de inteligência artificial para realizar a detecção e em nenhum deles foi identificado como método de detecção a medição de parâmetros de uma onda eletromagnética como fonte do sinal a ser analisado, tendo sido reportado em todos os estudos citados somente a utilização de ondas mecânicas, tais como vibração ou ultrassom.

1.3 MOTIVAÇÃO

Diante da necessidade real de monitorar a condição das hastes de âncora das torres estaiadas das linhas de transmissão de energia elétrica e dos problemas relacionados ao método tradicional de verificação por meio de escavação e inspeção visual, se faz necessária a busca de uma solução alternativa que garanta a segurança e reduza os custos na operação dessas linhas. Nesse sentido, avaliar a condição de uma haste de âncora a partir da sua resposta a sinais eletromagnéticos seria uma alternativa que possibilitaria o diagnóstico de forma não-invasiva e não-destrutiva a partir da superfície, sem que houvesse a necessidade de escavação ou outra intervenção de qualquer natureza no sistema de ancoragem da torre.

É possível medir os parâmetros eletromagnéticos que caracterizam uma haste de âncora fazendo com que ela se comporte como uma linha de transmissão para altas frequências. Esta medição pode ser realizada com instrumentos portáteis, acoplando-se a haste de âncora a uma haste de referência por meio de um conector especialmente projetado para esse fim, conforme reportado em trabalhos anteriores (BEZERRA, DE MEDEIROS, *et al.*, 2014), (LOURENÇO NOVO, DE MELO, *et al.*, 2014), (LOURENÇO NOVO, 2015). Nessa abordagem, é possível medir as respostas no domínio da frequência de parâmetros como impedância de entrada (Z_{in}), coeficiente de onda estacionária de tensão (*voltage standing wave ratio* – *VSWR*) e perda de retorno na porta de entrada (parâmetro de espalhamento S_{11}) relativos a essas hastes.

Devido a mudanças no meio de propagação, uma onda eletromagnética incidente sofre reflexões nos pontos onde encontra as discontinuidades causadas pelas falhas na estrutura da haste, pois nelas, há transformação química do material de alguma liga metálica para algum tipo de óxido. Dessa forma, os sinais coletados apresentam diferenças nos casos da haste normal e defeituosa que podem ser exploradas para fins de diagnóstico (COUTINHO, LOURENÇO NOVO, *et al.*, 2019). No entanto, essas diferenças são sutis e, em uma avaliação inicial, não aparentam estar concentradas de forma significativamente marcante em nenhuma parte específica do espectro analisado, e sim espalhadas por todo o sinal, sendo praticamente indistinguíveis para um analista humano, ainda que treinado.

O desenvolvimento de uma aplicação específica para o processamento e classificação automática dos sinais eletromagnéticos medidos possibilitaria a detecção de possíveis falhas estruturais nas hastes de âncora, permitindo seu diagnóstico de forma não-destrutiva e não-invasiva. Nesse sentido, devido às características do problema proposto, utilizar inteligência artificial como ferramenta de análise se mostra uma alternativa adequada e viável.

Por conta da complexa geometria do sistema composto pela haste de âncora, conector e haste de referência, não seria viável buscar uma expressão matemática analítica, ainda que obtida de forma numérica ou aproximada, que fosse capaz de modelar e descrever os sinais de resposta desse sistema na presença de uma falha arbitrária. Por outro lado, uma aplicação de inteligência artificial utilizando métodos de aprendizado de máquina seria naturalmente capaz de mapear as relações subjacentes entre os sinais de entrada e as saídas desse sistema, sem a necessidade de modelos físicos formais. Essa aplicação poderia extrair automaticamente dos dados disponíveis as principais características capazes de diferenciar uma haste normal de uma haste defeituosa e, a partir do conhecimento adquirido através de um conjunto restrito de exemplos, generalizar os resultados para prever situações futuras com um alto grau de precisão.

O sistema proposto é inovador, pois, no limite do conhecimento do autor, não há na literatura, até o momento da elaboração dessa tese, nenhum registro de aplicação de técnicas de aprendizado de máquina para diagnóstico de falhas estruturais em hastes de âncora operacionais por meio do processamento de parâmetros eletromagnéticos. Nesse cenário, o sistema proposto pretende viabilizar uma alternativa prática para o diagnóstico preciso, automático, não-invasivo, não-destrutivo, de menor custo e que minimize o risco de acidente aos trabalhadores e às instalações durante a sua execução, trazendo ganhos significativos em relação ao método tradicionalmente utilizado para a inspeção dessas hastes.

1.4 OBJETIVOS

O objetivo dessa tese é propor um sistema para o diagnóstico de falhas estruturais em hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica. O sistema proposto recebe como entradas um conjunto de sinais oriundos de aquisições de parâmetros eletromagnéticos da haste em uma faixa de frequências pré-definida e produz como saída o seu diagnóstico, classificando-a primariamente como normal ou defeituosa. O processamento dos sinais adquiridos é realizado por um modelo de aprendizado de máquina (*machine learning*), capaz de extrair automaticamente dos dados um conjunto de características que permitem ao sistema distinguir, com um alto nível de confiança, os padrões de uma haste normal e de uma defeituosa, sem a necessidade de intervenções humanas na análise dessas informações. Logo, os objetivos específicos desta tese consistem em:

- a) Analisar e definir os parâmetros eletromagnéticos mais adequados para a extração das características que serão utilizadas na detecção das falhas nas hastes de âncora;
- b) Analisar e definir a faixa de frequências e resolução requeridas para a aquisição dos dados desses parâmetros;
- c) Gerar e estruturar uma base de dados representativa para realização do treinamento e dos testes de validação de desempenho dos modelos de aprendizado de máquina que serão desenvolvidos;
- d) Analisar e definir eventuais necessidades de pré-processamento dos dados coletados;
- e) Analisar e definir o conjunto de modelos de aprendizado de máquina mais adequado para o processamento dos sinais coletados;
- f) Implementar a aplicação de aprendizado de máquina para o sistema de detecção, definir seus hiperparâmetros e testar seu desempenho utilizando a base de dados gerada.

1.5 ESTRUTURA DA TESE

Este trabalho está organizado em um total de sete capítulos, incluindo este capítulo introdutório. No Capítulo 2 são apresentados os parâmetros de altas frequências que serão utilizados como sinais de entrada do sistema de diagnóstico proposto nesta tese. São introduzidos os conceitos de linhas de transmissão, impedância de entrada, coeficiente de onda estacionária e parâmetros de espalhamento para circuitos de altas frequências.

No Capítulo 3 são apresentados conceitos fundamentais em inteligência artificial. É dada ênfase nos algoritmos de aprendizado de máquina utilizados para o desenvolvimento dos classificadores binários que serão treinados para executar a tarefa de detectar falhas estruturais nas hastes de âncora. São descritos cinco dos algoritmos de aprendizado supervisionado mais utilizados e difundidos para tarefas de reconhecimento de padrões e classificação, sendo eles: *K-NN*, *Random Forests*, *SVM*, Regressão Logística e RNA. Também é descrita a técnica de análise de componentes principais (*PCA – Principal Component Analysis*), um algoritmo de aprendizado não-supervisionado utilizado para redução da dimensionalidade dos dados.

No Capítulo 4, o sistema de diagnóstico proposto é apresentado, sendo descritos o seu princípio básico de funcionamento e sua arquitetura. É apresentada a infraestrutura utilizada para desenvolvimento do sistema. São detalhados o esquema de medição, os modelos computacionais desenvolvidos e a metodologia adotada para geração da base de dados para treinamento dos classificadores por aprendizado de máquina utilizados.

No Capítulo 5 são descritos os experimentos realizados para avaliação do desempenho do sistema proposto, considerando diversas estratégias. Serão apresentados os esquemas de teste para avaliação dos parâmetros eletromagnéticos e dos algoritmos de aprendizado de máquina para definição da configuração mais indicada para solução do problema.

No Capítulo 6 são avaliados os resultados obtidos pelo sistema proposto nos testes de avaliação de desempenho realizados. As características observadas nos sinais eletromagnéticos adquiridos são apresentadas. Uma análise comparativa do desempenho do sistema proposto para cada parâmetro de entrada e para cada algoritmo de aprendizado de máquina é realizada. Finalmente, são discutidos o significado dos resultados obtidos e sua comparação com outras metodologias propostas na literatura.

O Capítulo 7 trata das conclusões desta tese. São apresentadas as considerações finais e sugestões para continuidade das investigações em trabalhos futuros. Por fim, são apresentadas as contribuições científicas geradas no desenvolvimento desta tese, listando as publicações dos resultados produzidos por este trabalho.

2 PARÂMETROS ELETROMAGNÉTICOS DE ALTAS FREQUÊNCIAS

Em circuitos de corrente contínua e baixas frequências as dimensões físicas dos dispositivos costumam tipicamente ser desprezadas para fins de análise e modelagem. Entretanto, quando se trata de redes que operam em altas frequências, essas dimensões precisam ser consideradas por influenciar de forma significativa a resposta elétrica do sistema. Nessas redes, as dimensões elétricas dos componentes do sistema são comparáveis ao comprimento de onda (λ) e as tensões e correntes variam consideravelmente em magnitude e fase ao longo da extensão física de seus elementos. Para a modelagem de tais sistemas, a teoria padrão de circuitos de Kirchoff não pode ser aplicada diretamente, e a análise necessita ser realizada por meio de uma teoria de circuitos estendida, cujos resultados sejam aderentes aos previstos pelas equações de Maxwell (POZAR, 2011).

Sistemas que operam em altas frequências são normalmente caracterizados por meio de parâmetros, dentre os quais podem ser citados como exemplos:

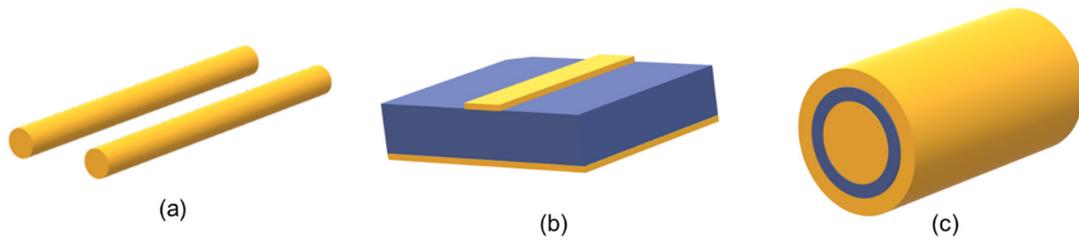
- a) Parâmetros de espalhamento (*Scattering Parameters*);
- b) Impedância de entrada (Z_{in}) e
- c) Coeficiente de onda estacionária de tensão (*VSWR*).

Cada conjunto de parâmetros oferece uma forma distinta de descrever o comportamento da rede ao ser excitada por uma onda eletromagnética de alta frequência. Dessa forma, os valores exatos desses parâmetros caracterizam a resposta do sistema e são função das interações que ocorrem entre a onda incidente e as suas reflexões no meio de propagação que pode ser, dentre outros, uma linha de transmissão de altas frequências.

2.1 LINHAS DE TRANSMISSÃO DE ALTAS FREQUÊNCIAS

Em algumas aplicações, é necessário que a energia eletromagnética seja eficientemente direcionada de um ponto a outro do circuito. Uma linha de transmissão é uma estrutura física que possibilita a transmissão guiada ponto-a-ponto da energia e da informação transportadas por uma onda eletromagnética em uma rede de altas frequências (CHENG, 1989). No que se refere à estrutura física, uma linha de transmissão é composta basicamente por um par de condutores separados por um dielétrico (inclusive o ar) que podem estar dispostos em diversas configurações geométricas. A Figura 5 apresenta alguns exemplos comuns de linhas de transmissão utilizadas em aplicações de transporte de dados ou de energia, sendo eles: um par de fios paralelos, um par de placas paralelas (linha de microfita) ou um par de condutores formando um cabo coaxial.

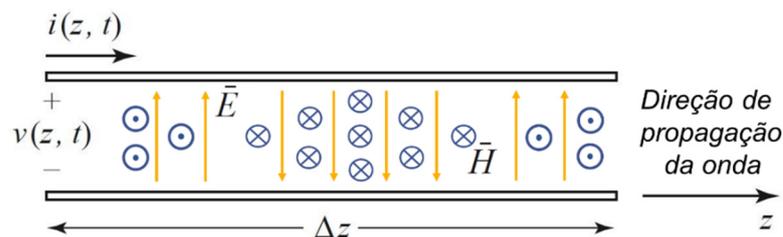
Figura 5 – Exemplos de linhas de transmissão de altas frequências. Par de fios paralelos (a). Linha de microfita (b). Cabo coaxial (c).



Fonte: O autor, 2020.

Para uma onda cujas componentes dos campos elétrico \vec{E} e magnético \vec{H} se encontram num plano ortogonal à direção de propagação (modo transversal eletromagnético - TEM), pode-se representar um trecho Δz de uma linha de transmissão por dois condutores paralelos, como mostrado de forma simplificada na Figura 6. Ao longo dessa linha, as magnitudes e fases das tensões e correntes variam em função tanto do tempo t quanto da posição z .

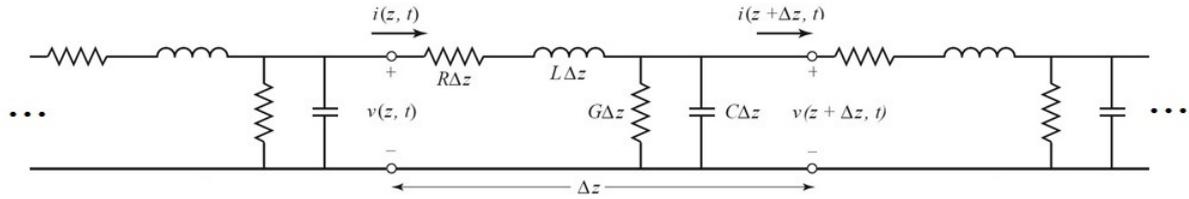
Figura 6 – Representação de um trecho de uma linha de transmissão por um par de fios paralelos.



Fonte: Modificada de Pozar, 2011 e Cheng, 1989.

Considerando que as características elétricas estão distribuídas uniformemente ao longo de uma linha de transmissão infinita, um trecho longitudinal de comprimento infinitesimal Δz pode ser modelado pelo circuito equivalente de parâmetros concentrados mostrado na Figura 7, no qual as grandezas elétricas são definidas por unidade de comprimento. A capacitância paralela C (F/m) e a indutância em série L (H/m) são os elementos que representam o armazenamento e troca de energia elétrica e magnética entre os condutores, enquanto a resistência em série R (Ω/m) e a condutância paralela G (S/m), modelam as perdas no condutor e no material dielétrico, respectivamente (POZAR, 2011).

Figura 7 – Modelagem de um trecho infinitesimal (Δz) de uma linha de transmissão.



Fonte: Adaptado de Pozar, 2011.

Aplicando as leis de Kirchoff de tensão e corrente no circuito da Figura 7 tem-se respectivamente as equações (1) e (2).

$$v(z, t) - R\Delta z i(z, t) - L\Delta z \frac{\partial i(z, t)}{\partial t} - v(z + \Delta z, t) = 0 \quad (1)$$

$$i(z, t) - G\Delta z v(z + \Delta z, t) - C\Delta z \frac{\partial v(z + \Delta z, t)}{\partial t} - i(z + \Delta z, t) = 0 \quad (2)$$

Após manipulações algébricas nas equações (1) e (2) e admitindo o regime harmônico senoidal para a tensão e a corrente, são obtidas, respectivamente, as formas mostrada nas equações (3) e (4).

$$\frac{d^2 V(z)}{dz^2} - \gamma^2 V(z) = 0 \quad (3)$$

$$\frac{d^2 I(z)}{dz^2} - \gamma^2 I(z) = 0 \quad (4)$$

Assim, as soluções para as Equações (3) e (4) são expressas, respectivamente, por (5) e (6), onde γ é definida como a constante de propagação complexa da linha de transmissão, dada pela expressão (7).

$$V(z) = V_0^+ e^{-\gamma z} + V_0^- e^{\gamma z} \quad (5)$$

$$I(z) = I_0^+ e^{-\gamma z} + I_0^- e^{\gamma z} \quad (6)$$

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \quad (7)$$

A impedância característica (Z_0) é definida como a razão entre a tensão e a corrente em um ponto qualquer da linha de transmissão, sendo expressa em termos das suas grandezas elétricas pela expressão (8).

$$Z_0 = \frac{V_0^+}{I_0^+} = \frac{-V_0^-}{I_0^-} = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (8)$$

Convertendo a equação (5) de volta ao domínio do tempo e substituindo (7) e (8), tem-se a expressão da tensão e corrente na linha em função do tempo e posição sendo dadas, respectivamente pelas expressões (9) e (10).

$$v(z, t) = |V_0^+| \cos(\omega t - \beta z + \phi^+) e^{-\alpha z} + |V_0^-| \cos(\omega t + \beta z + \phi^-) e^{\alpha z} \quad (9)$$

$$i(z, t) = \frac{1}{Z_0} (|V_0^+| \cos(\omega t - \beta z + \phi^+) e^{-\alpha z} - |V_0^-| \cos(\omega t + \beta z + \phi^-) e^{\alpha z}) \quad (10)$$

Tomando como exemplo a Equação (9), observa-se na linha de transmissão a combinação de uma onda incidente com magnitude $|V_0^+|$ e ângulo de fase ϕ^+ que se propaga no sentido $+z$ e uma onda refletida com magnitude $|V_0^-|$ e ângulo de fase ϕ^- que se propaga no sentido $-z$. O número de onda β (rad/m) representa a quantidade de ciclos por unidade de comprimento e a constante de amortecimento α (dB/m ou Np/m) define a taxa de desvanecimento da onda devido às perdas no meio de propagação. Para a corrente poderia ser feita a análise dual semelhante, a partir da Equação (10).

Considerando que a onda possui frequência angular $\omega = 2\pi f$ (rad/s), o comprimento de onda λ (m) e a velocidade de fase v_p (m/s) na linha de transmissão podem ser dados respectivamente por (11) e (12).

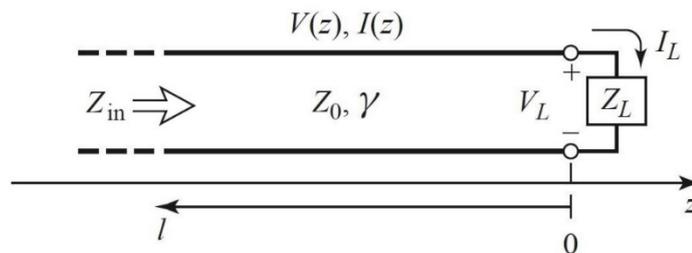
$$\lambda = \frac{2\pi}{\beta} \quad (11)$$

$$v_p = \frac{\omega}{\beta} = \lambda f \quad (12)$$

2.2 IMPEDÂNCIA DE ENTRADA

Como mostrado na equação (8), a relação entre a corrente e a tensão em um dado ponto de uma linha de transmissão infinita é igual a sua impedância característica Z_0 . No entanto, quando essa linha é terminada por uma carga $Z_L \neq Z_0$ (Figura 8), a relação entre tensão e a corrente na carga é forçada a ser Z_L , portanto, deve haver uma onda refletida para que se possa satisfazer essa condição em $z = 0$ (POZAR, 2011).

Figura 8 – Linha de transmissão terminada em uma carga Z_L .



Fonte: Modificada de Pozar, 2011.

A relação entre as amplitudes da onda de tensão refletida e incidente nesse ponto é definida como o coeficiente de reflexão, Γ (POZAR, 2011):

$$\Gamma = \frac{V_0^-}{V_0^+} = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (13)$$

Usando (13), as Equações (5) e (6) podem ser, respectivamente, expressas em função do coeficiente de reflexão por (14) e (15).

$$V(z) = V_0^+ (e^{-\gamma z} + \Gamma e^{\gamma z}) \quad (14)$$

$$I(z) = \frac{V_0^+}{Z_0} (e^{-\gamma z} - \Gamma e^{\gamma z}) \quad (15)$$

Dessa forma, o valor da impedância de entrada Z_{in} vista na linha de transmissão exemplificada na Figura 8 a uma distância $l = -z$ da carga pode ser calculada a partir das Equações (13), (14) e (15) pela expressão (16).

$$Z_{in}(l) = \frac{V(-l)}{I(-l)} = Z_0 \frac{Z_L + Z_0 \tanh \gamma l}{Z_0 + Z_L \tanh \gamma l} \quad (16)$$

2.3 COEFICIENTE DE ONDA ESTACIONÁRIA DE TENSÃO

Quando a impedância da carga não está casada com a impedância característica ($Z_L \neq Z_0$), uma onda refletida deve existir para que a Lei de Ohm seja atendida no ponto $z = 0$ da linha de transmissão. A interação entre a onda incidente e as reflexões produzidas pelo descasamento gera uma onda estacionária que faz com que o valor da tensão ao longo da linha não seja constante (POZAR, 2011). Os valores máximo e mínimo dessa onda estacionária podem ser expressos pelas equações (17) e (18), respectivamente.

$$V_{max}(z) = |V_0^+|(1 + |\Gamma|) \quad (17)$$

$$V_{min}(z) = |V_0^+|(1 - |\Gamma|) \quad (18)$$

À medida em que $|\Gamma|$ aumenta, a razão entre V_{max} e V_{min} também aumenta, logo, uma medida do grau de descasamento entre as impedâncias da linha, chamada de Coeficiente de Onda Estacionária de Tensão ou VSWR pode ser definida pela expressão (19) (POZAR, 2011).

$$VSWR = \frac{V_{max}}{V_{min}} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (19)$$

O $VSWR$ é uma grandeza adimensional representada por um número real positivo tal que $1 \leq VSWR < \infty$, onde um valor de $VSWR = 1$ indica que a impedância da carga está perfeitamente casada com a impedância característica da linha de transmissão ($Z_L = Z_0$).

2.4 PARÂMETROS S

Os parâmetros de espalhamento ou parâmetros S (do inglês *Scattering*) quantificam como um sinal de tensão de alta frequência se propaga através de uma rede multiportas, relacionando as componentes refletida e incidente da onda eletromagnética em cada uma das portas da rede. Uma porta de uma rede de micro-ondas pode ser entendida por qualquer acesso ao circuito onde é possível ser injetado ou medido um sinal eletromagnético, normalmente por meio de um analisador de redes vetorial.

Para uma rede de N portas os parâmetros-S são normalmente apresentados na forma de uma matriz $N \times N$ contendo N^2 parâmetros, como mostrado em (20), onde cada elemento da matriz é dado pela expressão (21).

$$\begin{bmatrix} V_1^- \\ V_2^- \\ \vdots \\ V_N^- \end{bmatrix} = \begin{bmatrix} S_{11} & \cdots & S_{1N} \\ \vdots & \ddots & \vdots \\ S_{N1} & \cdots & S_{NN} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ \vdots \\ V_N^+ \end{bmatrix} \quad (20)$$

$$S_{ij} = \left. \frac{V_i^-}{V_j^+} \right|_{V_k^+ = 0 \text{ para } k \neq j} \quad (21)$$

O parâmetro S_{ij} da matriz S representa a relação entre a tensão da onda refletida na porta i quando as ondas incidentes em todas as demais portas são anuladas, com exceção da porta j . Isso significa que todas as portas $k \neq j$ são terminadas em uma carga casada para evitar outras reflexões. Os parâmetros S são números complexos e podem ser expressos alternativamente em função de sua magnitude $|S_{ij}|$ e fase ϕ_{ij} pela expressão (22).

$$S_{ij} = |S_{ij}| e^{j\phi_{ij}} \quad (22)$$

Um dos parâmetros S mais relevantes é a perda de retorno na porta de entrada (RL_{in} do inglês *Input Return Loss*). Esse parâmetro é geralmente dado em decibéis (dB), em função do S_{11} , pela equação (23).

$$RL_{in} = -20 \log_{10} |S_{11}| \quad (23)$$

3 INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

Sistemas baseados em inteligência artificial deixaram de ser oníricas projeções da ficção científica e passaram a fazer parte do cotidiano das pessoas, sendo difícil, nos dias atuais, encontrar uma atividade humana em que não haja, direta ou indiretamente, alguma de suas diversas aplicações. A presença da inteligência artificial pode hoje ser facilmente identificada em diagnósticos médicos (CHEN, HAO, *et al.*, 2017), (WIENS e SHENOY, 2018), assistentes virtuais (POLYAKOV, MAZHANOV, *et al.*, 2018), (HOY, 2018), aplicações de tradução simultânea (JIMÉNEZ, BENALCÁZAR e SOTOMAYOR, 2017), (AYALA e CHEN, 2017), (MOORKENS, 2018), reconhecimento de imagens (SUPONENKOV, SISOJEVS, *et al.*, 2017), (PAK e KIM, 2017), propaganda (ZHAI, CHANG, *et al.*, 2016), (REN, ZHANG, *et al.*, 2017), (PRABHAKAR, SURESH KUMAR, *et al.*, 2019), sistemas de recomendação (BERTENS, GUITART, *et al.*, 2018), (KHANAL, PRASAD, *et al.*, 2019), previsões meteorológicas (HOLMSTROM, LIU e VO, 2016), (SCHER e MESSORI, 2018), detecção de fraudes financeiras (AWOYEMI, ADETUNMBI e OLUWADARE, 2017), (DHANKHAD, MOHAMMED e FAR, 2018) e controladores de veículos autônomos (FUJIYOSHI, HIRAKAWA e YAMASHITA, 2019), (SHI, WONG, *et al.*, 2020).

Apesar de seus conceitos filosóficos remontarem a tempos antigos, com a formulação por parte de Aristóteles (384-322 A.C.) do primeiro conjunto preciso de leis que governam a parte racional da mente humana, a área de inteligência artificial é um dos mais novos campos da ciência e engenharia. Pode-se considerar que os primeiros trabalhos nessa área foram iniciados logo após a Segunda Guerra Mundial, com as contribuições de Alan Turing (1912-1954), embora sua denominação tenha apenas sido cunhada no ano de 1956 (RUSSEL e NORVIG, 2010).

Embora os conceitos relacionados a essa vasta e recente área do conhecimento convirjam para a possibilidade de uma máquina realizar alguma tarefa de forma autônoma e com o mínimo de intervenção humana, não existe um consenso em relação ao que seja exatamente a inteligência artificial. O que há, no entanto, são diversas definições que abordam a capacidade da máquina de agir ou de pensar de modo semelhante ao de um ser humano ou de forma racionalmente ótima, cujos exemplos podem ser organizados na forma de um quadro, mostrado na Tabela 1.

A área de inteligência artificial é um guarda-chuva multidisciplinar e amplo que agrega contribuições dos campos da filosofia, matemática, economia, neurociência, psicologia,

linguística, computação e engenharia, e que compreende um conjunto de conceitos acessórios que não estão restritos a, mas incluem os seguintes (RUSSEL e NORVIG, 2010):

- a) **processamento de linguagem natural:** capacidade da máquina se comunicar naturalmente em um dado idioma, como o inglês ou o português;
- b) **representação do conhecimento:** capacidade da máquina armazenar aquilo que sabe ou aprende;
- c) **raciocínio automático:** capacidade da máquina usar o conhecimento adquirido para responder a questões, solucionar problemas ou tirar novas conclusões;
- d) **visão por máquina:** capacidade da máquina perceber o ambiente que a cerca;
- e) **robótica:** capacidade da máquina de se mover, manipular objetos e interagir com o ambiente;
- f) **aprendizado de máquina:** capacidade da máquina se adaptar a novas circunstâncias, detectar e extrapolar padrões.

Tabela 1 – Quatro categorias de definições de inteligência artificial.

	Humanamente	Racionalmente
Pensar	“O novo e excitante esforço de fazer computadores pensarem ... máquinas com mentes, no sentido amplo e literal” (HAUGELAND, 1985).	“O estudo das faculdades mentais por meio do uso de modelos” (CHARNIAK e MCDERMOTT, 1985).
	“A automação de atividades que são associadas ao pensamento humano, atividades como tomada de decisões, resolução de problemas, aprendizado, ...” (BELLMAN, 1978).	“O estudo das computações que fazem ser possível perceber, raciocinar e agir” (WINSTON, 1992).
Agir	“A arte de criar máquinas que realizam funções que requerem inteligência quando realizadas por seres humanos” (KURZWEIL, 1990).	“Inteligência computacional é o estudo e o desenvolvimento de agentes inteligentes” (POOLE, 1993).
	“O estudo de como fazer computadores fazerem coisas para as quais, no momento atual, pessoas são melhores” (RICH e KNIGHT, 1991).	“Inteligência artificial ... diz respeito ao comportamento inteligente em artefatos” (NILSSON, 1998).

Fonte: Adaptado de Russel e Norvig, 2010.

O aprendizado de máquina é um ramo da inteligência artificial baseado na ideia de disponibilizar um conjunto de dados a uma máquina e permitir que ela aprenda de forma autônoma a executar uma determinada tarefa, tornando-se cada vez mais habilidosa na medida em que sua experiência aumenta. Em termos formais, pode-se dizer que um programa de computador aprende a partir da experiência E a realizar alguma classe de tarefas T que pode

ser avaliada por alguma medida de desempenho P , se o seu desempenho medido por P nas tarefas contidas em T melhora com a experiência E (MITCHELL, 1997). Assim, o aprendizado de máquina pode ser entendido como um conjunto de métodos computacionais que utilizam experiência para melhorar o desempenho em determinada tarefa ou fazer previsões mais precisas.

Esta tese propõe um sistema que fará uso de aprendizado de máquina para a detecção de padrões em um conjunto de parâmetros eletromagnéticos que possam identificar a presença de falhas estruturais em hastes de âncora. Tomando esse caso específico como exemplo ilustrativo do conceito, pode-se dizer que a tarefa T seria detectar a falha em uma nova haste de âncora, a experiência E seria observar um conjunto de dados previamente rotulados que contivesse exemplos tanto de sinais de hastes defeituosas quanto de hastes normais e, finalmente, a medida de desempenho P seria a taxa de acerto obtida no diagnóstico de novas hastes cujos sinais não tenham sido utilizados durante o treinamento do modelo.

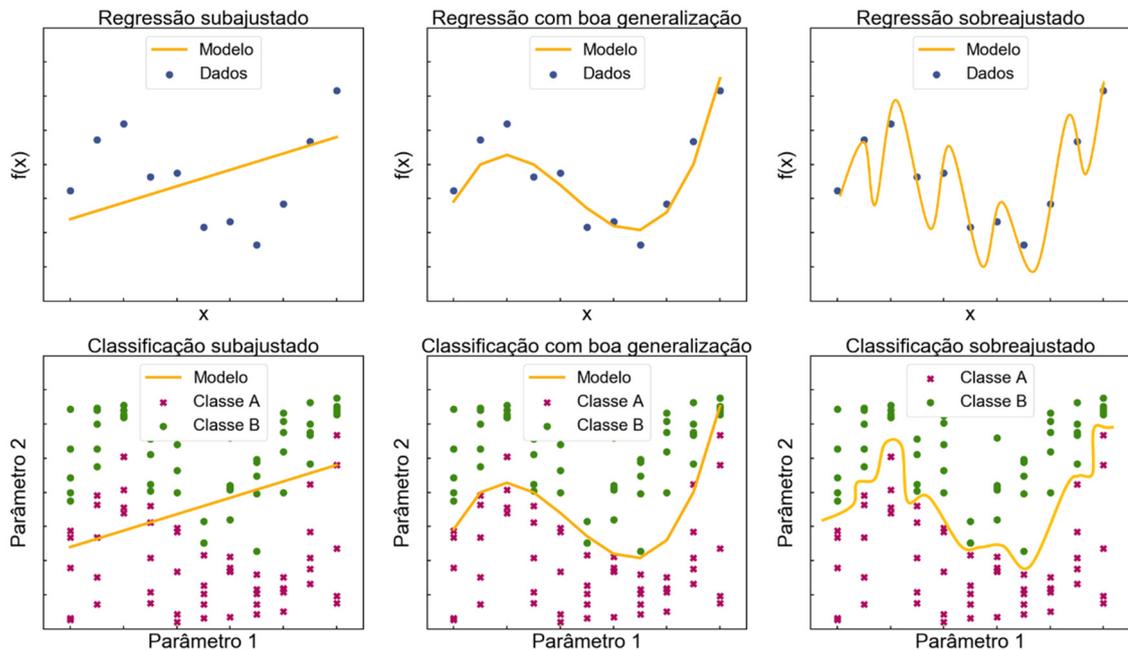
No contexto de aprendizado de máquina, experiência se refere às informações passadas, apresentadas como exemplos ao algoritmo e que tipicamente estão na forma de dados eletronicamente coletados, estruturados, pré-processados e disponibilizados para análise. Esses dados podem ter sido obtidos por meio da interação automática do sistema com o ambiente através de sensores ou estar na forma de conjuntos de treinamento estruturados e que, em algumas vezes, foram classificados previamente por um ser humano. No entanto, é importante ressaltar que, em todos os casos, a qualidade do conjunto de treinamento e a escolha apropriada do algoritmo de aprendizado de máquina para esses dados são fatores cruciais para o bom desempenho do modelo desenvolvido (SÁNCHEZ, BARANDELA, *et al.*, 2003).

Um dos conceitos mais importantes em aprendizado de máquina é o de generalização. Esse conceito, oriundo da psicologia, pode ser definido como o grau em que o algoritmo de aprendizado de máquina consegue mapear corretamente pares entrada-saída que não foram utilizados na geração do modelo, considerando que os dados de teste foram retirados da mesma população que os de treinamento (HAYKIN, 2009). Um modelo que apresenta boa generalização se encontra entre dois limites que devem ser evitados. Em um dos extremos ocorre o fenômeno do subajustamento (*underfitting*), que ocorre quando os exemplos utilizados para treinar o modelo são pouco representativos, ou quando o modelo é muito simples, não sendo capaz de se ajustar adequadamente ao conjunto de treinamento. O subajustamento leva a um baixo desempenho do modelo tanto para o conjunto de treinamento quanto para o de teste. No outro extremo está o sobreajustamento (*overfitting*). Nesse caso, o modelo se ajusta

excessivamente aos dados de treinamento, atingindo altos níveis de desempenho durante o treino, porém, apresenta um baixo desempenho quando confrontado com os novos dados do conjunto de teste (MONARD e BARANAUSKAS, 2003).

A Figura 9 exibe exemplos de modelos com boa generalização e de modelos que apresentam problemas de sobreajustamento ou subajustamento, para os casos de classificação e regressão. Enquanto modelos com subajustamento são pobres e não conseguem mapear a variância intrínseca existente nos dados, modelos sobreajustados acabam mapeando uma variância que é devida ao ruído existente nos dados, e não à sua informação útil.

Figura 9 – Exemplos de modelos com boa generalização e com ocorrência de subajuste e sobreajuste para os casos de classificação e regressão.



Fonte: O autor, 2020.

Uma vez que o sucesso do algoritmo depende fortemente do conjunto de dados utilizado, a área de aprendizado de máquina é inerentemente relacionada com as áreas de análise de dados e estatística. De forma geral, pode-se dizer que as técnicas de aprendizado de máquina são métodos orientados a dados que combinam conceitos fundamentais em computação com ideias oriundas das áreas de estatística, probabilidade e otimização (MOHRI, ROSTAMIZADEH e TALWALKAR, 2012).

Existem basicamente dois paradigmas de aprendizado de máquina. O primeiro paradigma é o aprendizado com professor, também chamado de aprendizado supervisionado.

Nesse tipo de aprendizado é disponibilizado um conjunto de dados rotulados nos quais para cada observação há uma resposta ou saída associada. O objetivo é encontrar um modelo que relacione as entradas às saídas para que seja possível prever os resultados de futuras observações (previsão) ou melhor entender as relações entre as respostas e as entradas (inferência) (JAMES, WITTEN, *et al.*, 2013). Nesse sentido, o objetivo do aprendizado supervisionado pode ser definido como aprender uma certa tarefa representada pela função $f: X \rightarrow Y$, a partir de um conjunto de dados de treinamento $L = (X, y)$, de tal modo que as predições feitas pelo modelo $f(x) = \hat{y}$, sejam tão acuradas quanto possível.

No aprendizado supervisionado, se Y é uma variável categórica, a tarefa aprendida visa resolver um problema de classificação. Por outro lado, se Y é uma variável numérica, então a tarefa aprendida tem por objetivo solucionar um problema de regressão. Diante disso, um classificador pode ser definido como uma função $f: X \rightarrow Y$, onde Y é um conjunto de classes ou rótulos finitos, denotados por $\{C_1, C_2, \dots, C_j\}$, e, de forma análoga, um regressor pode ser definido como uma função $f: X \rightarrow y$, onde $y \in \mathbb{R}$.

O segundo paradigma é o aprendizado sem professor, que, por sua vez, pode ser subdividido em aprendizado não-supervisionado e aprendizado por reforço. No aprendizado não-supervisionado não são disponibilizadas ao algoritmo as respostas associadas às observações realizadas, e o que se busca é compreender as relações entre as variáveis ou entre as observações em si (JAMES, WITTEN, *et al.*, 2013). Nesse caso, os dados de treinamento consistem em vetores X de entrada sem nenhum vetor Y de valores-alvo correspondentes. O objetivo desses algoritmos pode ser, por exemplo, descobrir grupos ou conjuntos de elementos similares nos dados, o que é chamado de agrupamento ou *clustering*; determinar a distribuição dos dados no espaço de entradas, o que é chamado de estimação de densidade, ou projetar os dados de um hiperespaço de dimensão n em outro de dimensão k (em que $k < n$), reduzindo a dimensionalidade dos dados para diminuir seu tempo de processamento ou para que possam ser representados em duas ou três dimensões, permitindo a sua visualização (BISHOP, 2006).

O outro tipo de aprendizado sem professor é o aprendizado por reforço. Nesse tipo de aprendizado, o algoritmo deve buscar tomar ações em determinadas situações com o objetivo de maximizar algum tipo de recompensa. Como não são disponibilizadas para o algoritmo as respostas ótimas para o problema, ele deve explorar o espaço de entradas interagindo com o ambiente através de uma sequência de estados, ações, recompensas e penalidades por um processo de tentativa e erro (BISHOP, 2006).

A adoção de um ou outro paradigma depende do tipo e da estrutura dos dados disponíveis, bem como do problema específico que se deseja abordar. Na composição das soluções propostas nesta tese, será utilizado aprendizado do tipo supervisionado para o desenvolvimento de classificadores e aprendizado do tipo não-supervisionado para redução da dimensionalidade dos dados.

Dentre as diversas soluções existentes, serão abordados nas seções 3.1 a 3.5 seguintes os algoritmos K-NN; *random forests*; SVM; regressão logística e redes neurais artificiais, os quais foram selecionados pelo seu destaque em termos de popularidade e desempenho em tarefas de classificação por meio de aprendizado supervisionado (WU, KUMAR, *et al.*, 2008). Ao longo deste trabalho, os desempenhos dos classificadores binários desenvolvidos com cada um desses algoritmos serão avaliados e comparados, na tarefa de identificar o mais indicado para compor o sistema proposto para o diagnóstico de falhas estruturais em hastes de âncora a partir dos dados de parâmetros eletromagnéticos. Adicionalmente, será introduzido na seção 3.6 o conceito de análise de componentes principais, uma técnica de aprendizado não-supervisionado amplamente adotada para redução da dimensionalidade dos dados em aplicações de aprendizado de máquina (JAMES, WITTEN, *et al.*, 2013). Essa técnica será utilizada para diminuição da quantidade de processamento requerida nas simulações e testes necessários para definição de hiperparâmetros dos classificadores.

Ressalta-se que cada um dos algoritmos que serão apresentados é, por si só, uma fonte extensa de pesquisas e aplicações, não sendo pretensão desta tese esgotar sua análise, mas sim, apresentá-los em linhas gerais, de modo que seja possível compreender os seus princípios de funcionamento e seu papel no sistema de diagnóstico proposto.

3.1 K-NEAREST NEIGHBORS

Como seu nome sugere, o algoritmo de K-NN ou K-vizinhos mais próximos realiza a classificação de um novo elemento encontrando os $K > 0$ pontos que possuem as menores distâncias a esse elemento dentro do conjunto de teste e, então, associando esse novo elemento à classe que possuir mais representantes dentre os K localizados. Essa medida de distância não se restringe à distância euclidiana, podendo ser utilizada qualquer outra medida de distância considerada adequada para comparação dos pontos do conjunto de treinamento específico.

Métodos baseados em avaliação de pontos vizinhos ao ponto de teste são conhecidos como métodos de aprendizado de máquina não generalizados, porque eles simplesmente memorizam todos os dados de conjunto de treinamento para poder realizar suas análises.

Entretanto, embora haja estratégias para otimizar a busca de elementos nesses dados, a necessidade de armazenar todo o conjunto de treinamento pode requerer uma extensiva carga computacional, especialmente se o conjunto de treinamento for extenso, (BISHOP, 2006).

Seja um conjunto de dados de treinamento com N elementos no total, os quais devem ser classificados em j classes C_1, C_2, \dots, C_j . Nesse conjunto há N_l elementos pertencentes à classe C_l , tal que $\sum_l N_l = N$. A classificação de uma nova amostra x de dimensão n pelo algoritmo K-NN começa com a construção de uma hipersfera n -dimensional centrada em x , cujo raio é iniciado em um pequeno valor e é incrementado até que ela contenha precisamente K elementos do conjunto de treinamento, como mostrado na Figura 10.

Considerando que dentro do volume V dessa hipersfera estão contidos exatamente K_l elementos que pertençam à classe C_l , e que $\sum_l K_l = K$, então uma estimativa da densidade associada a cada classe em torno do ponto x é dada por (24).

$$p(x|C_l) = \frac{K_l}{N_l V} \quad (24)$$

Similarmente, a densidade incondicional é dada por (25) enquanto a probabilidade *a priori* por (26).

$$p(x) = \frac{K}{NV} \quad (25)$$

$$p(C_l) = \frac{N_l}{N} \quad (26)$$

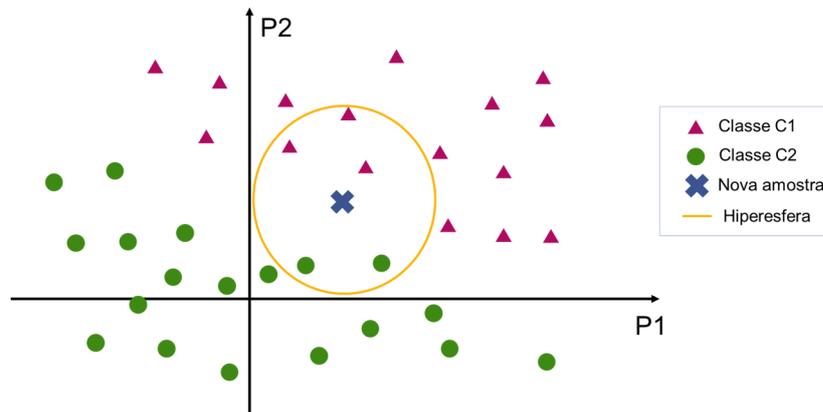
Assim, usando o Teorema de Bayes, pode-se dizer que a probabilidade *a posteriori* do elemento x ser um membro da classe C_l é dada por (27).

$$p(C_l|x) = \frac{p(x|C_l)p(C_l)}{p(x)} = \frac{K_l}{K} \quad (27)$$

Utilizando o resultado apresentado em (27), se observa que é possível minimizar a probabilidade de erro de classificação associando o ponto de teste x à classe que apresente a maior probabilidade *a posteriori*, ou seja, o maior valor de K_l/K . Dessa forma, para classificar

uma nova amostra x , basta identificar, por meio de alguma medida de distância, o conjunto dos K elementos mais próximos desse ponto, e então associá-lo à classe que contenha o maior número de representantes dentro desse conjunto.

Figura 10 – Exemplo de construção da hiperesfera do algoritmo K-NN para $K=5$.



Fonte: O autor, 2020.

Apesar de sua simplicidade, o algoritmo K-NN tem sido bem sucedido em um grande número de problemas de classificação e regressão, incluindo reconhecimento de dígitos manuscritos e análise de cenas em imagens de satélites. Sendo um método não-paramétrico, ele costuma apresentar bom desempenho em situações de classificação nas quais a fronteira do limite de decisão é bastante irregular (PEDREGOSA, VAROQUAUX, *et al.*, 2011).

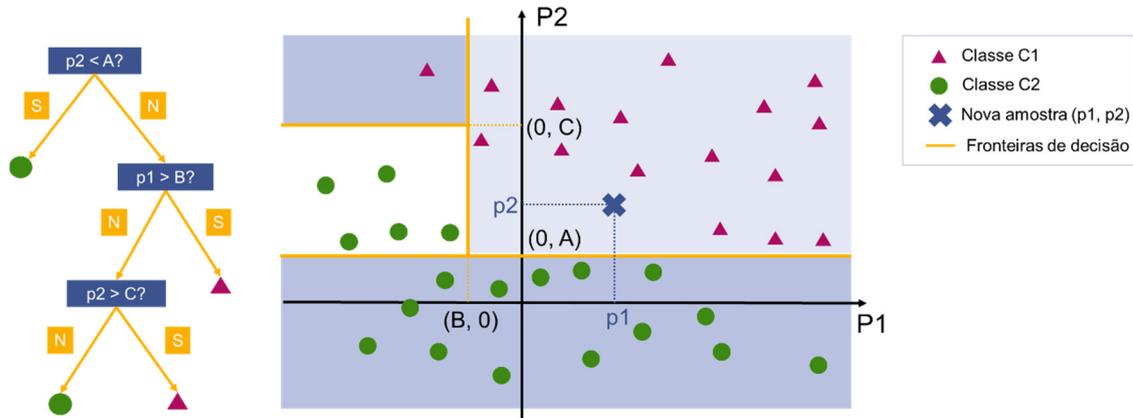
3.2 RANDOM FORESTS

O *random forests* é um algoritmo de aprendizagem supervisionada que faz uso do método de predição por árvores de decisão para a construção de modelos mais robustos. Uma árvore de decisão é um método que realiza a predição por meio da segmentação recursiva do conjunto de dados em um número de regiões disjuntas mais simples, de acordo com os valores de seus parâmetros, formando uma estrutura ramificada em árvore (JAMES, WITTEN, *et al.*, 2013). Quando utilizado como classificador, a separação entre classes ocorre a cada nó da árvore, como mostrado na Figura 11.

Durante a etapa de treinamento, o algoritmo decide a cada iteração qual dos parâmetros é o que melhor divide o conjunto de dados, com o objetivo de tentar manter agrupados em cada partição apenas elementos de uma mesma classe. Essa estruturação da cadeia de decisão como uma árvore permite uma classificação rápida, direta e facilmente legível para um ser humano,

que é capaz de interpretar quais os critérios que foram adotados pelo algoritmo para realizar a predição, o que nem sempre é possível em outros métodos de aprendizado de máquina.

Figura 11 – Esquema de classificação de um novo ponto por árvore de decisão.



Fonte: O autor, 2020.

Considerando um conjunto $X_{m \times n}$ de m vetores de treinamento constituídos de n parâmetros $[x_1, x_2, \dots, x_n]$ e um conjunto de rótulos $Y_{m \times 1} = (y_1, y_2, \dots, y_m)^T$, em que cada elemento y_i pertence a uma certa classe C_i e Q o subconjunto de dados a ser dividido num dado nó k , iniciando com todo o conjunto de dados $Q = (X, Y)$. Seja também um critério de separação $\theta = (j, t_k)$, formado pelo par composto pelo parâmetro j e pelo seu respectivo valor de limiar t_k que divide Q em duas partições $Q_{esq}(\theta)$ e $Q_{dir}(\theta)$, tais que:

$$Q_{esq}(\theta) = \{x, y \in Q \mid x_j \leq t_k\} \quad (28)$$

$$Q_{dir}(\theta) = \{x, y \in Q \mid x_j > t_k\} \quad (29)$$

O algoritmo busca que, em cada nó k da árvore, a separação seja realizada de tal modo que maximize a pureza (ou minimize a impureza) das partições de Q . Idealmente, isso significa que, no limite, em cada partição Q_p haja apenas amostras pertencentes a uma única classe.

Medidas típicas de impureza H utilizadas pelo algoritmo na construção das árvores de decisão são o índice de Gini ou a entropia. Essas grandezas são expressas, respectivamente, pelas equações (30) e (31), considerando que p_{lk} é a proporção de observações da classe l feitas na partição Q_p de um dado nó k .

$$H(Q_p) = \sum_l p_{lk}(1 - p_{lk}) \quad (30)$$

$$H(Q_p) = - \sum_l p_{lk} \log(p_{lk}) \quad (31)$$

Seja N_k o número de elementos em Q , n_{esq} o número de elementos na partição $Q_{esq}(\theta)$ e n_{dir} o número de elementos na partição $Q_{dir}(\theta)$, com $N_k = n_{esq} + n_{dir}$. A impureza total a ser minimizada no nó k pode ser calculada através da expressão (32).

$$G(Q, \theta) = \frac{n_{esq}}{N_k} H(Q_{esq}(\theta)) + \frac{n_{dir}}{N_k} H(Q_{dir}(\theta)) \quad (32)$$

O algoritmo seleciona o critério de separação θ^* que minimiza $G(Q, \theta)$ em cada nó k , recursivamente, até que um critério de parada como, por exemplo, a profundidade máxima permitida para a árvore ou um número mínimo de folhas seja atingido (BREIMAN, FRIEDMAN, *et al.*, 1984).

Métodos de predição por árvores de decisão são simples e possuem uma fácil interpretação dos critérios adotados na predição (ALI, KHAN, *et al.*, 2012). No entanto, as árvores de decisão são modelos sensivelmente instáveis, variando demasiadamente em função de pequenas alterações nos dados e, tipicamente, não conseguem atingir resultados competitivos com outros métodos de aprendizado de máquina (JAMES, WITTEN, *et al.*, 2013).

Diante dessa questão, o algoritmo *random forests* consegue, por meio de uma estratégia engenhosa, contornar os problemas observados nas árvores de decisão e construir modelos denominados florestas aleatórias que possuem alto poder de predição e desempenho (SHOTTON, FITZGIBBON, *et al.*, 2011). Uma floresta aleatória é uma combinação de um número definido B de preditores para funções de valores discretos, em que cada função individual aprendida está estruturada como uma árvore de decisão (BREIMAN, 2001).

Ao invés de construir uma única árvore de decisão utilizando todos os m dados do conjunto de treinamento, o algoritmo *random forests* constrói B árvores de decisão independentes, cada uma sendo treinada com um subconjunto de tamanho hm , com $0 \leq h \leq 1$ (PANOV e DZEROSKI, 2007). Cada árvore é formada em função apenas dos dados presentes

no seu subconjunto específico de dados de treinamento, cujos elementos foram aleatoriamente selecionados com a mesma distribuição de probabilidade e de forma independente das demais árvores da floresta. Essa técnica é conhecida como *bagging* (*bootstrap aggregation*) e se apoia no fato de que, considerando um conjunto de B observações independentes Z_1, Z_2, \dots, Z_B , em que cada observação apresenta variância σ^2 , a variância da média das observações \bar{Z} é dada por σ^2/B . Assim, ao utilizar como predição a média de um conjunto de observações, se obtém uma redução da variância (HASTIE, TIBSHIRANI e FRIEDMAN, 2009).

Para descorrelacionar ainda mais os preditores, cada árvore da floresta é individualmente construída utilizando um subespaço aleatório de q parâmetros de entrada, em que tipicamente é adotado o critério $q \approx \sqrt{n}$. Dessa forma, cada árvore generaliza sua classificação de uma maneira distinta, o que ajuda a reduzir ainda mais a possibilidade de sobreajuste (HO, 1995). Finalmente, após a construção dos preditores, a classificação de um novo dado pelo algoritmo *random forests* é simplesmente realizada por um esquema de votação por maioria simples, tomada entre todas as B predições individuais que foram realizadas por cada árvore de decisão existente na floresta.

Em comparação aos demais modelos de aprendizado de máquina, pode-se dizer que a estratégia utilizada pelo algoritmo *random forests* para produzir previsões mais acuradas e estáveis ocorre simplesmente por meio da criação aleatória de uma grande quantidade de árvores de decisão. Dessa forma, ao invés de buscar pelo melhor parâmetro global de classificação, o algoritmo realiza essa busca em subconjuntos aleatórios extraídos dos dados, criando uma grande diversidade que geralmente leva a modelos melhores (BRAMER, 2016). No entanto, a principal desvantagem dessa estratégia é que o algoritmo pode se tornar demasiadamente lento quando é utilizada uma grande quantidade de árvores para evitar o sobreajuste.

Adicionalmente, uma das maiores vantagens do *random forests*, herdado das árvores de decisão, é a possibilidade de interpretação dos critérios utilizados pelo algoritmo na classificação, o que nem sempre é possível quando utilizados outros algoritmos de aprendizado de máquina. No caso de modelos implementados com o *random forests*, essa característica permite uma compreensão do grau de importância de cada componente do vetor de entrada no processo de decisão, indo no sentido da raiz para as folhas da árvore.

3.3 SUPPORT VECTOR MACHINES

Uma máquina de vetores de suporte ou SVM é um algoritmo de aprendizado de máquina binário que, com base em um conjunto de dados com amostras positivas e negativas, constrói como superfície de decisão o hiperplano de máxima margem de separação entre as classes (VAPNIK, 1995). Para classificar um conjunto de dados com dimensão N , um hiperplano de $(N-1)$ dimensões é gerado pelo SVM (Oommen, Misra, *et al.*, 2008).

Considerando o caso mais simples em que um conjunto $X_{m \times n}$ de m vetores de treinamento com n dimensões pertencentes a duas classes C_1 e C_2 , associado a um conjunto de rótulos binários $Y_{m \times 1} = (y_1, y_2, \dots, y_m)^T$, $y_i \in \{-1, 1\}$, em que os elementos $x_i \in X$ de classes diferentes possam ser perfeitamente separados pelo hiperplano (33), em que w é o vetor perpendicular a esse hiperplano e o valor $b/\|w\|$ a distância do hiperplano até a origem (Scholkopf e Smola, 2001).

$$f(x) = w \cdot x + b = 0 \quad (33)$$

O hiperplano (33) separa o espaço X em duas regiões, dadas por $f(x) > 0$ e $f(x) < 0$. A separação nas classes $C_1 = +1$ e $C_2 = -1$ pode ser feita com o auxílio da função $g(x) = \text{sgn}(x)$, sendo dada por (34):

$$g(x) = \text{sgn}(w \cdot x + b) = \begin{cases} +1 & \text{se } w \cdot x + b > 0 \\ -1 & \text{se } w \cdot x + b < 0 \end{cases} \quad (34)$$

Há infinitos hiperplanos que atendem à equação (33) e que podem ser obtidos simplesmente pela multiplicação de w e b por uma constante arbitrária. Dentre essa família de hiperplanos, será considerada a escolha particular dos valores de w e b que consistem na chamada forma canônica de (33), e que é definida por (35) (Scholkopf e Smola, 2001).

$$\min_{i=1, \dots, m} |w \cdot x_i + b| = 1 \quad (35)$$

A forma canônica (35) obriga que os pontos x_i do conjunto de dados X que estão mais próximos da superfície de separação (33) estejam a uma distância $1/\|w\|$ do hiperplano e $2/\|w\|$ do ponto mais próximo que pertença a outra classe (Scholkopf e Smola, 2001).

Assim, a classificação dos pontos nas classes pode ser dada pelas inequações (36), que podem ser reescritas em uma forma resumida por (37) (LORENA e DE CARVALHO, 2007).

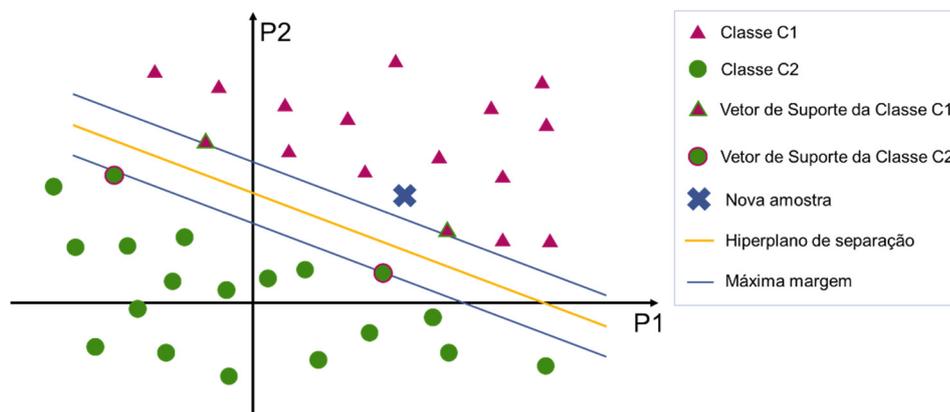
$$\begin{cases} w \cdot x_i + b \geq +1, \text{ se } y_i = +1 \\ w \cdot x_i + b \leq -1, \text{ se } y_i = -1 \end{cases} \quad (36)$$

$$y_i(w \cdot x_i + b) - 1 \geq 0 \quad (37)$$

Os pontos particulares para os quais (36) é satisfeita com a igualdade são denominados vetores de suporte. Esses vetores desempenham um papel fundamental para a separação entre classes, fazendo com que todos os demais pontos do conjunto de treinamento se tornem completamente irrelevantes. O SVM define como vetores de suporte os pontos mais facilmente confundíveis da base de dados e os utiliza como referência para classificação dos demais. Por confundíveis, entendem-se os pontos do conjunto X , pertencentes a uma determinada classe, que mais se assemelhem a elementos da classe oposta.

Em termos conceituais, os vetores de suporte são os pontos do conjunto de treinamento mais próximos do hiperplano de separação e que estão localizados exatamente sobre os hiperplanos que definem a margem, sendo, portanto, os mais difíceis de classificar. A Figura 12 apresenta visualmente os principais conceitos utilizados pelo algoritmo SVM para um espaço de parâmetros $P2 \times P1$ fictício (HAYKIN, 2009).

Figura 12 – Conceitos de hiperplano de separação, margem e vetores de suporte do SVM.



Fonte: O autor, 2020.

A margem ótima ou máxima margem é aquela que separa os elementos de classes distintas pela maior distância possível, ou seja, devem ser escolhidos parâmetros w e b que atendam a condição (38), sujeita à restrição (37). Em um conjunto de dados linearmente separáveis, pode haver infinitos hiperplanos de separação, no entanto, apenas um deles possui máxima margem, sendo denominado hiperplano ótimo. Os vetores de suporte que restringem a largura dessa margem estão localizados em hiperplanos paralelos ao hiperplano ótimo (OOMMEN, MISRA, *et al.*, 2008).

$$\operatorname{argmax}_{w,b} \frac{2}{\|w\|} \quad (38)$$

Por conveniência, a condição (38) pode ser reescrita na forma equivalente (39). Essa estratégia permite sua abordagem por técnicas de resolução de problemas de programação quadrática, em que se busca minimizar uma função quadrática sujeita a um conjunto de inequações lineares na variável w (BISHOP, 2006). Esses problemas possuem ampla e estabelecida teoria matemática (SCHOLKOPF e SMOLA, 2001).

$$\operatorname{argmin}_{w,b} \frac{\|w\|^2}{2} \quad (39)$$

Como a função objetivo a ser minimizada é convexa e os pontos que satisfazem as restrições também formam um conjunto convexo, o problema possui um único mínimo global, podendo ser solucionado com a introdução de uma função Lagrangiana (40), que engloba em única expressão as restrições à função objetivo associadas a coeficientes α_i denominados multiplicadores de Lagrange (LORENA e DE CARVALHO, 2007).

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i (y_i(w \cdot x_i + b) - 1) \quad (40)$$

A solução do problema de otimização está num ponto de sela da função Lagrangiana (BERTSEKAS, 1995), então, tomando as derivadas parciais de (40) em relação às variáveis w e b e igualando-as a zero, tem-se, com algumas manipulações algébricas, as condições de optimalidade (41) e (42) (HAYKIN, 2009).

$$\frac{\partial L(w, b, \alpha)}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^m \alpha_i y_i x_i \quad (41)$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = 0 \Rightarrow \sum_{i=1}^m \alpha_i y_i = 0 \quad (42)$$

Substituindo (41) e (42) em (40) tem-se a representação dual da Lagrangiana dada por (43), expressa apenas em função de produtos internos de α e dos dados conhecidos x e y , sujeita as restrições (44) e (45), cujos coeficientes α_i podem ser obtidos por programação quadrática (BISHOP, 2006).

$$\tilde{L}(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i x_j \quad (43)$$

$$\alpha_i \geq 0 \quad (44)$$

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad (45)$$

Substituindo (41) em (33), tem-se que o hiperplano ótimo é dado em função dos multiplicadores de Lagrange por (46).

$$f(x) = \sum_{i=1}^m \alpha_i y_i x_i x + b \quad (46)$$

Para o problema dual, obtidos os coeficientes α_i e utilizando as condições de Karush-Kuhn-Tucker (SCHOLKOPF e SMOLA, 2001), ocorre que as condições (47) devem ser obedecidas.

$$\begin{cases} \alpha_i \geq 0 \\ y_i f(x_i) - 1 \geq 0 \\ \alpha_i [y_i f(x_i) - 1] = 0 \end{cases} \quad (47)$$

Assim, para cada ponto de dados, ou tem-se $\alpha_i = 0$ ou $y_i f(x_i) = 1$. Qualquer dado que possua $\alpha_i = 0$ não tem influência na expressão (46) que define o hiperplano de separação, e, conseqüentemente, na classificação de novos dados (BISHOP, 2006). Os pontos restantes são os chamados vetores de suporte e por satisfazerem a condição $y_i f(x_i) = 1$, de acordo com a equação (37), se pode observar que eles estão localizados nos hiperplanos que definem a máxima margem de separação.

Como os vetores de suporte devem atender à condição $y_i f(x_i) = 1$, o termo b pode ser obtido utilizando essa expressão conjuntamente com (46) para qualquer vetor x_j que pertença ao conjunto S de vetores de suporte (BISHOP, 2006).

$$y_j \left(\sum_{i \in S} \alpha_i y_i x_i x_j + b \right) = 1 \quad (48)$$

No entanto, a solução numericamente mais estável (49) é geralmente obtida, multiplicando-se os dois membros da equação (48) por y_j , utilizando $y_j^2 = 1$, e obtendo a média para todos os N_s vetores de suporte existentes em S para só então, isolar o termo b (BISHOP, 2006).

$$b = \frac{1}{N_s} \sum_{j \in S} \left(y_j - \sum_{i \in S} \alpha_i y_i x_i x_j + b \right) \quad (49)$$

Tendo obtido o vetor w por (41) e o termo b por (46), basta substituir esses valores em (34), para que a classificação de qualquer novo ponto de dados seja, finalmente, realizada. Assim, após o treinamento do modelo, apenas os vetores de suporte são requeridos na classificação e os demais pontos de dados são simplesmente descartados (BISHOP, 2006).

Como apenas em raros casos os dados reais são linearmente separáveis, é tipicamente introduzida no algoritmo uma variável de folga (*slack variable*) $C > 0$, que controla o compromisso entre a linearidade do hiperplano de separação e um certo grau de tolerância a

presença de elementos dentro da margem de separação. Além disso, mapeamentos e transformações nas variáveis de entrada por meio de *kernels* também são tipicamente realizadas para permitir a operação do SVM em regime não-linear (SMOLA e SCHÖLKOPF, 2004).

3.4 REGRESSÃO LOGÍSTICA

Apesar de seu nome sugerir uma regressão, a regressão logística é na verdade um modelo paramétrico de classificação, no qual a saída é dada por uma função logística sigmoideal agindo em uma função linear do vetor de parâmetros de entrada. O modelo gerado pela regressão logística provê uma função $f: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ e um conjunto de coeficientes α_i que expressam a probabilidade *a posteriori* de um novo elemento x ser um membro da classe C_l , na forma da expressão (50) (DREISEITL e OHNO-MACHADO, 2002).

$$p(C_l|x) = f(x, \alpha_l) \quad (50)$$

Utilizando o Teorema de Bayes e considerando o caso mais simples, para a classificação em duas classes C_1 e C_2 , tem-se que (BISHOP, 2006):

$$p(C_1|x) = \frac{p(x|C_1) p(C_1)}{p(x)} \quad (51)$$

$$p(C_1|x) = \frac{p(x|C_1) p(C_1)}{p(x|C_1) p(C_1) + p(x|C_2) p(C_2)} \quad (52)$$

$$p(C_1|x) = \frac{1}{1 + \frac{p(x|C_2) p(C_2)}{p(x|C_1) p(C_1)}} \quad (53)$$

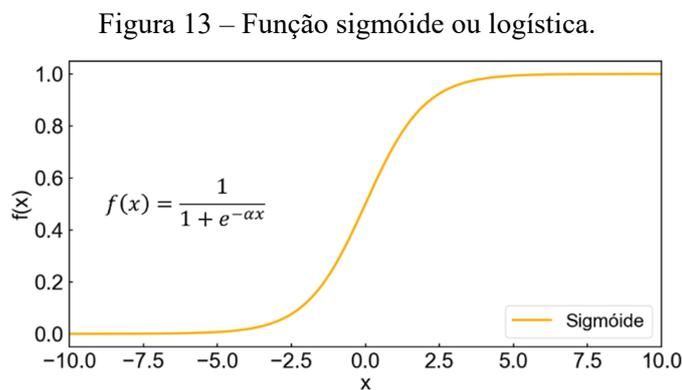
Definindo:

$$a = \ln \left(\frac{p(x|C_1) p(C_1)}{p(x|C_2) p(C_2)} \right) \quad (54)$$

Tem-se que:

$$p(C_1|x) = \frac{1}{1 + \exp(-a)} = \sigma(a) \quad (55)$$

A função $\sigma(a)$ é denominada função logística ou sigmóide e recebe esse nome por conta de seu gráfico que possui um formato semelhante ao da letra “S”, como mostrado na Figura 13. Por sua vez, a expressão a , que representa o logaritmo da razão entre as probabilidades $p(C_1|x)$ e $p(C_2|x)$, é a função inversa da função logística, sendo denominada de função logit.



Fonte: O autor, 2020.

Considerando que as classes C_1 e C_2 particionam todo o espaço, $p(C_1) = 1 - p(C_2)$. Assim, usando (51), a expressão (54) pode ser reescrita como:

$$a = \ln\left(\frac{p(C_1|x)}{p(C_2|x)}\right) = \ln\left(\frac{p(C_1|x)}{1 - p(C_1|x)}\right) = \ln\left(\frac{\sigma(a)}{1 - \sigma(a)}\right) \quad (56)$$

Na regressão logística, o vetor de parâmetros de entrada é dado por $x = [x_0, x_1, \dots, x_n]$, para o qual, por simplificação da notação é definido $x_0 = 1$. O vetor x é então convertido em um valor real por meio do seu produto escalar com um vetor de coeficientes $[\alpha_0, \alpha_1, \dots, \alpha_n]^T$, segundo uma transformação linear dada por (70).

$$a = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_n x_n \quad (57)$$

Usando (50), (55) e (57), e associando arbitrariamente à classe C_1 o valor 1 e à classe C_2 o valor 0, pode-se definir que o modelo de classificação por regressão logística, é governado pelas equações (58) e (59).

$$p(C_1|x) = \frac{1}{1 + \exp -(\alpha_0 + \alpha_1 x_1 + \dots + \alpha_n x_n)} \quad (58)$$

$$p(C_2|x) = 1 - p(C_1|x) \quad (59)$$

Os coeficientes α_i que parametrizam o modelo são determinados a partir das amostras da base de dados de treinamento, sendo estimados por máxima verossimilhança (AGGARWAL, 2018). O controle de sobreajuste do algoritmo de regressão logística é tipicamente realizado por meio de um fator de penalização $C > 0$ que controla o grau de complexidade do modelo através de estratégias de regularização L1 (*Lasso*) ou L2 (*Ridge*) (PEREIRA, BASTO e DA SILVA, 2016). A fronteira de separação entre as classes do modelo de regressão logística é o lugar geométrico dado pela expressão (60). Nesse ponto os valores das expressões (58) e (59) se igualam, ou seja, $p(C_1|x) = p(C_2|x) = 0,5$.

$$\alpha_0 + \alpha_1 x_1 + \dots + \alpha_n x_n = 0 \quad (60)$$

Finalmente, utilizando os resultados (58) e (59) no processo de classificação, o elemento x é associado à classe C_l que apresentar $p(C_l|x) > 0,5$.

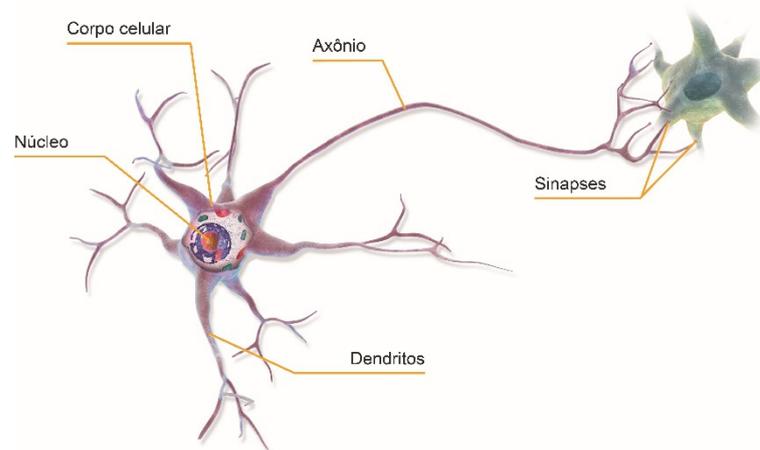
3.5 REDES NEURAIIS ARTIFICIAIS

As redes neurais artificiais são modelos de predição inicialmente inspirados no modo de funcionamento dos neurônios e do cérebro dos seres vivos. As redes neurais artificiais merecem especial destaque pela sua grande capacidade de generalização e de mapeamento de relações subjacentes entre variáveis de entrada e saída sem a necessidade de modelos físicos formais. Tal capacidade as tornam excelentes ferramentas para a realização de tarefas que envolvam classificação e regressão em análise de dados.

Um esquema simplificado da estrutura de um neurônio biológico é mostrado na Figura 14. O neurônio biológico recebe impulsos elétricos ou sinais químicos emitidos por outros neurônios por meio de prolongamentos finos e ramificados do seu corpo celular, denominados

dendritos. Ao atingir um determinado limiar de excitação, devido à quantidade e configuração dos estímulos recebidos, o neurônio responde com modificações no potencial elétrico de sua membrana celular. Esse impulso elétrico gerado como resposta do neurônio à excitação recebida é conduzido através de uma ramificação denominada axônio até atingir outro neurônio, por intermédio de uma sinapse. Na sinapse, o estímulo eletroquímico gerado pelo neurônio transmissor é passado por intermédio de neurotransmissores ao neurônio receptor, que dá continuidade ao processo até que seja gerada uma ação final de natureza motora, racional ou sensorial (HAYKIN, 1999).

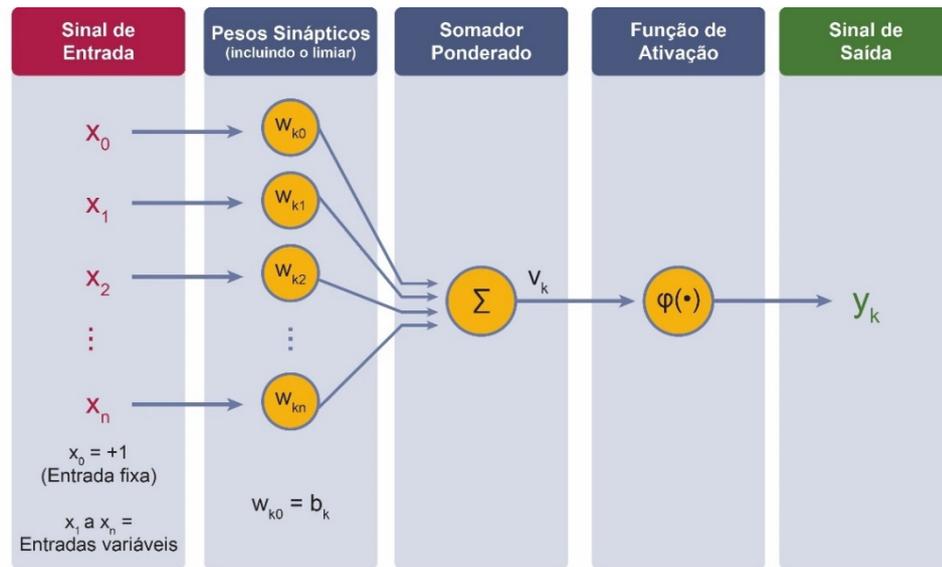
Figura 14 – Esquema simplificado da estrutura de um neurônio biológico.



Fonte: Adaptado de Blaus, 2013.

De um modo análogo ao dos neurônios biológicos em relação ao cérebro, os neurônios artificiais são as unidades básicas de processamento das redes neurais artificiais. A Figura 15 mostra um neurônio artificial e seus elementos básicos. O neurônio artificial é composto por um conjunto de entradas interligadas por meio de conexões ponderadas, ou sinapses, a um somador cuja saída alimenta uma função de ativação. A quantidade resultante da soma dos produtos entre os valores de entrada e os pesos sinápticos é chamada de campo local induzido. O valor do campo local induzido é então aplicado como entrada em uma função de ativação que, por fim, calcula a saída do neurônio.

Figura 15 – Modelo esquemático de um neurônio artificial.



Fonte: Adaptado de Haykin, 2009.

Consideremos que o neurônio artificial k da Figura 15 possui um valor de limiar b_k , um conjunto de pesos sinápticos $W_k' = [w_{k1}, w_{k2}, \dots, w_{kn}]$ e é excitado por um vetor de entrada com n componentes $X' = [x_1, x_2, \dots, x_n]$. Sem perda da generalidade, para fins de simplificação da notação, serão definidos os novos vetores $X = [x_0, x_1, \dots, x_n]$ e $W = [w_{k0}, w_{k1}, \dots, w_{kn}]$, para os quais $x_0 = +1$ e $w_{k0} = b_k$. Dessa forma, o campo local induzido desse neurônio pode ser dado pela expressão (61), que transforma o vetor X em um escalar v_k .

$$v_k = \sum_{i=0}^n w_{ki} x_i \quad (61)$$

A função de ativação φ define a saída y_k do neurônio k em função do seu campo local induzido v_k . Assim, considerando que o neurônio k possui uma função de ativação $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, a sua saída é dada pela expressão (62):

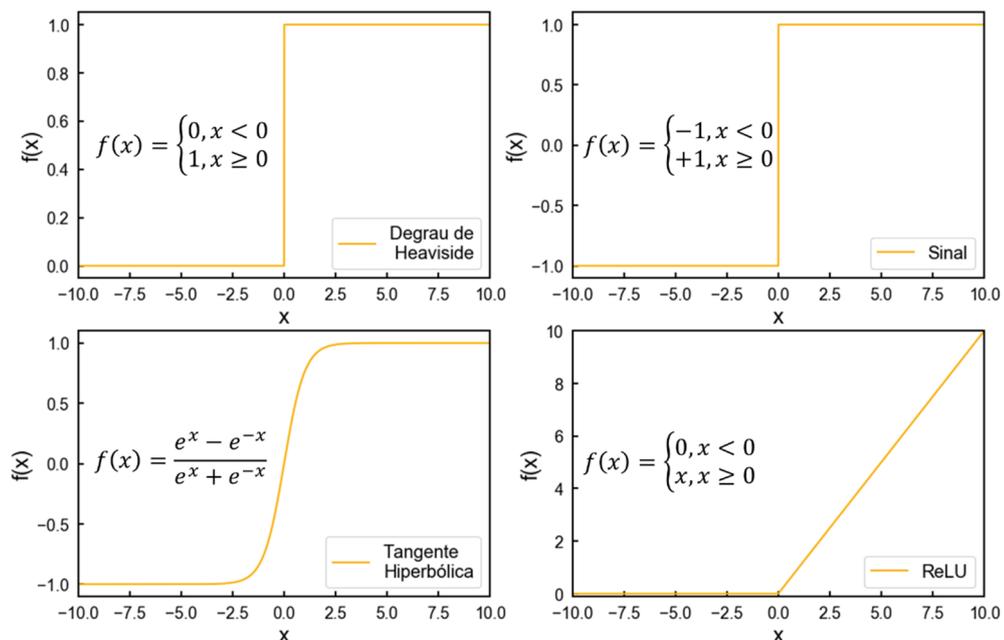
$$y_k = \varphi(v_k) \quad (62)$$

As funções de ativação, em geral, têm como objetivos limitar a saída do neurônio ou introduzir não-linearidades ao sistema, o que permite que as redes neurais modelem sistemas complexos e não-lineares. Além da função sigmoide, já mostrada na Figura 13, outros exemplos

de funções de ativação comumente utilizadas na implementação de redes neurais artificiais são exibidos na Figura 16 (KARLIK e OLGAC, 2011), (GLOROT, BORDES e BENGIO, 2011). Algumas dessas funções de ativação como a sigmóide, degrau de Heaviside, sinal e tangente hiperbólica restringem a saída do neurônio a um conjunto ou faixa de valores limitados, sendo tipicamente utilizadas na camada de saída das RNAs por possuírem algum significado associado em termos de classificação ou probabilidades. Por outro lado, nas camadas ocultas, a função conhecida como ReLU (*Rectified Linear Unit*) e suas variações têm obtido excelentes resultados e sido amplamente adotadas (HE, ZHANG, *et al.*, 2015).

Nos seres vivos, os neurônios não atuam isoladamente, mas interconectados em estruturas complexas chamadas redes neurais. Como exemplo, no cérebro humano, se estima que haja algo na ordem de 100 bilhões de neurônios (WASSERMAN, 1989). Uma RNA funciona de modo análogo, sendo um processador paralelo composto por um conjunto de unidades simples de processamento – os neurônios artificiais – que possui a capacidade natural de acumular conhecimento experimental e torná-lo disponível para uso. (HAYKIN, 2009). Pode-se dizer que uma RNA se assemelha ao cérebro pois também é uma máquina adaptativa em que o conhecimento é adquirido do meio ambiente por meio de um processo de aprendizado, e onde os pesos das conexões entre neurônios, chamados de pesos sinápticos, são utilizados para armazenar o conhecimento adquirido.

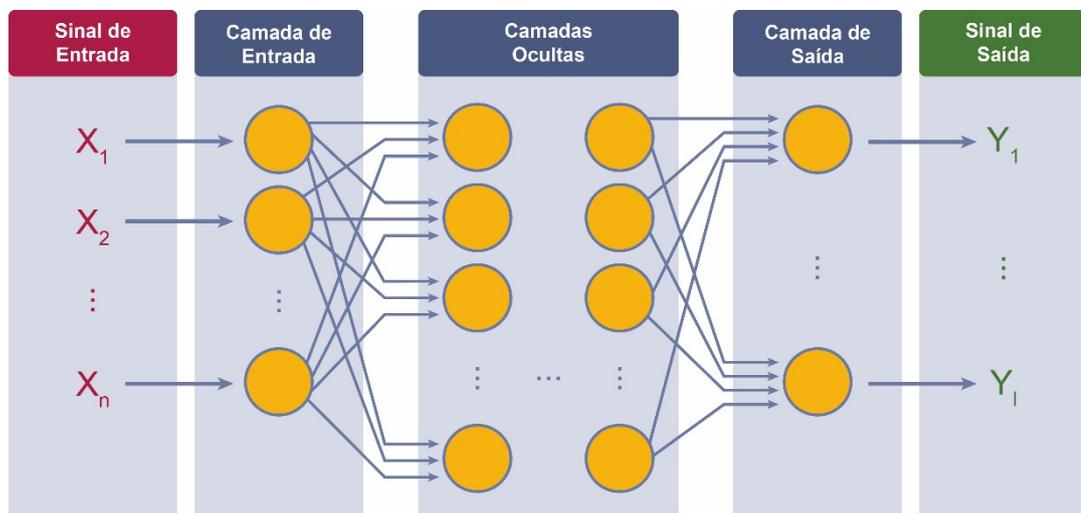
Figura 16 – Exemplos de funções de ativação.



Fonte: O autor, 2020.

Dentre as formas possíveis de se organizar os neurônios e formar uma rede neural artificial, uma das mais comuns e amplamente utilizadas é a arquitetura do tipo FF-MLP (*Feed Forward Multi-Layered Perceptron*). Nessas RNAs os neurônios artificiais são estruturados em camadas sucessivas, como mostrado na Figura 17. As saídas dos neurônios de uma camada alimentam as entradas dos neurônios da camada seguinte, sucessivamente, e o sinal se propaga através da RNA em um único sentido desde a camada de entrada até a saída, não existindo caminhos de retorno ou ciclos fechados (*loops*).

Figura 17 – Estrutura básica de uma Rede Neural Artificial tipo MLP.



Fonte: O autor, 2020.

O número de neurônios que serão utilizados na camada de entrada de uma FF-MLP é diretamente relacionada com a quantidade de parâmetros do vetor de entrada que essa RNA irá processar. A quantidade de neurônios da camada de saída, por sua vez, depende do uso da RNA como classificador ou como regressor. Assim, para modelos de classificação em que os dados estão divididos em l classes, a camada de saída possuirá l neurônios com saídas geralmente binárias. Por outro lado, para os modelos de regressão a RNA possui tipicamente um único neurônio cuja saída é um valor real.

As camadas intermediárias, não visíveis, de uma rede FF-MLP são denominadas de camadas ocultas (*hidden layers*). Essa denominação é dada porque os neurônios das camadas ocultas extraem características altamente abstratas dos dados e seus valores contém informações que, muitas vezes, não são explicitamente observáveis ou facilmente

interpretáveis. Nas camadas ocultas o modelo é quem determina de forma autônoma quais conceitos e representações são úteis para explicar as relações existentes entre os dados de entrada e saída (GOODFELLOW, BENGIO e COURVILLE, 2016).

Em uma RNA do tipo MLP projetada para atuar como um classificador, o conhecimento é adquirido do ambiente através de um processo de aprendizado supervisionado. Isso significa que a RNA aprende através de exemplos pré-classificados de amostras passadas que são continuamente apresentados à rede durante a etapa de treinamento. Esse conhecimento adquirido é então armazenado nos pesos das interconexões entre neurônios e fica disponível para o uso em futuras classificações. Assim, similarmente ao cérebro humano, as RNAs aprendem a partir de experiências, e não por meio da programação de um conjunto de regras.

O objetivo do treinamento da RNA pode ser entendido como a busca de um conjunto de pesos sinápticos W que minimiza o erro E na saída Y para uma base de dados (X, T) específica. Um dos processos mais consagrados para treinamento de uma rede MLP é o de propagação reversa (*Backpropagation*) (LECUN, BOTTOU, *et al.*, 2012). Para ilustrar o algoritmo de *backpropagation*, por meio de equações mais compactas, será utilizada a notação matricial das expressões (61) e (62), que são dadas, respectivamente por (63) e (64).

$$V = WX \quad (63)$$

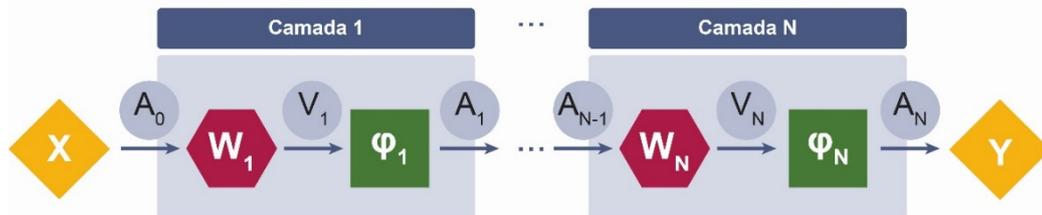
$$y = \varphi(V) \quad (64)$$

O processo básico de treinamento de uma RNA se inicia com um conjunto $X_{m \times n}$ de m vetores de treinamento, constituídos de n parâmetros $[x_1, x_2, \dots, x_n]$, um conjunto de valores-alvo associados $T_{m \times 1} = (t_1, t_2, \dots, t_m)^T$ e o vetor $Y_{m \times 1} = (y_1, y_2, \dots, y_m)^T$ contendo as respectivas saídas da rede FF-MLP genérica de N camadas mostrada na Figura 18. Nessa RNA, W_k representam as matrizes de pesos sinápticos dos neurônios da k -ésima camada da RNA ($k = 1, 2, \dots, N$) e φ_k suas respectivas funções de ativação.

Inicialmente, as matrizes de pesos W_k são inicializadas com um conjunto de valores aleatórios. Em seguida, as entradas X são aplicadas à RNA e a saída Y é obtida pelas equações (63) e (64) aplicadas a cada camada, desde a entrada até a saída, sequencialmente. Finalmente, o erro E dos valores Y previstos pela RNA em relação aos valores alvo T é obtido, onde a função de erro $E = f(Y, T)$ pode ser qualquer medida de distância entre os valores de Y e T ,

sendo tipicamente utilizados o erro absoluto ou o erro médio quadrático entre esses valores. Essa primeira etapa do algoritmo é denominada etapa de passagem direta (*forward pass*), pois a informação flui, sucessivamente, camada após camada, desde a entrada até a saída da RNA.

Figura 18 – Fluxo de informação no sentido direto da rede FF-MLP.



Fonte: O autor, 2020.

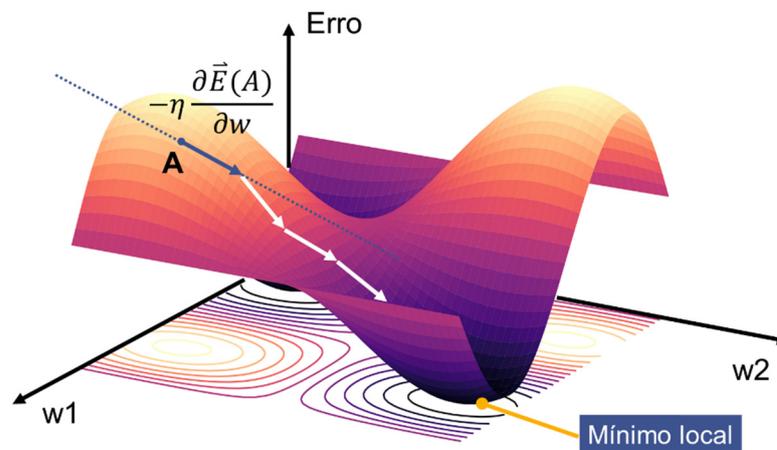
A atualização dos valores dos pesos sinápticos da camada k após a iteração i é realizada em função do erro, por meio da expressão (65). O termo η , denominado taxa de aprendizado, controla o equilíbrio entre velocidade do treinamento e garantia de convergência para um mínimo local. Esse termo é, no caso mais simples, uma constante escalar positiva, no entanto, abordagens mais complexas utilizam um valor de η variável (LECUN, BOTTOU, *et al.*, 2012). Dependendo da curvatura da superfície da função erro no espaço de pesos, o valor da taxa de aprendizado adequado será maior ou menor, buscando atingir um mínimo local a uma velocidade razoável e, ao mesmo tempo, evitar a divergência do algoritmo.

$$W_k \leftarrow W_k - \eta \cdot \frac{\partial E}{\partial W_k} \quad (65)$$

O gradiente de uma função define a direção de seu máximo crescimento, portanto, tomando o negativo do gradiente, é obtida a direção de máxima diminuição do valor da função erro no espaço de pesos W^n . Por utilizar como estratégia de atualização dos pesos uma série de passos iterativos na direção da diminuição do gradiente da função erro, esse método de otimização é denominado gradiente descendente (*gradient descent - GD*). O GD busca iterativamente um mínimo local da função erro no espaço de pesos W^n , atualizando os pesos das camadas a cada iteração com um incremento proporcional ao produto da sua taxa de influência no erro (gradiente) pela taxa de aprendizado que controla a convergência do algoritmo.

A Figura 19 ilustra a sequência fictícia de passos do GD partindo de um ponto A do espaço de pesos W^2 , até atingir um mínimo local da função erro E . A cada iteração, o tamanho do incremento dado aos pesos W da RNA é definido pelo módulo dos componentes do gradiente da função E no ponto atual, escalonado pela taxa de aprendizado utilizada. Por sua vez, como se busca atingir um mínimo local, a direção do incremento dado aos pesos W pelo algoritmo GD é a oposta da apresentada pelo gradiente da função E no ponto atual.

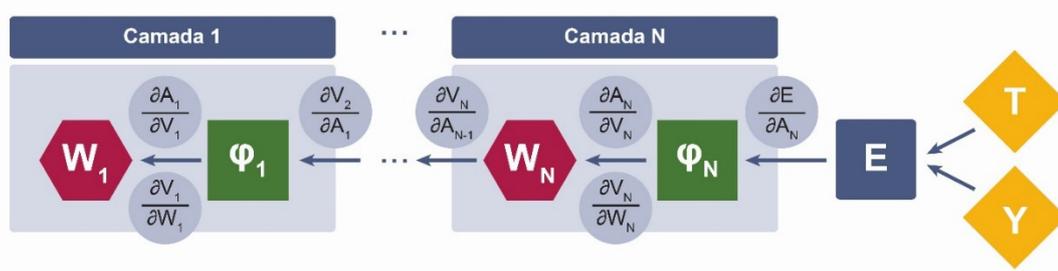
Figura 19 – Exemplo de funcionamento do método do Gradiente Descendente.



Fonte: O autor, 2020.

Para a atualização dos valores dos pesos sinápticos da camada k , é necessário o cálculo do gradiente da função erro em relação ao conjunto de pesos dessa camada. Esse valor é obtido a partir da camada de saída utilizando a regra da cadeia. Essa etapa é chamada de passagem reversa (*backward pass*) pois, a partir do valor do erro calculado na iteração, os pesos sinápticos dos neurônios são ajustados um a um desde a camada de saída até a primeira chamada, na direção oposta à do fluxo de informação na operação normal da RNA. O fluxo da passagem reversa do *backpropagation* é ilustrado na Figura 20, onde são mostrados os cálculos das derivadas parciais que são utilizadas na definição dos novos valores dos pesos das camadas.

Figura 20 – Fluxo de informação no sentido reverso para ajuste dos pesos sinápticos da rede FF-MLP no algoritmo de *backpropagation*.



Fonte: O autor, 2020.

Considerando como exemplo o caso de uma FF-MLP com uma única camada oculta com pesos W_1 e função de ativação φ_1 e cuja camada de saída possui pesos W_2 e função de ativação φ_2 , as equações da passagem direta são definidas por:

$$A_0 = X \quad (66)$$

$$V_1 = W_1 \cdot A_0 \quad (67)$$

$$A_1 = \varphi_1(V_1) \quad (68)$$

$$V_2 = W_2 \cdot A_1 \quad (69)$$

$$A_2 = \varphi_2(V_2) = Y \quad (70)$$

Definindo convenientemente em (71) uma função erro E , que mede a diferença entre o valor Y obtido pela RNA em (70) e o valor desejado T , os ajustes nos pesos são realizados por meio das equações (72) e (73) do método gradiente descendente.

$$E = \frac{1}{2}(Y - T)^2 \quad (71)$$

$$W_2 \leftarrow W_2 - \eta \cdot \frac{\partial E}{\partial W_2} \quad (72)$$

$$W_1 \leftarrow W_1 - \eta \cdot \frac{\partial E}{\partial W_1} \quad (73)$$

Nas equações (72) e (73), os valores dos gradientes são obtidos, respectivamente, pela regra da cadeia, conforme as equações (74) e (75).

$$\frac{\partial E}{\partial W_2} = \frac{\partial E}{\partial A_2} \cdot \frac{\partial A_2}{\partial V_2} \cdot \frac{\partial V_2}{\partial W_2} \quad (74)$$

$$\frac{\partial E}{\partial W_1} = \frac{\partial E}{\partial A_2} \cdot \frac{\partial A_2}{\partial V_2} \cdot \frac{\partial V_2}{\partial A_1} \cdot \frac{\partial A_1}{\partial V_1} \cdot \frac{\partial V_1}{\partial W_1} \quad (75)$$

Após algumas manipulações nas equações (74) e (75) e utilizando as definições de (66) a (71), se obtém (76) e (77).

$$\frac{\partial E}{\partial W_2} = (Y - T) \cdot \varphi_2'(V_2) \cdot A_1 \quad (76)$$

$$\frac{\partial E}{\partial W_1} = (Y - T) \cdot \varphi_2'(V_2) \cdot W_2 \cdot \varphi_1'(V_1) \cdot X \quad (77)$$

Considerando que as funções de ativação são deriváveis, (76) e (77) podem ser reescritas apenas em função de termos conhecidos por meio das expressões (78) e (79).

$$\frac{\partial E}{\partial W_2} = (Y - T) \cdot \varphi_2'(W_2 \cdot \varphi_1(W_1 \cdot X)) \cdot \varphi_1(W_1 \cdot X) \quad (78)$$

$$\frac{\partial E}{\partial W_1} = (Y - T) \cdot \varphi_2'(W_2 \cdot \varphi_1(W_1 \cdot X)) \cdot W_2 \cdot \varphi_1'(W_1 \cdot X) \cdot X \quad (79)$$

Atualizados os valores de W_1 e W_2 o processo se repete e continua até que o número máximo de iterações definido seja alcançado ou algum outro critério de parada seja atingido.

O projeto desse tipo de modelo deve ser feito com atenção especial, uma vez que as RNAs possuem a característica de serem aproximadores universais (HORNIK,

STINCHCOMBE e WHITE, 1989). Isso significa que uma rede MLP com a uma quantidade suficiente de neurônios, camadas e rodadas de treinamento é teoricamente capaz de modelar qualquer função existente, o que acaba por torná-las mais facilmente sujeitas ao problema de sobreajustamento do modelo. O sobreajustamento pode ser identificado via testes robustos, como o de validação cruzada (KOHAVI, 1995) e contornado por meio de estratégias como a diminuição da complexidade da RNA ou o aumento da base de dados disponível, bem como pela utilização de técnicas como parada antecipada (*early stopping*) (LI, SOLTANOLKOTABI e OYMAK, 2020), exclusão seletiva (*dropout*) (SRIVASTAVA, HINTON, *et al.*, 2014) ou regularização (GOODFELLOW, BENGIO e COURVILLE, 2016).

3.6 ANÁLISE DE COMPONENTES PRINCIPAIS

A análise de componentes principais ou PCA é uma técnica que consiste em modificar, em termos vetoriais, a base de representação dos vetores de entrada, permitindo a redução da dimensionalidade desses dados com uma perda de informação controlada. Após a mudança de base realizada pelo PCA, é possível representar os vetores de entrada por um conjunto minimizado de componentes dimensionais que reduzem o esforço computacional necessário para seu processamento, mas que ainda são capazes de expressar a maior parte da variabilidade dos dados originais, mantendo suas características principais (JAMES, WITTEN, *et al.*, 2013). O PCA é considerado uma técnica de aprendizado de máquina não-supervisionado pois envolve apenas o conjunto de parâmetros X de entrada, mas nenhum conjunto de respostas Y associado.

O processo de decomposição da base de dados $X_{m \times n}$, de m vetores de entrada constituídos de n parâmetros $[x_1, x_2, \dots, x_n]$, em suas componentes principais, se inicia com o cálculo da matriz de covariância $\Sigma_{n \times n}$, por meio da expressão (80).

$$\Sigma = \frac{1}{m} X^T X \quad (80)$$

A matriz de covariância Σ é simétrica e pode ser representada em função da sua matriz diagonal de autovalores $\Lambda_{n \times n} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, em que, por conveniência, os autovalores λ_i são dispostos em ordem decrescente $\lambda_1 > \lambda_2 > \dots > \lambda_n$, e da matriz $Q_{n \times n} = [q_1, q_2, \dots, q_n]$ de respectivos autovetores ortonormais associados, de acordo com a expressão (81).

$$\Sigma = Q \Lambda Q^T \quad (81)$$

Os vetores q_i dispostos nas colunas da matriz Q são denominados os componentes principais dos dados contidos em X . Esses componentes são ordenados em sentido decrescente, de acordo com a sua contribuição na variância σ_i^2 dos dados. Por sua vez, na matriz diagonal Λ , cada autovalor λ_i possui valor igual ao da variância σ_i^2 contida nos dados na direção do autovetor q_i . Assim, a variância total nos dados pode ser dada por (82), onde σ_i^2 é a variância do i -ésimo componente principal q_i (HAYKIN, 2009).

$$var_{total} = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i \quad (82)$$

Em muitas aplicações, os primeiros poucos componentes principais contabilizam praticamente toda a variação dos dados, e, para muitos propósitos, é possível trabalhar somente com esse diminuto conjunto de componentes principais e descartar os demais (RUPPERT e MATTERSON, 2015). De acordo com a expressão (82), a fração p da variância retida nos dados ao se usar apenas os primeiros k componentes principais pode ser dada por (83).

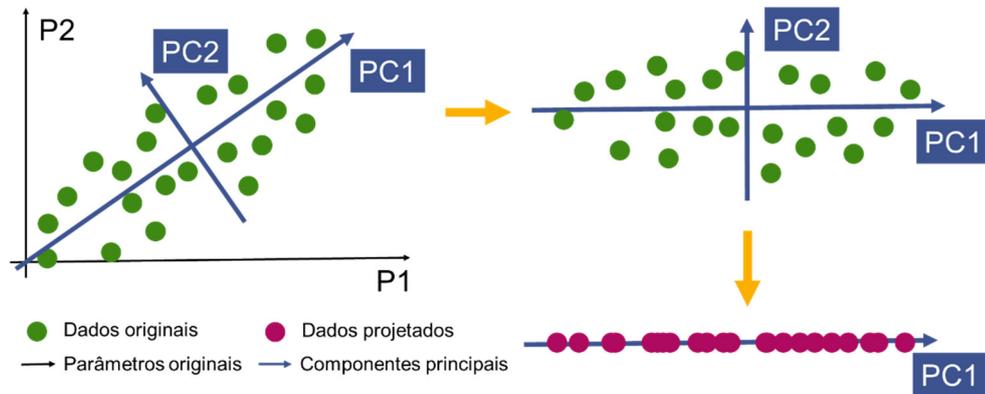
$$p = \frac{1}{var_{total}} \sum_{i=1}^k \lambda_i \quad (83)$$

A partir de (83), se pode determinar o número k de componentes principais necessários para se reter uma determinada fração p de variância nos dados. Valores de p tipicamente utilizados são 0,95 e 0,99, significando, respectivamente, que, após a redução da dimensionalidade, será retida 95% ou 99% da variância total contida nos dados. De posse de k , é possível converter os dados n -dimensionais contidos originalmente em $X_{m \times n}$ em dados k -dimensionais, com $k < n$, por meio da transformação (84), onde $X'_{m \times k}$ é a matriz dos dados contidos em X representados nas k componentes principais e $Q'_{n \times k}$ é a matriz formada pelas k primeiras colunas de $Q_{n \times n}$.

$$X' = XQ' \quad (84)$$

Finalmente, na forma apresentada em (84), se obtém a redução de dimensionalidade dos dados de n para k dimensões, como mostrado na Figura 21.

Figura 21 – Esquema do processo de redução de dimensionalidade com PCA



Fonte: O autor, 2020.

O resíduo de variância que foi descartado juntamente com os componentes de ordem mais baixa dá origem ao erro de aproximação associado a essa redução, e pode ser expresso por (85).

$$var_{residual} = \sum_{i=k+1}^n \sigma_i^2 = \sum_{i=k+1}^n \lambda_i \quad (85)$$

Por fim, a recuperação do dado no espaço n -dimensional original $\tilde{X}_{m \times n}$ é feita pela transformação reversa dada pela expressão (86). No entanto, é importante ressaltar que devido à perda de informação associada à variância residual descartada em (85), a recuperação do dado após a redução de dimensionalidade não é exata, e sim aproximada.

$$\tilde{X} = X'Q'^T \quad (86)$$

O aumento crescente da disponibilidade de dados estruturados e da capacidade computacional, tornou comum o uso de modelos de aprendizado de máquina nos mais variados campos do conhecimento. Essa popularização produziu implementações otimizadas e com alto desempenho computacional dos algoritmos descritos nas seções de 3.1 a 3.6 deste Capítulo. Tais implementações estão disponíveis na forma de bibliotecas de *software* prontas para o uso

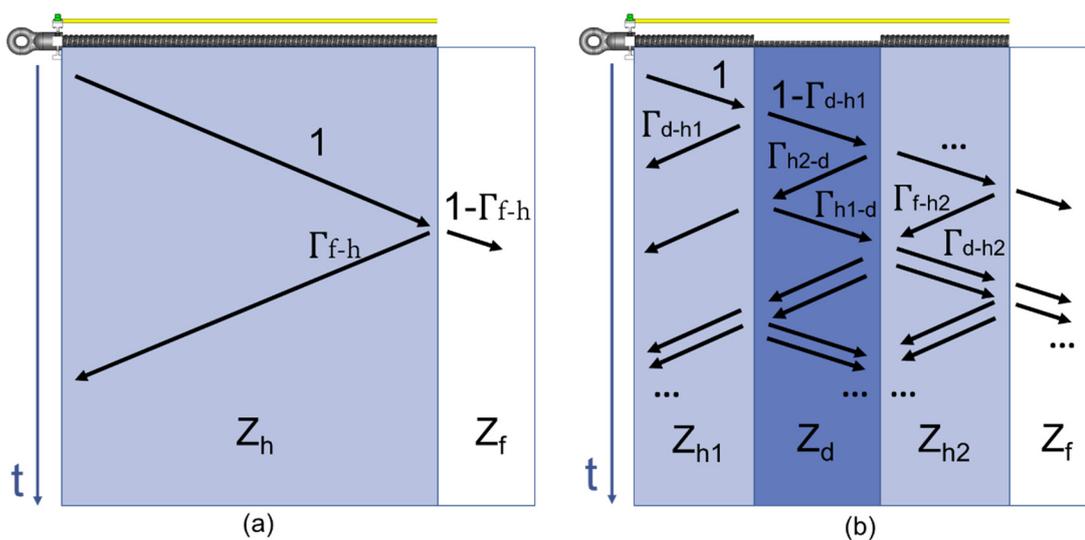
modular em soluções de engenharia, nas mais diversas linguagens e plataformas de desenvolvimento tais como R (FOUNDATION, 2020), Python (PYTHON, 2020), Java (JAVA, 2020), Matlab (MATLAB, 2020), Octave (OCTAVE, 2020) e outras. Dessa forma, não há necessidade, portanto, de se desenvolver implementações particulares desses algoritmos para seu uso em aplicações práticas como a proposta neste trabalho, que não buscam melhorias no algoritmo em si, mas que são focadas, simplesmente, na utilização desses modelos como ferramenta de análise de dados.

4 DESENVOLVIMENTO DO SISTEMA DE DIAGNÓSTICO

O princípio de funcionamento do sistema de diagnóstico proposto, se baseia no fato de que, para efeito de modelagem, uma haste de âncora é, essencialmente, um cilindro condutor metálico. Assim, ao se posicionar um segundo condutor ao longo da haste de âncora, denominado haste de referência, pode-se considerar, em uma primeira análise, que esse par de hastes se comporte como uma linha de transmissão (LOURENÇO NOVO, 2015). Considerando que haja uma conexão de baixas perdas entre o par de hastes e uma fonte de sinais de tensão, é possível aplicar uma onda de alta frequência no sistema e medir um conjunto de parâmetros eletromagnéticos da linha de transmissão formada por essas hastes (BEZERRA, DE MEDEIROS, *et al.*, 2014). Parâmetros eletromagnéticos são afetados por descasamentos causados por discontinuidades no meio em que a onda se propaga (POZAR, 2011) (CHENG, 1989), e, portanto, são capazes de transportar informação relativa à presença de eventuais falhas na estrutura da haste de âncora.

A Figura 22 mostra o princípio básico de funcionamento do sistema de diagnóstico proposto. Uma onda eletromagnética incidente se propagando em um meio físico, ao encontrar uma interface com um outro meio de impedância característica diferente, sofre uma reflexão e uma transmissão. A proporção entre o quanto da onda é refletida em relação à quantidade incidente depende das características dos meios e é dada pelo coeficiente de reflexão Γ , o qual é representado por um número complexo.

Figura 22 – Esquema simplificado do comportamento da onda eletromagnética para os casos da haste normal (a) e defeituosa (b).



Fonte: O autor, 2020.

Considerando que a linha de transmissão formada pela haste de âncora e a haste de referência seja uniforme e tenha impedância característica Z_h , para o caso da haste normal mostrado em (a), a onda incidente aplicada no sistema sofre uma reflexão apenas ao chegar ao final da haste, onde encontrará uma interface de mudança de meio (Z_h-Z_f) com uma nova impedância característica Z_f . Por outro lado, considerando um caso genérico de um único defeito inserido no meio da haste mostrado em (b), devido à alteração da composição do material causada pelo defeito no trecho com impedância Z_d , são geradas as interfaces ($Z_{h1}-Z_d$), (Z_d-Z_{h2}) e ($Z_{h2}-Z_f$) em cada ponto de mudança de meio. Apesar dos trechos Z_{h1} e Z_{h2} mostrados na figura serem constituídos do mesmo material físico (aço, por exemplo), eles são representados por impedâncias distintas pois, devido às alterações geométricas causadas pela falha eles podem vir a se comportar como meios eletromagnéticos distintos.

Cada vez que a onda eletromagnética encontra uma das interfaces mostradas em (b) (independentemente do sentido de propagação) um percentual dela é transmitido e outro refletido. Assim, o que é medido na porta de entrada posicionada no início do meio Z_{h1} para o caso (b) é a combinação de todas as componentes que retornam a esse ponto pelos múltiplos caminhos ilustrados, o que é diferente do medido no caso (a). As múltiplas reflexões e refrações no caso (b), em teoria seriam infinitas, mas na prática desvanecem devido a perdas nos materiais (SANTOS e SILVA, 2014). Dessa forma, o princípio de funcionamento do sistema de diagnóstico se apoia no fato de que, devido aos diferentes níveis de interação entre as ondas incidentes e refletidas para os casos de hastes normais e defeituosas, é possível distinguir essas duas condições a partir dos sinais medidos na porta de entrada do sistema.

A modelagem de sistemas que operam em altas frequências por meio de um conjunto de parâmetros é uma técnica amplamente utilizada, sendo diversos os parâmetros utilizados para esse fim (POZAR, 2011), (CHENG, 1989). Dentre as alternativas existentes, as análises apresentadas nesta tese irão se concentrar nos parâmetros S_{11} , Z_{in} e $VSWR$, por serem medidos diretamente pelo analisador de redes vetorial portátil, modelo *Keysight Fieldfox N9952A*, que será utilizado na aquisição de dados das hastes de âncora testadas.

Utilizando os resultados das equações (16), (19) e (21) para uma rede de uma única porta, os parâmetros eletromagnéticos citados podem ser expressos em função do coeficiente de reflexão Γ pelas expressões (87), (88) e (89), respectivamente (POZAR, 2011).

$$Z_{in} = Z_0 \frac{1 + \Gamma}{1 - \Gamma} \quad (87)$$

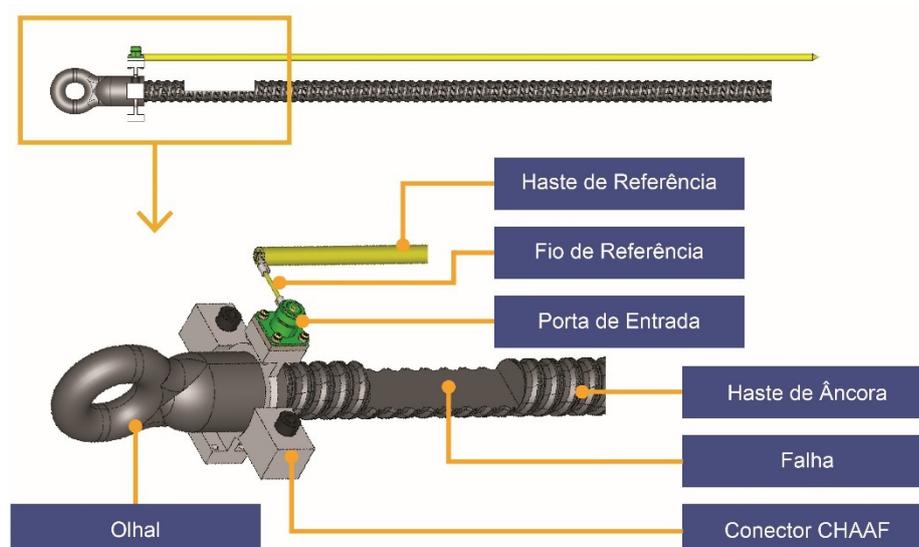
$$VSWR = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (88)$$

$$S_{11} = \Gamma \quad (89)$$

Os elementos do sistema são integrados através de um conector denominado CHAAF (Conector Haste-Analisador para Altas Frequências) cujas características foram especialmente projetadas para garantir o acoplamento eletromagnético entre as hastes a serem medidas e o equipamento de medição (LOURENÇO NOVO, DE MELO, *et al.*, 2014). Uma visão detalhada do esquema de conexão entre as hastes, mostrando a porta de entrada dos sinais de alta frequência, é apresentada na Figura 23.

Em termos construtivos, o conector CHAAF basicamente é composto por duas partes de formato aproximadamente semicircular que fazem contato mecânico e elétrico com a haste de âncora e que são ligados por dois parafusos de fixação em cada uma das pontas. A geometria exata do conector é de tal forma que há um casamento de impedância entre o analisador e o sistema composto pelas hastes. No topo do CHAAF há uma porta de conexão com o analisador de redes com um conector fêmea tipo N para altas frequências, e um fio de referência que interliga o conjunto à haste de referência.

Figura 23 – Detalhe da conexão entre a haste de âncora e haste de referência.

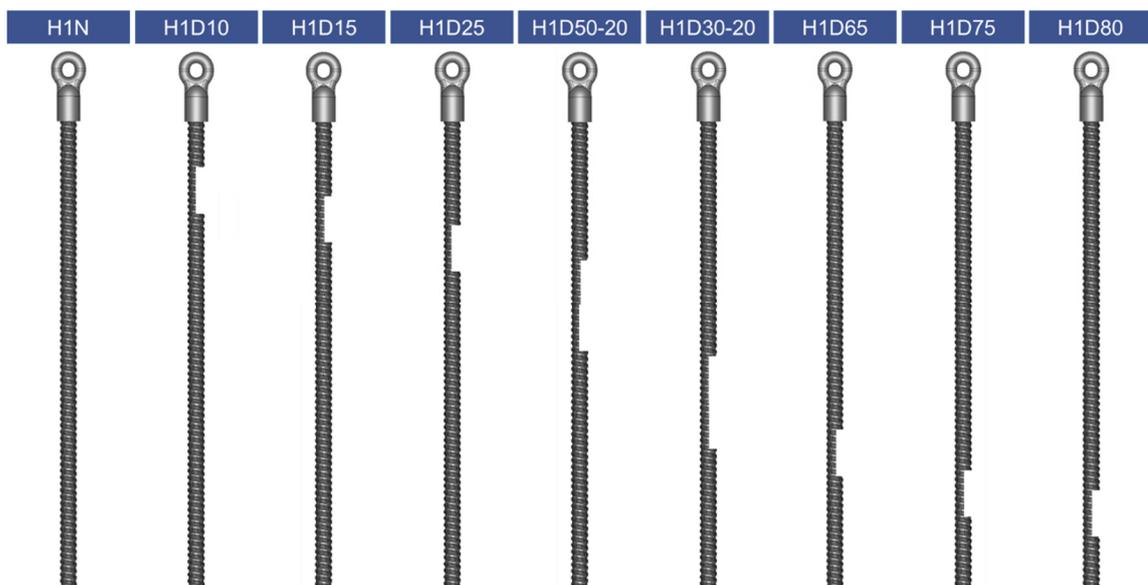


Fonte: O autor, 2020.

Os sinais medidos nas hastes serão analisados por uma aplicação de inteligência artificial especializada em diagnosticar falhas a partir desses dados. Essa aplicação consiste em um modelo de aprendizado de máquina treinado por meio da apresentação de um conjunto de exemplos rotulados previamente de sinais oriundos tanto de hastes normais quanto de hastes defeituosas. A base de dados utilizada como entrada para os algoritmos de aprendizado de máquina será composta por amostras oriundas de medição e de simulação, para obtenção de uma quantidade maior de dados para treinamento dos modelos.

Para compor a base de dados de treinamento e validação do sistema, foram confeccionadas 09 hastes de âncora para testes, cujo aspecto é apresentado na Figura 24. Essas hastes de teste possuem 1,0 m de comprimento e são compostas pelo mesmo material e configuração estrutural que aquelas utilizadas nas torres em campo. A haste de referência utilizada foi uma haste padrão de aterramento, feita em cobre, com 1,0 m de comprimento e 11,0 mm de diâmetro posicionada paralelamente a 8,0 cm de distância da haste de âncora. Uma das hastes de âncora foi mantida intacta e nas 08 restantes foram inseridas falhas de diferentes configurações em pontos distintos. A nomenclatura utilizada para a identificação das hastes de teste tem por objetivo informar o comprimento da haste, bem como a localização e extensão da falha inserida, conforme apresentado na Tabela 2 e detalhado a seguir com exemplos (COUTINHO, LOURENÇO NOVO, *et al.*, 2019).

Figura 24 – Conjunto de hastes de teste confeccionadas para medição.



Fonte: O autor, 2020.

A nomenclatura se inicia com uma letra “H”, indicando haste, seguida de um dígito informando seu comprimento total. Dessa forma, a nomenclatura H1N se refere a uma haste de 1,0 m de comprimento “H1”, sem nenhum tipo de desgaste, o que é indicado pelo código “N”. Hastes defeituosas possuem o código “D” e o defeito padrão inserido nas falhas possui 10 cm de comprimento. Para o caso de hastes em que foi inserido o defeito padrão, a informação referente ao comprimento do defeito é omitida na nomenclatura, para não carregar desnecessariamente a notação. No caso de hastes defeituosas, os dois dígitos que seguem o código “D”, representam o ponto de início do defeito em relação a posição da porta do conector. Por exemplo, a nomenclatura H1D15 é relativa a uma haste de 1,0 m de comprimento “H1”, com um defeito “D” de 10,0 cm de extensão (padrão) se iniciando a 15,0 cm do conector. Finalmente, nos casos em que o defeito inserido possuir um comprimento diferente do padrão, essa informação é inserida na nomenclatura como um sufixo. Por exemplo, seja a haste rotulada como H1D30-20. Neste caso, se trata de uma haste de 1,0 m de comprimento “H1” com um defeito que se inicia a 30,0 cm do conector “D30” e que possui 20 cm de extensão “-20”.

Tabela 2 – Configuração das hastes de âncora de teste utilizadas.

Identificação da haste	H1N	H1D10	H1D15	H1D25	H1D30-20	H1D50-20	H1D65	H1D75	H1D80
Comprimento da haste (cm)	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Diâmetro da haste (mm)	33,0	33,0	33,0	33,0	33,0	33,0	33,0	33,0	33,0
Início da falha (cm)	N/A	10,0	15,0	25,0	30,0	50,0	65,0	75,0	80,0
Fim da falha (cm)	N/A	20,0	25,0	35,0	50,0	70,0	75,0	85,0	90,0
Comprimento da falha (cm)	N/A	10,0	10,0	10,0	20,0	20,0	10,0	10,0	10,0
Profundidade da falha (mm)	N/A	16,5	16,5	16,5	16,5	16,5	16,5	16,5	16,5

Fonte: O autor, 2020.

Para cada uma das hastes de teste, o procedimento de medição completo contemplando a calibração do analisador de redes, a conexão das hastes e a aquisição dos dados foi repetido 30 vezes com o intuito de capturar erros e variâncias inerentes ao processo de medição. Esses erros incluem eventuais folgas nos conectores, variações no posicionamento da conexão com a âncora e ausência de paralelismo entre as hastes. Após a coleta dos dados, o sinal médio entre as 30 medições realizadas para cada haste foi calculado. Visando minimizar a presença de

amostras espúrias na base de dados, são descartadas as 02 medições dentre as 30 coletadas que apresentam o menor coeficiente de correlação cruzada de Pearson (HEUMANN e SHALABH, 2016) em relação ao sinal médio. Dessa forma, restam, ao final do processo, 28 aquisições válidas para cada uma das 09 hastes medidas, totalizando 252 sinais úteis que serão utilizados para compor a base de dados.

4.1 MODELOS COMPUTACIONAIS SIMULADOS

Medições realizadas em hastes de âncora operacionais em que fossem encontrados os mais variados tipos de falhas estruturais poderiam ser utilizadas para construir uma base de dados e treinar os algoritmos de aprendizado de máquina. Embora essa metodologia seja, em teoria, possível, tal procedimento para composição da base de dados não seria viável em termos práticos. Além da infraestrutura logística e tempo requeridos para realizar as medições em campo, os reais estados das hastes só poderiam ser indubitavelmente definidos após a escavação e inspeção visual das mesmas. Dessa forma, a medição das hastes em campo e a classificação dos dados medidos seria um processo destrutivo, de alto custo e demandaria uma grande quantidade de tempo, além de não haver nenhuma garantia *a priori* de que haveria dentre as hastes escavadas uma diversidade de condições de falha tal que garanta uma base de treinamento supervisionado adequada para os algoritmos de aprendizado de máquina.

Devido às restrições citadas, uma forma alternativa e viável de montar uma base de dados suficientemente extensa e que permita o devido treinamento dos algoritmos se dá por meio de modelagem computacional em simuladores de alta fidelidade. Esses simuladores utilizam ferramentas de edição gráfica e processamento dos dados por métodos finitos, permitindo a avaliação das respostas dos parâmetros eletromagnéticos do sistema em diversas configurações distintas de falha, o que seria impraticável de ser realizado apenas com medições.

Os modelos computacionais foram desenvolvidos no *software* HFSS (*High Frequency Structure Simulator*) da Ansys, que utiliza a técnica de *tangential vector finite elements* para eliminação de soluções numéricas espúrias, mas que atendem as equações do modelo (CENDES, 2016). Um exemplo de modelo computacional desenvolvido no HFSS é mostrado na Figura 25, onde são descritas as quatro propriedades da falha que são modificadas a cada simulação: formato, posição, comprimento e profundidade. Um total de 1925 simulações distintas foram geradas para as hastes de 1,0 m de comprimento, pela modificação conjunta de cada uma dessas quatro propriedades, cujas faixas de variação são apresentadas na Tabela 3.

A estratégia de aumentar a quantidade de dados por simulação, apesar de viável, ainda requer uma considerável quantidade de recursos de processamento computacional e memória para a solução dos modelos numéricos. Assim, os modelos computacionais foram simulados em dois computadores com processador Xeon 14 núcleos, 28 *Hyper Threading*, capacidade de armazenamento de 256 GB SSD e 05 TB HD e placa de vídeo NVIDIA Quadro P4000 com 128 GB de RAM.

Figura 25 – Modelo computacional de simulação das hastes experimentais utilizadas.



Fonte: O autor, 2020.

Tabela 3 – Características dos modelos computacionais simulados.

Característica	Valor ou faixa de variação
Comprimento da haste de âncora	100,0 cm
Diâmetro da haste de âncora	33,0 mm
Comprimento da haste de referência	100,0 cm
Diâmetro da haste de referência	11,0 mm
Distância entre as hastes	8,0 cm
Formato da falha	Corte reto em formato de “U” e em anel
Posição da falha	De 6,5 cm a 94,5 cm a cada 1,0 cm
Comprimento da falha	5,0 cm; 10,0 cm e 15,0 cm
Profundidade da falha	4,1 mm; 8,2 mm; 12,3 mm e 16,5 mm

Fonte: O autor, 2020.

4.2 ESTRUTURA DA BASE DE DADOS

A resposta de cada um dos parâmetros foi levantada na faixa de frequências compreendida entre 2 MHz e 1 GHz e amostrada a cada intervalo de 998 kHz, totalizando 1001

pontos para cada parâmetro, por aquisição. Essa faixa de frequências foi escolhida por permitir a observação de múltiplos pontos de ressonância, possuir um baixo deslocamento de frequência entre os sinais medido e simulado para uma mesma haste e possibilitar o desenvolvimento de um conector mecanicamente robusto para ser utilizado no arranjo experimental.

Por sua vez, a quantidade de pontos por amostragem (1001 pontos) foi definida visando a obtenção da melhor resolução possível, por ser o máximo permitido pela configuração do analisador de redes vetorial modelo *Keysight Fieldfox N9923A*, utilizado inicialmente nos experimentos. Optou-se por manter essa configuração, mesmo nas medições realizadas com o analisador de redes vetorial modelo *Keysight Fieldfox N9952A*, que possui resolução máxima de 10001 pontos, para que fosse garantida a compatibilidade dos dados, independentemente do modelo do equipamento utilizado e, assim, fosse possível utilizar ambos os analisadores.

Das três grandezas avaliadas (S_{11} , VSWR e Z_{in}), duas possuem valores complexos (S_{11} e Z_{in}), necessitando, portanto, de magnitude e fase ou de parte real e imaginária para serem corretamente representadas. Por isso, para cada haste avaliada, é gerado um vetor de parâmetros de entrada contendo a concatenação dos dados referentes aos seguintes parâmetros: magnitude de S_{11} - $Mag\{S_{11}\}$; fase de S_{11} - $Fas\{S_{11}\}$; coeficiente de onda estacionária de tensão - VSWR; parte real da impedância de entrada - $Re\{Z_{in}\}$; e parte imaginária da impedância de entrada - $Im\{Z_{in}\}$. Com o objetivo de manter a consistência e a estrutura dos dados, quer seja por medição, quer seja por simulação, cada aquisição gera 1001 pontos para cada um dos cinco parâmetros levantados, totalizando 5005 pontos por haste.

As amostras obtidas das m hastes, tanto por meio de simulações de modelos computacionais quanto pelas medições de hastes reais em bancada são organizadas em uma base de dados estruturada na forma de uma matriz híbrida de parâmetros $X_{m \times n}$, em que cada linha $x_i = (x_{i1}^1, x_{i2}^1, \dots, x_{in}^1) | (x_{i1}^2, x_{i2}^2, \dots, x_{in}^2) | \dots | (x_{i1}^5, x_{i2}^5, \dots, x_{in}^5)$ contém um vetor de entrada composto pela concatenação dos parâmetros eletromagnéticos referentes à i -ésima haste, amostrados em $n = 1001$ pontos na faixa de frequências de interesse. Adicionalmente, para permitir o treinamento supervisionado dos algoritmos, um vetor coluna $Y_{m \times 1} = (y_1, y_2, \dots, y_m)^T$ de m rótulos binários é definido, no qual a cada elemento y_i que está associado à i -ésima linha da matriz X é atribuído o valor 1 (um) se os dados são oriundos de uma haste considerada defeituosa e valor 0 (zero) caso contrário. O formato geral da base de dados pode ser visualizado na Figura 26.

Figura 26 - Formato da base de dados de parâmetros eletromagnéticos utilizada para treinamento e teste dos algoritmos de aprendizado de máquina.

Vetor	Entrada					Saída
Parâmetro	$Mag\{S_{11}\}$	$Fas\{S_{11}\}$	$VSWR$	$Re\{Z_{in}\}$	$Im\{Z_{in}\}$	<i>Estado</i>
Matriz	$X_{Mag\{S_{11}\}}$	$X_{Fas\{S_{11}\}}$	X_{VSWR}	$X_{Re\{Z_{in}\}}$	$X_{Im\{Z_{in}\}}$	<i>Y</i>
Dimensão	$m \times 1001$	$m \times 1001$	$m \times 1001$	$m \times 1001$	$m \times 1001$	$m \times 1$
Tipo	Valor Real	Valor Real	Valor Real	Valor Real	Valor Real	Valor Binário
Haste 01	DADOS: 1001 amostras em frequência na faixa de 2 MHz a 1 GHz para cada um dos 05 parâmetros Dimensão: ($m \times 5005$)					y_1
Haste 02						y_2
...						...
Haste m						y_m

Fonte: O autor, 2020.

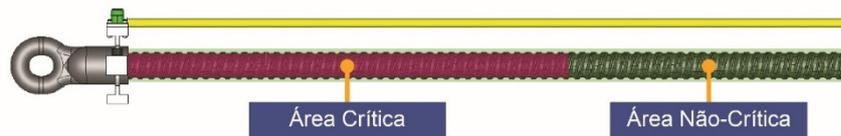
Na manutenção das hastes de âncora em campo, as operadoras das linhas de transmissão estabeleceram, por experiência e práticas rotineiras, um critério que determina que falhas superficiais e pouco extensas não comprometem significativamente a resistência à tração da haste. Por outro lado, falhas nos trechos finais da haste podem ser desprezadas por não comprometerem a ancoragem da haste, já que nesses pontos as falhas estarão normalmente cobertas da nata de cimento que compõe a fundação do sistema. Trazendo esses critérios de tolerância como referência para classificar as hastes geradas pelos modelos simulados computacionalmente, uma haste i é considerada defeituosa ($y_i = 1$) se as condições seguintes são satisfeitas: a falha tem profundidade superior a 4,1 mm; é mais extensa que 5,0 cm e ocorre dentro dos 60,0 cm iniciais da haste, contados a partir do conector.

A Figura 27 mostra de forma gráfica o critério utilizado para classificar a haste como defeituosa ou normal na base de dados. A haste é considerada defeituosa se a corrosão atinge a área crítica da haste, marcada na cor vermelha, e considerada normal se a corrosão atinge a área não-crítica marcada na cor verde. Dessa forma, quando uma haste é classificada como normal, não significa necessariamente que há ausência completa de corrosão, mas sim que o grau da falha estrutural identificada ainda não justifica a substituição da haste. Complementarmente, quando a haste é diagnosticada como defeituosa, significa que a falha causada pela corrosão é de tal forma que compromete a estrutura da ancoragem, requerendo uma ação de manutenção para substituição da haste defeituosa por uma nova.

A base de dados completa é composta por um total de $m = 2177$ amostras, sendo 252 amostras de medição e 1925 de simulação. A utilização de uma base de dados híbrida, tem a

vantagem de, por meio das simulações disponibilizar uma maior quantidade de exemplos para a generalização do classificador e, por meio das medições, introduzir no classificador efeitos dificilmente obtidos por meio de simulação, como ruídos, interferências, imperfeições físicas dos equipamentos e materiais, bem como variações inerentes ao próprio processo de medição.

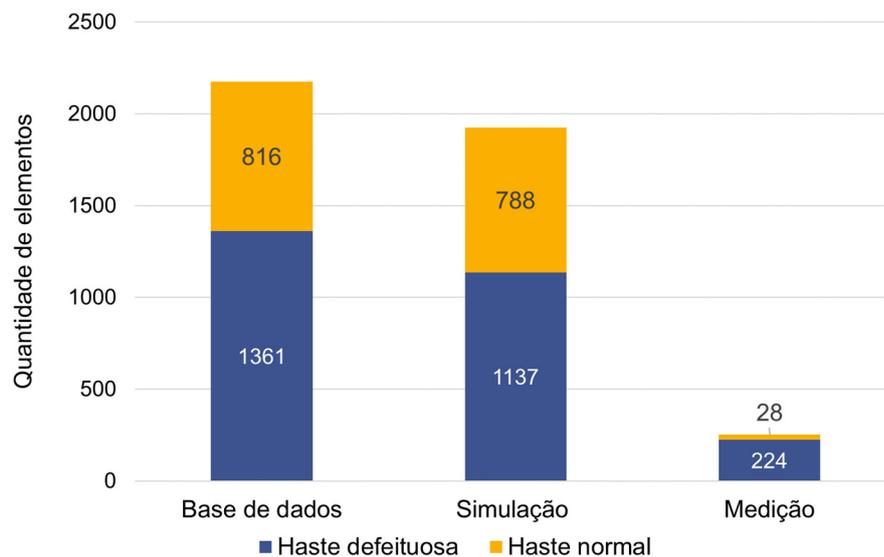
Figura 27 – Critério utilizado para classificar a haste como normal ou defeituosa.



Fonte: O autor, 2020.

Utilizando os critérios descritos para classificação das hastes, tem-se que a base de dados possui 1361 amostras normais e 816 defeituosas, estando dividida conforme mostrado na Figura 28. Essa base pode ser considerada relativamente equilibrada, pois, a proporção observada de aproximadamente 3:2 entre amostras normais e defeituosas não chega a ser uma assimetria que demande tratamento especial.

Figura 28 – Características de composição da base de dados utilizada.



Fonte: O autor, 2020.

4.3 DESENVOLVIMENTO DAS APLICAÇÕES

As aplicações de aprendizado de máquina do sistema inteligente de diagnóstico foram desenvolvidas na linguagem de programação Python (PYTHON, 2020). O Python gera um código de alto nível facilmente legível, o que otimiza o tempo de desenvolvimento dos sistemas. Além disso, Python é uma linguagem de código aberto, podendo ser livremente usada e distribuída mesmo para fins comerciais, o que a torna, também do ponto de vista econômico, uma opção extremamente viável. Foram utilizados objetos e métodos já implementados nas bibliotecas Keras (KERAS, 2016) e Scikit-Learn (PEDREGOSA, VAROQUAUX, *et al.*, 2011) que permitem a computação numérica de alto desempenho em uma variedade de plataformas e a prototipagem de soluções de inteligência artificial em alto nível, permitindo realização de simulações e experimentos de forma mais eficiente.

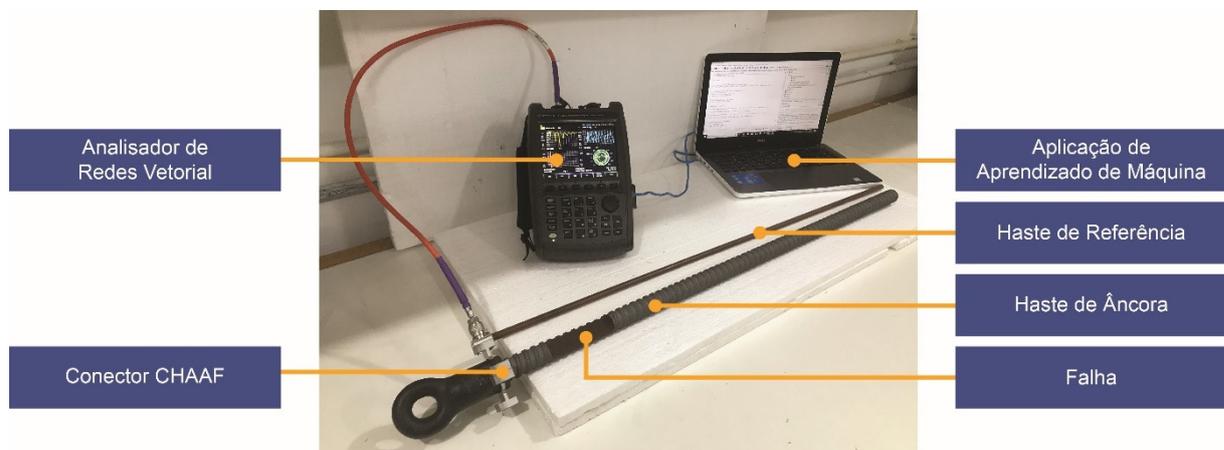
Em modo de desenvolvimento, a base de dados é devidamente estruturada e aleatoriamente dividida em dois subconjuntos disjuntos: o conjunto de treino e o conjunto de teste. O conjunto de treino contém as amostras que serão utilizadas para que o modelo aprenda a diferenciar entre os sinais correspondentes a hastes normais daqueles provenientes de hastes defeituosas. Por outro lado, o conjunto de teste é reservado para avaliar o desempenho do classificador. Dessa forma, por não terem sido utilizados durante o treinamento, os elementos do conjunto de teste são inéditos para o sistema e simulam sinais de uma amostra futura a ser classificada pelo modelo treinado.

Durante a etapa de busca de hiperparâmetros, a divisão entre a quantidade de dados nos conjuntos de treino e teste adotada foi de 80% e 20%, respectivamente. Durante a etapa de medição de desempenho, no entanto, foi adotado o processo de validação cruzada por segmentação dos dados em k subconjuntos (*k-fold cross validation*) (HASTIE, TIBSHIRANI e FRIEDMAN, 2009). Nesse processo iterativo de avaliação de desempenho, os dados são aleatoriamente separados em k subconjuntos disjuntos e, a cada iteração, um desses conjuntos é destinado para teste e os demais $k - 1$ são utilizados no treinamento do modelo. Neste trabalho, foi adotada a quantidade $k = 10$ de subconjuntos recomendada para equilibrar o viés e a variância (*bias-variance tradeoff*) na seleção de modelos (KOHAVI, 1995). Isso significa que, a cada iteração, os modelos foram treinados com 90% das amostras disponíveis e testados com as 10% restantes. Dessa forma, é considerado como valor do desempenho do modelo a média aritmética das acurácias medidas nas 10 iterações que cobrem toda a base de dados.

4.4 MODO DE OPERAÇÃO DO SISTEMA

Uma visão geral dos componentes do sistema de diagnóstico é mostrada na Figura 29. Após a conexão dos equipamentos à haste de âncora, o operador envia, por meio de um computador o comando para início da operação. O sistema então automaticamente realiza a medição dos parâmetros eletromagnéticos e o processamento dos dados adquiridos, informando ao operador em uma tela o diagnóstico referente ao estado da haste, classificando-a como normal ou defeituosa.

Figura 29 – Visão geral dos componentes do sistema de diagnóstico de falhas em hastes de âncora.



Fonte: O autor, 2020.

Internamente ao sistema, em modo de operação, ocorre que, após o comando do operador, o conjunto de parâmetros medidos é exportado em um arquivo formato “.csv” (*comma-separated-values*) gerado pelo próprio analisador de redes para um computador ou eletrônica embarcada que possui a aplicação de inteligência artificial responsável pelo tratamento e análise dos dados. Essa aplicação é composta por sistemas acessórios de pré-processamento e de algoritmos de aprendizado de máquina treinados previamente para extrair informação referente à condição das hastes a partir dos dados de cada parâmetro eletromagnético, interpretar os sinais medidos e, finalmente, classificar o estado da haste como normal ou defeituosa sem a necessidade de nenhuma intervenção humana.

4.5 INFRAESTRUTURA DISPONÍVEL PARA O PROJETO

O Departamento de Engenharia Elétrica da UFPE possui um campo experimental contendo 14 hastes de âncora com diversos comprimentos e diferentes níveis de corrosão e desgaste. Essas hastes encontram-se implantadas no solo de maneira similar à adotada nas

linhas de transmissão reais, possibilitando a medição dos parâmetros em condições semelhantes às aquelas encontradas realmente em campo, como mostrado na Figura 30.

Figura 30 – Conjunto de hastes do Campo Experimental da UFPE.



Fonte: O autor, 2020.

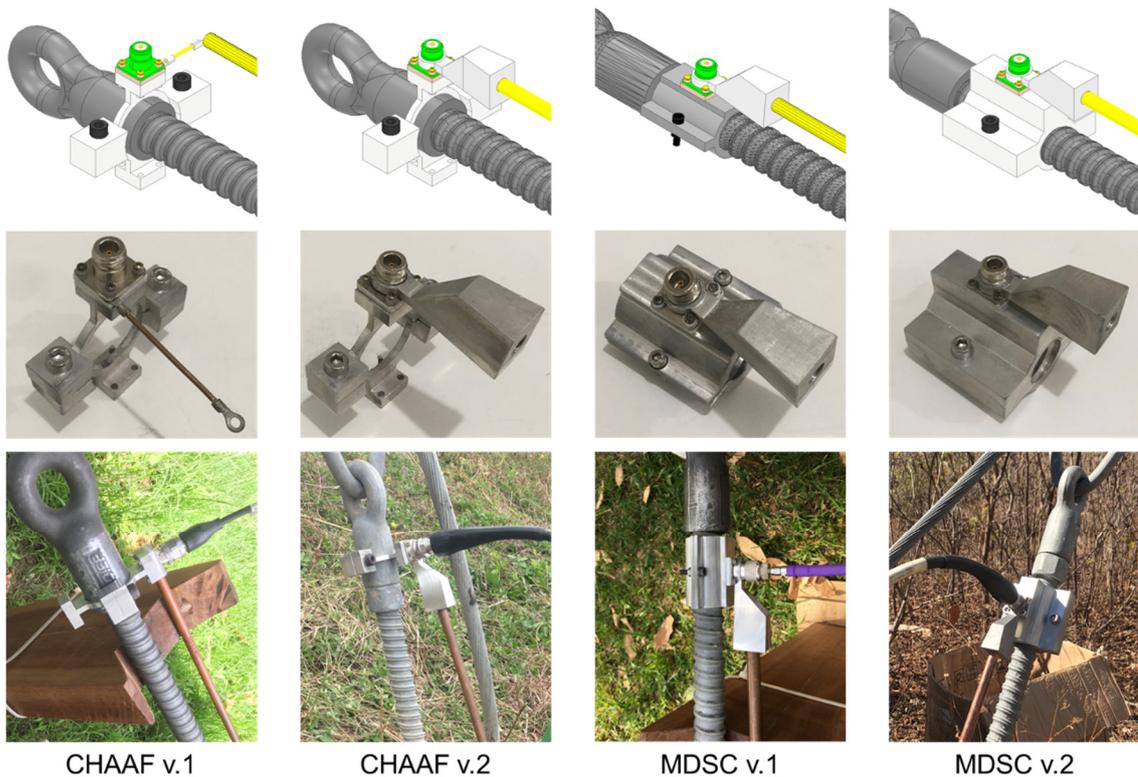
O sistema de diagnóstico protótipo foi totalmente desenvolvido, simulado e medido no Laboratório de Micro-ondas¹ da UFPE. Os modelos computacionais de alta fidelidade do sistema foram projetados e simulados no *software* HFSS (*High Frequency Structure Simulator*) da Ansys (CENDES, 2016), amplamente utilizado para simulação eletromagnética em estruturas tridimensionais. Esses *softwares* rodam em dois computadores de alto desempenho processador Xeon 14 núcleos, 28 *Hyper Threading*, capacidade de armazenamento de 256 GB SSD e 05 TB HD e placa de vídeo NVIDIA Quadro P4000 com 128 GB de RAM. O Laboratório de Micro-ondas está equipado ainda com um conjunto de analisadores de rede vetoriais (modelos Keysight, Rohde & Schwartz e Agilent Technologies).

Amostras de hastes de âncora com diferentes comprimentos e tipos de falha foram confeccionadas para a realização de medições de teste em bancada. Quatro diferentes tipos de conectores dedicados a acoplar a haste de âncora ao analisador de redes vetorial foram desenvolvidos por pesquisadores do Laboratório de Micro-ondas e estão sendo produzidos e

¹ www.ufpe.br/laboratoriomicroondas

estudados para a realização das medições em campo. Esses conectores foram denominados CHAAF v.1 (Conector Haste-Analisador para altas Frequências – versão 1) (LOURENÇO NOVO, DE MELO, *et al.*, 2014); CHAAF v.2; MDSC v.1 (*Microwave Device of Supporting and Connection – version 1*); e MDSC v.2. Os modelos de simulação computadorizada, os protótipos confeccionados e os testes em campo desse conjunto de conectores são mostrados na Figura 31.

Figura 31 – Conectores em desenvolvimento para realização de medições em campo.



Fonte: O autor, 2020.

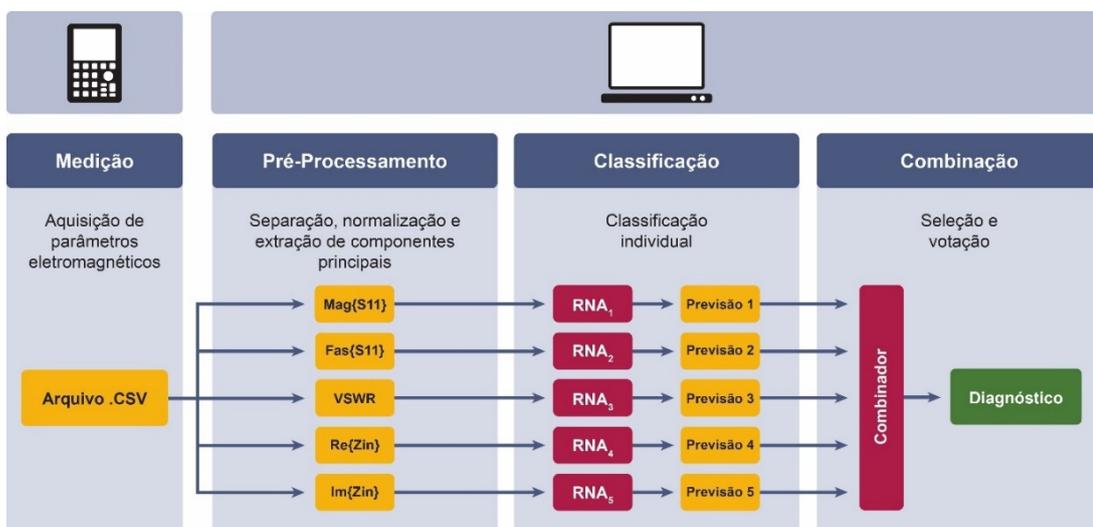
5 EXPERIMENTOS

Após a estruturação de uma base de dados para treinamento dos modelos de aprendizado de máquina, resta definir qual algoritmo é o mais adequado para a tarefa de detecção das falhas nas hastes a partir desses dados. Similarmente, é necessário avaliar qual dos parâmetros de entrada é capaz de produzir os melhores resultados, já que todos eles são, essencialmente, representações distintas e relacionadas pelo coeficiente de reflexão Γ do comportamento da onda eletromagnética que se propaga na haste. No entanto, como não há uma teoria fechada para seleção de algoritmos de aprendizado de máquina e dos dados de entrada ideais a partir de características do problema específico, tal escolha tipicamente necessita ser realizada por meio de uma bateria de testes comparativos experimentais, que avaliam um subconjunto de arquiteturas previamente selecionadas por serem soluções potenciais para o problema.

5.1 DESENVOLVIMENTO DO CLASSIFICADOR MULTIPARÂMETRO

Com o objetivo de avaliar comparativamente o desempenho do sistema de diagnóstico em função de cada diferente parâmetro eletromagnético utilizado como entrada, foi desenvolvida uma solução composta por um comitê de classificadores (*ensemble*), como mostrado na Figura 32. Cada um dos cinco classificadores autônomos (membro do comitê) é formado por uma rede neural artificial independente das demais e que recebe como entrada um vetor contendo as amostras de um dos parâmetros eletromagnéticos medidos pelo analisador de redes vetorial, após terem sido tratadas por um estágio de pré-processamento.

Figura 32 – Arquitetura do sistema proposto para avaliação de parâmetros eletromagnéticos.



Fonte: O autor, 2020.

As RNAs do estágio de classificação foram construídas com funções de ativação ReLU na camada oculta e sigmóide na camada de saída, para que o resultado gerado por cada classificador tenha uma interpretação probabilística em relação às duas classes possíveis, com um valor de limiar de 0,5 para a decisão. Para o treinamento das RNAs, foi adotada a função de perda por entropia cruzada binária (*binary cross-entropy*) (GOLIK, DOETSCH e NEY, 2013), e o algoritmo de treinamento Adam (*Adaptive Moment Estimation*) (KNIGMA e BA, 2015), por terem apresentado melhor desempenho nos testes realizados.

O comitê de classificadores do estágio combinador é capaz de selecionar qualquer associação dentre as possíveis 31 combinações não vazias existentes utilizando os cinco parâmetros disponíveis, as quais serão identificadas pelos índices mostrados na Tabela 4. Para uma determinada combinação, o estágio combinador final recebe como entradas os valores estimados por cada classificador individual ativo, calcula a média aritmética desses valores e produz como saída o resultado do diagnóstico global do sistema, associando o valor lógico 1 (um), indicando que a haste é do tipo defeituosa, caso a média obtida seja superior a 0,5 ou valor lógico 0 (zero), caso contrário.

No estágio de pré-processamento, a aplicação coleta o arquivo .csv gerado pelo analisador de redes, extrai a informação útil e a organiza em uma matriz NumPy (NumPy, 2020), o que permite a representação dos dados em um formato adequado para a realização de computação numérica de alto desempenho utilizada pelos algoritmos de aprendizado de máquina na linguagem Python (OLIPHANT, 2007). Em seguida, os dados pertencentes a cada um dos cinco parâmetros medidos são separados uns dos outros e estruturados em vetores individuais para que possam ser encaminhados pelo sistema até as suas respectivas RNAs.

Durante a etapa de treinamento do sistema, os sinais já separados por parâmetro são normalizados para que cada componente dos vetores de entrada na base de dados de treinamento tenha uma distribuição aproximadamente gaussiana com média zero e variância unitária. A normalização dos vetores de entrada da base de dados de treinamento elimina quaisquer vieses entre os seus componentes e garante que todas as amostras recebam igual tratamento pela RNA. Esse procedimento acelera o processo de aprendizado e diminui a probabilidade de o algoritmo de treinamento ficar estagnado em um ponto de mínimo local, possibilitando a obtenção de maiores taxas de acerto (SOLA e SEVILLA, 1997).

Tabela 4 – Relações entre os índices de combinação e os parâmetros eletromagnéticos.

Índice de Combinação	Parâmetros associados	Índice de combinação	Parâmetros associados
1	MagS11	17	MagS11 + FasS11 + ReZin
2	FasS11	18	MagS11 + FasS11 + ImZin
3	VSWR	19	MagS11 + VSWR + ReZin
4	ReZin	20	MagS11 + VSWR + ImZin
5	ImZin	21	MagS11 + ReZin + ImZin
6	MagS11 + FasS11	22	FasS11 + VSWR + ReZin
7	MagS11 + VSWR	23	FasS11 + VSWR + ImZin
8	MagS11 + ReZin	24	FasS11 + ReZin + ImZin
9	MagS11 + ImZin	25	VSWR + ReZin + ImZin
10	FasS11 + VSWR	26	MagS11 + FasS11 + VSWR + ReZin
11	FasS11 + ReZin	27	MagS11 + FasS11 + VSWR + ImZin
12	FasS11 + ImZin	28	MagS11 + FasS11 + ReZin + ImZin
13	VSWR + ReZin	29	MagS11 + VSWR + ReZin + ImZin
14	VSWR + ImZin	30	FasS11 + VSWR + ReZin + ImZin
15	ReZin + ImZin	31	MagS11 + FasS11 + VSWR + ReZin + ImZin
16	MagS11 + FasS11 + VSWR		

Fonte: O autor, 2020.

Considerando a base de dados de treinamento do k -ésimo parâmetro eletromagnético $X_{m \times n}^k$, em que cada linha é um vetor de entrada de n componentes na forma $x_i^k = (x_{i1}^k, x_{i2}^k, \dots, x_{in}^k)$. O processo de normalização se inicia calculando a média u_j^k e o desvio padrão s_j^k , de cada uma das n componentes dispostas ao longo das colunas do conjunto de treinamento, conforme (90) e (91), respectivamente.

$$u_j^k = \frac{1}{m} \sum_{i=1}^m x_{ij}^k \quad (90)$$

$$s_j^k = \left[\frac{1}{m} \sum_{i=1}^m (x_{ij}^k - u_j^k)^2 \right]^{\frac{1}{2}} \quad (91)$$

De posse dos valores de u_j^k e s_j^k para todas as colunas da base de dados de treinamento, é feita a normalização dos dados, substituindo cada elemento x_{ij}^k pelo seu valor normalizado z_{ij}^k por meio da expressão (92). Assim, a matriz de treinamento original $X_{m \times n}^k$ é convertida na matriz normalizada $Z_{m \times n}^k$.

$$z_{ij}^k = \frac{x_{ij}^k - u_j^k}{s_j^k} \quad (92)$$

Após a normalização dos dados, os valores da média \bar{z}_j^k e do desvio padrão σ_j^k , para cada um dos parâmetros ao longo de qualquer coluna $1 \leq j \leq n$ da matriz de treinamento normalizada $Z_{m \times n}^k$ estarão distribuídos de modo que se observe as condições (93) e (94). Os valores da média u_j^k e do desvio padrão s_j^k de normalização das componentes, obtidos durante o treinamento do modelo, são armazenados internamente no sistema para utilização posterior, no modo de operação.

$$\bar{z}_j^k = \frac{1}{m} \sum_{i=1}^m z_{ij}^k \cong 0 \quad (93)$$

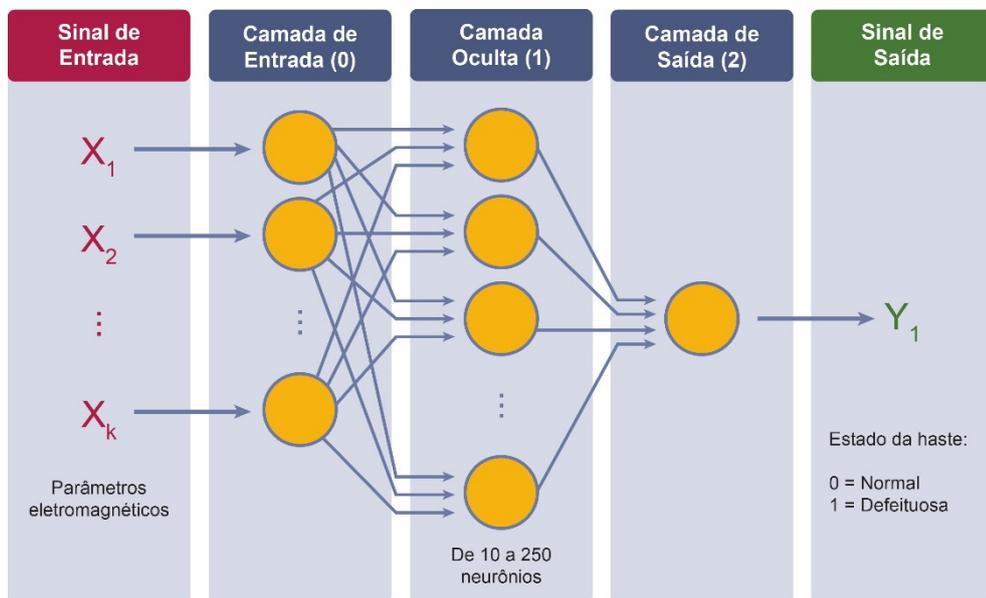
$$\sigma_j^k = \left[\frac{1}{m} \sum_{i=1}^m (z_{ij}^k - \bar{z}_j^k)^2 \right]^{\frac{1}{2}} \cong 1 \quad (94)$$

Na base de dados normalizada, é realizada uma redução de dimensionalidade dos dados através da técnica de análise de componentes principais (BALDI e HORNÍK, 1989). Como será mostrado adiante na Seção 6.1, essa estratégia permite reduzir a quantidade de dados a serem processados pelo algoritmo e, conseqüentemente, o tempo necessário para treinamento dos modelos durante a etapa de busca dos hiperparâmetros das RNAs. Para cada um dos cinco parâmetros eletromagnéticos foi utilizada apenas a quantidade de componentes principais k suficiente para reter 99% da variância do sinal. As bases utilizadas para a decomposição dos sinais originais em suas componentes principais são armazenadas internamente no sistema para utilização quando em modo de operação, se desejado.

Os dados separados por parâmetro, normalizados e com dimensionalidade reduzida são então utilizados para definir o conjunto de hiperparâmetros das RNAs. A busca por hiperparâmetros foi realizada por meio de testes e simulações, em que o desempenho dos modelos foi avaliado em função da quantidade de neurônios da camada oculta. A quantidade de neurônios testada variou numa faixa entre $(k_{min}/2)$ e $(2k_{max})$ em que k é o número de componentes principais que retém 99% da variância para cada um dos parâmetros. Em termos numéricos, foram exploradas RNAs com camadas ocultas de 10 a 250 neurônios, em passos incrementais de 10 neurônios, totalizando 25 configurações distintas.

Após a definição do conjunto de hiperparâmetros, as RNAs de 1 a 5 do estágio de classificação apresentaram a configuração exibida na Figura 33, diferindo entre si apenas no que se refere à quantidade de neurônios na camada oculta, dada em função do parâmetro eletromagnético que foi utilizado como entrada para cada RNA. Uma vez definidas as arquiteturas de RNA com melhor desempenho para cada parâmetro individualmente, foi finalmente realizada a avaliação dos resultados do comitê de classificadores.

Figura 33 – Estrutura geral das RNAs no estágio de classificação.



Fonte: O autor, 2020.

Após a realização das simulações para definição dos hiperparâmetros das RNAs, foram selecionadas as arquiteturas descritas na Tabela 5, que apresentaram o melhor desempenho individual durante os testes.

Tabela 5 – Hiperparâmetros selecionados para as RNAs do classificador multiparâmetro.

Hiperparâmetro	RNA 1	RNA 2	RNA 3	RNA 4	RNA 5
Parâmetro de entrada	$\text{Mag}\{S_{11}\}$	$\text{Fas}\{S_{11}\}$	VSWR	$\text{Re}\{Z_{in}\}$	$\text{Im}\{Z_{in}\}$
Componentes Principais (entradas)	40	102	44	42	36
Neurônios na camada oculta	100	100	180	220	180
Neurônios na camada de saída	1	1	1	1	1

Fonte: O autor, 2020.

Para cada um dos 31 índices relacionados na Tabela 4 foi realizado um conjunto de 3 baterias de teste, nas quais o desempenho dos modelos foi avaliado pelo método de validação cruzada com 10 subconjuntos disjuntos, totalizando 30 iterações por combinação. Por fim, os desempenhos das 31 combinações de parâmetros foram comparados, considerando um intervalo de confiança bilateral de 95%, em relação à acurácia média obtida por cada configuração nas 30 iterações de teste realizadas.

5.2 AVALIAÇÃO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA

Os cinco algoritmos de aprendizado supervisionado descritos no Capítulo 3 foram avaliados na tarefa de detecção de falhas estruturais nas hastes de âncora. Tais algoritmos foram selecionados para avaliação devido ao fato de serem amplamente difundidos, apresentarem bom desempenho em tarefas de classificação (WU, KUMAR, *et al.*, 2008) e de já possuírem uma grande parte de suas funções básicas implementadas em bibliotecas de *software* de código aberto confiáveis (PEDREGOSA, VAROQUAUX, *et al.*, 2011), o que reduz quaisquer parcialidades nos resultados devido à diferenças na eficiência de implementações particulares feitas localmente.

Os hiperparâmetros dos algoritmos de aprendizado de máquina, embora sejam a essência dos modelos, são tipicamente escolhidos por regras empíricas, heurísticas, ou por experiência prática (YUEN e LAM, 2006), uma vez que são altamente dependentes das características do problema específico. O dimensionamento ótimo da quantidade de camadas e neurônios de uma rede neural artificial, por exemplo, é uma questão ainda em aberto e que envolve um vasto campo de pesquisa (KARSOLIYA, 2012), (SHEELA e DEEPA, 2013). Assim, o modo mais eficaz de definir esses parâmetros é através de ciclos iterativos de simulações, testes e ajustes de refinamento. Os hiperparâmetros selecionados para os algoritmos testados são listados na Tabela 6. As configurações consideradas satisfatórias foram obtidas

após uma série de simulações e ajustes iterativos nos modelos, em função dos parâmetros configuráveis de cada algoritmo.

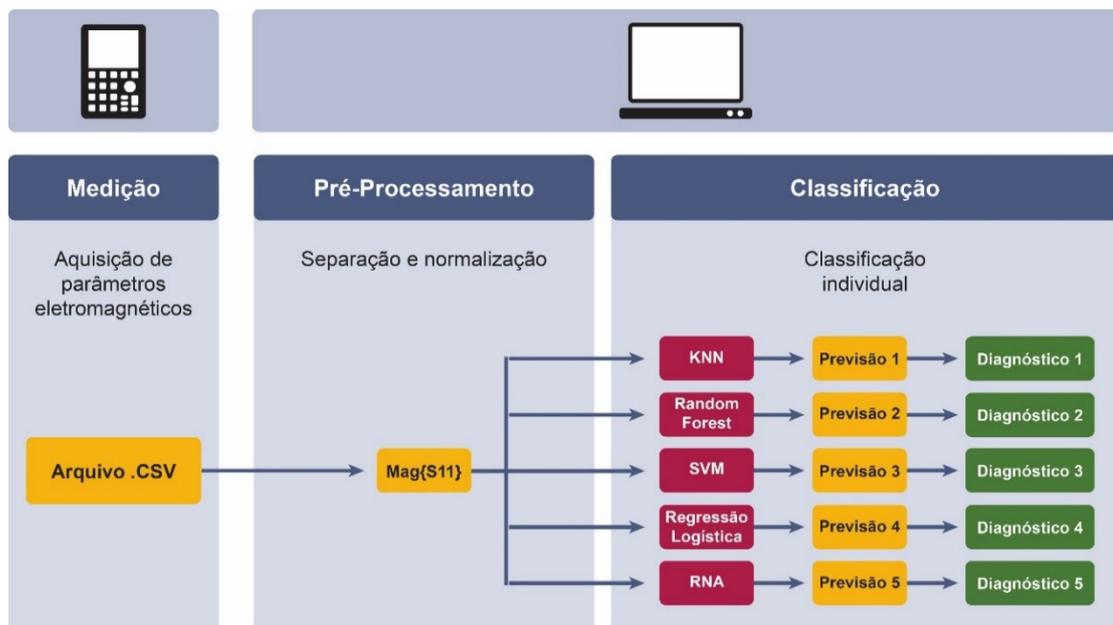
Tabela 6 – Hiperparâmetros escolhidos para cada um dos algoritmos.

Algoritmo	Hiperparâmetros
Regressão Logística	$C = 1$
SVM	Kernel: Gaussian; $C = 125.000$
<i>Random Forests</i>	Número de árvores: 300; Critério: Entropia
K-NN	Métrica: Distância euclidiana; Número k de vizinhos: 05
RNA	Tipo: FF-MLP; Neurônios: 500x100x1; Funções de ativação: ReLU (camada oculta), sigmóide (camada de saída); Função de perda: Entropia cruzada binária; Otimizador: Adam; Épocas: 200.

Fonte: O autor, 2020.

Para comparação do desempenho dos algoritmos, foi utilizada o esquema de teste apresentado na Figura 34. O sinal referente ao parâmetro $Mag\{S_{11}\}$ foi extraído dos dados coletados, normalizado e então aplicado como entrada em cada um dos cinco diferentes modelos de classificador avaliados. Neste estudo particular, apenas o desempenho individual obtido por cada algoritmo foi avaliado. A análise dos desempenhos obtidos por comitês de diferentes classificadores ou por associação de parâmetros foi reservada a estudos futuros.

Figura 34 – Esquema de teste utilizado para análise comparativa entre algoritmos.



Fonte: O autor, 2020.

Visando a obtenção de um classificador mais simples e rápido para a comparação dos algoritmos nos testes, foi utilizado como parâmetro de entrada apenas o sinal $Mag\{S_{11}\}$ (magnitude de S_{11}). Esse parâmetro adimensional, medido em dB (decibéis), é o mais comumente disponível para medição em analisadores de rede portáteis de baixo custo, além de ter demonstrado nos experimentos ser o que oferece a melhor visualização da resposta da haste em relação a seus pontos de ressonância.

A avaliação do desempenho obtido por cada algoritmo de classificação foi feita por meio dos indicadores: erro, precisão, sensibilidade e f1score. A definição desses indicadores é dada na Tabela 7, sendo t_p o número de verdadeiros positivos, t_n o de verdadeiros negativos, f_p o de falsos positivos e f_n o de falsos negativos (GOUTTE e GAUSSIER, 2005). As faixas de variação dos indicadores para cada algoritmo testado foram obtidas por meio do processo de validação cruzada por segmentação dos dados em 10 subconjuntos, e consideradas dentro de um intervalo de confiança bilateral de 95% (MITCHELL, 1997).

Tabela 7 – Indicadores utilizados para avaliação do desempenho dos algoritmos.

Erro	Precisão (<i>Pre</i>)	Sensibilidade (<i>Sen</i>)	f1Score
$\frac{f_p + f_n}{t_p + f_p + t_n + f_n}$	$\frac{t_p}{t_p + f_p}$	$\frac{t_p}{t_p + f_n}$	$\frac{2 \cdot Pre \cdot Sen}{Pre + Sen}$

Fonte: O autor, 2020.

A interpretação desses indicadores é importante para avaliar eventuais desequilíbrios ou parcialidades indesejadas nos resultados do classificador. Para uma determinada base de dados, esses indicadores são dependentes dos hiperparâmetros dos modelos e servem para que seja estabelecido um equilíbrio entre a quantidade de falsos negativos e falsos positivos gerados pelo classificador. Como esses fatores são concorrentes, o ponto de equilíbrio ideal é fortemente dependente do problema específico que está sendo abordado, uma vez que há casos mais tolerantes a falsos positivos que a falsos negativos e vice-versa.

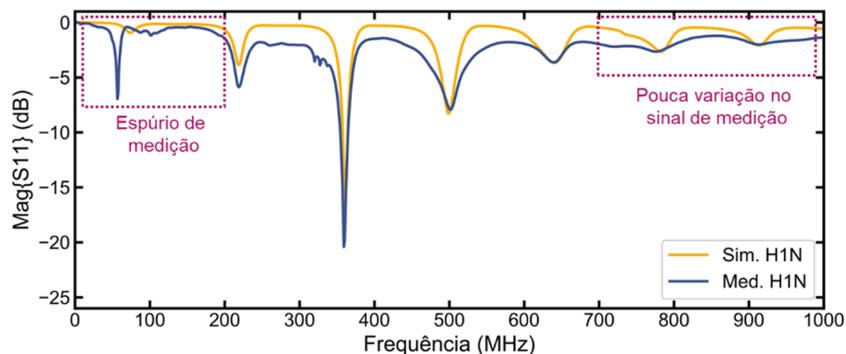
Para o problema específico de detecção de falhas nas hastes de âncora, a precisão pode ser interpretada como a proporção de hastes verdadeiramente defeituosas dentre todas as que foram classificadas pelo modelo como tal. Dessa forma, a precisão estima o grau de certeza que o classificador possui quando indica que uma determinada haste é defeituosa. Por outro lado, a sensibilidade pode ser entendida como a proporção de hastes verdadeiramente defeituosas que

o modelo conseguiu classificar corretamente dentre todas as hastes defeituosas existentes na base de dados. Assim, a sensibilidade avalia a capacidade do modelo em detectar uma haste defeituosa. Por sua vez, o f1score é uma medida alternativa de avaliar a acurácia, que leva em consideração a proporção entre a quantidade de elementos em cada categoria, sendo mais indicada, portanto, para bases de dados com fortes assimetrias entre o tamanho das classes.

6 RESULTADOS E DISCUSSÕES

Avaliar as características dos sinais de entrada do sistema é uma atividade de fundamental importância, pois pode direcionar ações e ajustes na configuração e arquitetura dos modelos de aprendizado de máquina adotados, de indicar a necessidade de pré-processamento adicional ou mesmo da geração de novos dados. A Figura 35 exibe o sinal de magnitude de S_{11} medido e simulado na faixa de 2 MHz a 1 GHz para um exemplo de haste normal, sem falhas estruturais, de 1,0 m de comprimento e identificada como H1N. Do gráfico, se observa uma grande aderência do modelo simulado em relação ao sinal medido, apesar da alta complexidade geométrica do sistema real. É possível notar, no entanto, a presença de um pico espúrio no sinal medido na faixa de frequências mais baixas, compreendida entre 2 MHz e 200 MHz e que causa uma diferença significativa entre as respostas observadas entre esses sinais nessa faixa. Essa diferença observada entre os sinais é gerada por imperfeições da haste real que não são abrangidas pelos modelos simulados e pela existência de submúltiplos harmônicos da frequência fundamental que são evidenciados por sofrerem relativamente menos atenuação nas frequências mais baixas do espectro.

Figura 35 – Comparação entre os sinais medido e simulado da magnitude de S_{11} para uma haste normal (H1N) na faixa de frequências de 2 MHz a 1 GHz.



Fonte: O autor, 2020.

Na faixa de frequências mais altas, compreendida entre 700 MHz e 1 GHz, o sinal medido apresentou variação pouco significativa, pois, com o aumento da frequência, os múltiplos harmônicos da frequência fundamental sofrem relativamente mais atenuação que aqueles posicionados em frequências mais baixas. Por conta desse comportamento, as componentes de frequência superiores a 700 MHz conferem pouca contribuição em termos de informação adicional ao sistema de diagnóstico. Assim, analisando comparativamente os dados

obtidos, adotou-se como faixa útil o trecho compreendido entre 200 MHz e 700 MHz, que é o que melhor atende ao binômio aderência entre os resultados dos modelos simulados e medidos *versus* quantidade de informação útil fornecida pelas componentes de frequência do sinal.

A estratégia de considerar apenas as amostras contidas na banda de 200 MHz a 700 MHz reduz a dimensão dos vetores de entrada originais de 1001 para 500 pontos e demanda o processamento apenas do trecho dos sinais simulados que apresentou maior grau de similaridade em relação aos sinais experimentais medidos. Tomando como referência o caso da haste normal íntegra H1N, a correlação de Pearson (HEUMANN e SHALABH, 2016) entre os sinais medido e simulado nessa faixa variou em função do parâmetro eletromagnético avaliado como mostrado na Tabela 8 para a faixa de 200 MHz a 700 MHz.

Tabela 8 – Correlação entre os sinais medidos e simulados da haste normal H1N.

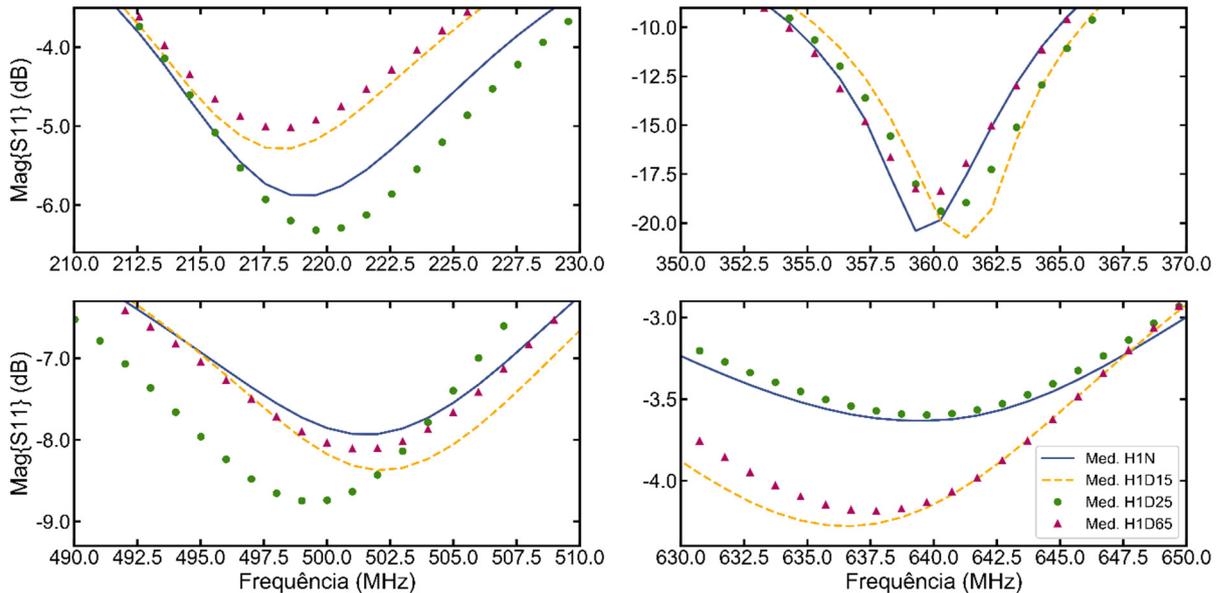
Parâmetro	$Mag\{S_{11}\}$	$Fas\{S_{11}\}$	$VSWR$	$Re\{Z_{in}\}$	$Im\{Z_{in}\}$
Correlação	0,931	0,913	0,763	0,905	0,827

Fonte: O autor, 2020.

As diferenças entre as respostas das hastes normais e defeituosas são sutis e distribuídas por todo o sinal, podendo ser observadas nas vizinhanças das frequências de ressonância. A Figura 36 exibe o detalhamento dos sinais medidos da haste normal H1N e das hastes defeituosas H1D15, H1D25 e H1D65, apresentadas na Figura 24 e descritas na Tabela 2, em uma faixa de 20,0 MHz de largura centrada nos pontos de ressonância.

Devido às interações entre as ondas incidentes e refletidas em cada trecho da haste, interferências construtivas e destrutivas ocorrem, modificando a resposta do sinal. A influência desse comportamento ondulatório periódico, faz com que os deslocamentos na frequência de ressonância e as variações de amplitude do sinal não variem monotonamente de acordo com a posição da falha. Esse aspecto, associado ao fato de o sistema possuir uma complexa geometria, torna inviável buscar uma expressão analítica ou formular algum tipo de critério gráfico de classificação dos sinais para diagnóstico das falhas e, por isso, a construção de um classificador baseado em regras, por meio de programação convencional se torna uma tarefa extremamente complexa e inviável. Tais fatores reforçam a necessidade da utilização de um algoritmo de inteligência artificial que seja capaz de realizar o diagnóstico das falhas por meio da identificação de padrões nos sinais de resposta das hastes à uma excitação eletromagnética de alta frequência que não sejam facilmente perceptíveis por um analista humano.

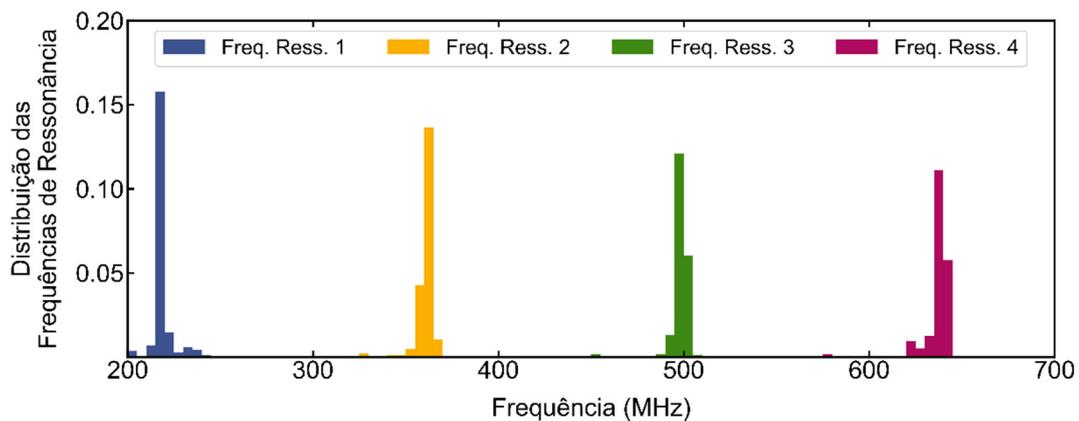
Figura 36 – Detalhe do parâmetro Magnitude de S_{11} medido para as hastes H1N, H1D15, H1D25 e H1D65 nas vizinhanças dos 04 principais pontos de ressonância.



Fonte: O autor, 2020.

A Figura 37 exibe graficamente na forma de histogramas os perfis de distribuição das frequências de ressonância nas respostas dos parâmetros na faixa de 200 MHz a 700 MHz para todos os elementos da base de dados. O primeiro ponto de ressonância ocorre em uma frequência por volta de 220 MHz e os demais são aproximadamente espaçados por um intervalo de frequência em torno de 140 MHz. Os histogramas são bem estreitos e apresentam pouca dispersão lateral, indicando que as posições das frequências de ressonância sofrem pouca influência das diferentes configurações de haste avaliadas.

Figura 37 – Distribuição das frequências de ressonância.



Fonte: O autor, 2020.

A similaridade entre os sinais medidos e simulados também pode ser observada pelos perfis das frequências de ressonância. A Tabela 9 apresenta um resumo comparativo das características dos sinais de magnitude de S_{11} medido e simulado para o caso da haste normal H1N nas frequências de ressonância. Complementarmente, a Tabela 10 apresenta essas mesmas características, mas para o exemplo de haste defeituosa H1D15.

Tabela 9 – Características do Sinal de Magnitude do S_{11} para a haste normal (H1N).

Característica	Ress1	Ress2	Ress3	Ress4
Frequência de ressonância – medição (MHz)	217,564	357,284	499,998	637,722
Frequência de ressonância – simulação (MHz)	216,566	359,280	497,004	638,720
Desvio de frequência absoluto (med-sim) (MHz)	0,998	1,996	2,994	0,998
Desvio de frequência percentual (med-sim) (%)	0,46	0,55	0,60	0,16
Amplitude na ressonância – medição (dB)	-5,85	-21,10	-8,12	-3,72
Amplitude na ressonância – simulação (dB)	-3,88	-16,61	-8,29	-3,68

Fonte: O autor, 2020.

Tabela 10 – Características do Sinal de Magnitude do S_{11} para a haste defeituosa (H1D15).

Característica	Ress1	Ress2	Ress3	Ress4
Frequência de ressonância – medição (MHz)	216,566	359,280	499,998	634,728
Frequência de ressonância – simulação (MHz)	216,566	360,278	498,002	637,722
Desvio de frequência absoluto (med-sim) (MHz)	0,00	0,998	1,996	2,994
Desvio de frequência percentual (med-sim) (%)	0,00	0,28	0,40	0,47
Amplitude na ressonância – medição (dB)	-5,28	-20,75	-8,37	-4,28
Amplitude na ressonância – simulação (dB)	-3,88	-23,37	-7,77	-3,99

Fonte: O autor, 2020.

Cada uma das 1.925 configurações distintas de falha simuladas levou em média cerca 22 minutos para ser concluída, totalizando aproximadamente 700 horas de processamento para geração da base de dados completa. Apesar do relativo alto custo computacional, a abordagem de geração da base de dados por modelos simulados ainda é significativamente mais viável do que por meio de medições em hastes físicas em campo ou em laboratório.

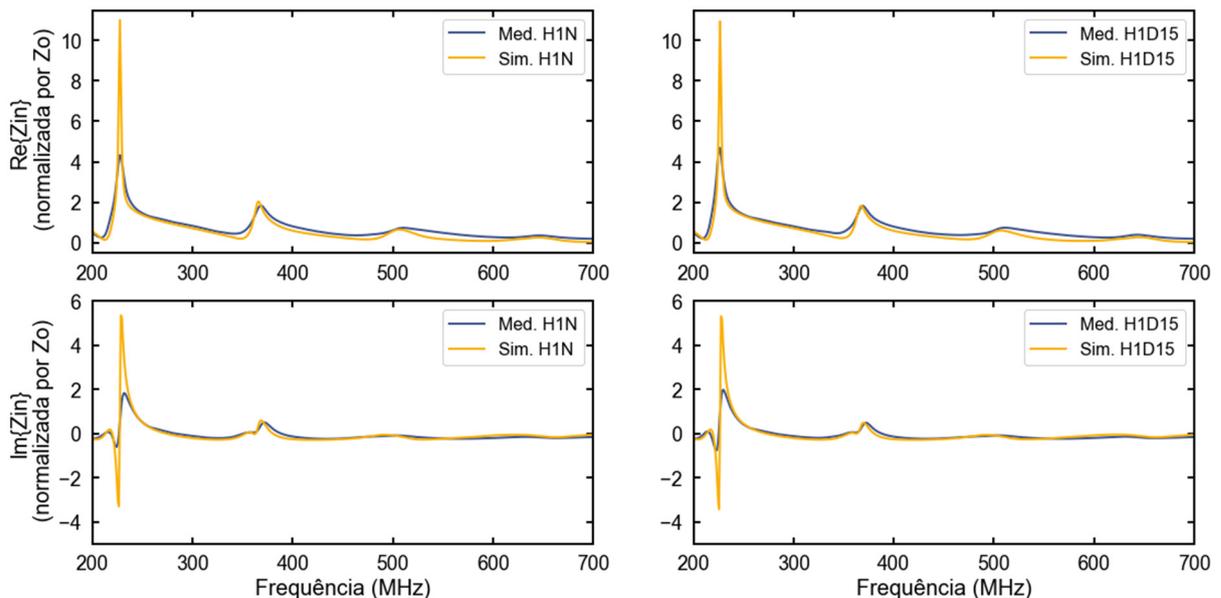
Em termos comparativos, conforme dados e observações realizadas *in loco*, cada medição em campo leva em torno de 30 minutos para ser concluída, mas deve ser seguida da

escavação da haste, que consome aproximadamente 12 horas de trabalho para ser completada. Mesmo sem considerar improdutividade e eventuais imprevistos, numa estimativa otimista essa alternativa demandaria mais de 24.000 horas de trabalho para construção de uma base de dados semelhante. Por outro lado, em laboratório, o processo de medição de uma haste leva aproximadamente 10 minutos para ser realizado, porém, demandaria a confecção, transporte, medição e catalogação de cerca de duas mil hastes físicas distintas, o que também se mostra uma alternativa inviável em termos de custo e prazo.

6.1 ANÁLISE DOS SINAIS MEDIDOS E SIMULADOS DA BASE DE DADOS

A base de dados é composta por sinais de cinco parâmetros eletromagnéticos distintos, que possuem comportamentos diversos. A Figura 38 exibe o comportamento típico dos sinais dos parâmetros parte real e imaginária da impedância de entrada normalizados em função da impedância característica de 50Ω . Os sinais são mostrados para o caso da haste normal H1N e para um exemplo de haste defeituosa contendo uma falha de 10,0 cm de comprimento, 16,5 mm de profundidade e localizada a 15,0 cm de distância da porta, denominada H1D15, cujos detalhes construtivos são apresentados na Figura 24 e descritas na Tabela 2. Dos gráficos apresentados na Figura 38 pode-se observar uma grande superposição entre os sinais medido e simulado tanto pro caso da haste normal quanto pro caso da haste defeituosa.

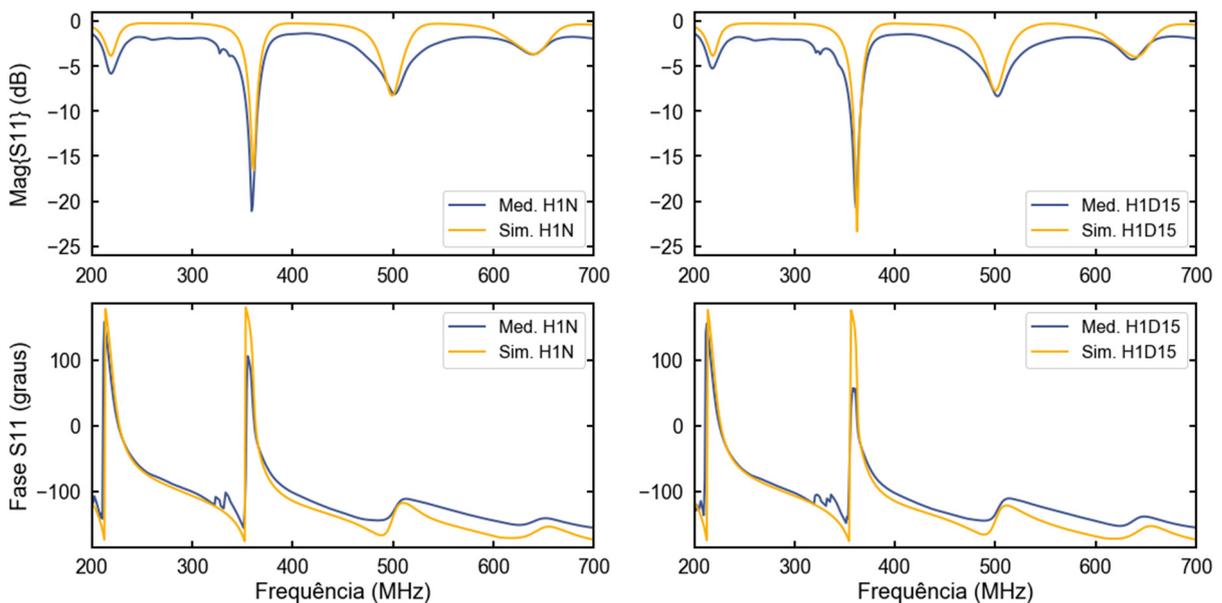
Figura 38 – Comparação entre os sinais medido e simulado das partes real e imaginária do Z_{in} (normalizado por Z_0) para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15).



Fonte: O autor, 2020.

Os sinais medido e simulado da magnitude e fase do parâmetro S_{11} para as hastes H1N e H1D15 são mostrados na Figura 39. Nas respostas do parâmetro S_{11} pode ser observada com maior clareza a existência de quatro pontos de ressonância, onde os sinais apresentaram valores de magnitude abaixo de -3 dB na faixa entre 200 MHz e 700 MHz. Comparando os resultados gerados pelas simulações com aqueles obtidos por medição, nota-se que as frequências de ressonância se encontram bem alinhadas, denotando que há um baixo desvio de frequência entre esses sinais.

Figura 39 – Comparação entre os sinais medido e simulado da magnitude e fase de S_{11} para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15).



Fonte: O autor, 2020.

As ligeiras variações na posição das frequências de ressonância entre os diversos sinais, bem como nas suas amplitudes são definidas pelas características específicas das falhas. Considerando toda a base de dados para a faixa avaliada é possível observar que os pontos de ressonância seguem um padrão definido, onde os valores exatos do conjunto de frequências de ressonância variam seguindo uma distribuição com média e desvio padrão mostrados na Tabela 11. Os baixos valores dos desvios padrão obtidos em relação às médias sugerem que os impactos gerados nas frequências de ressonância devido às diferentes configurações de falha são sutis, e que essas variações devem estar dispersas ao longo de todo o sinal.

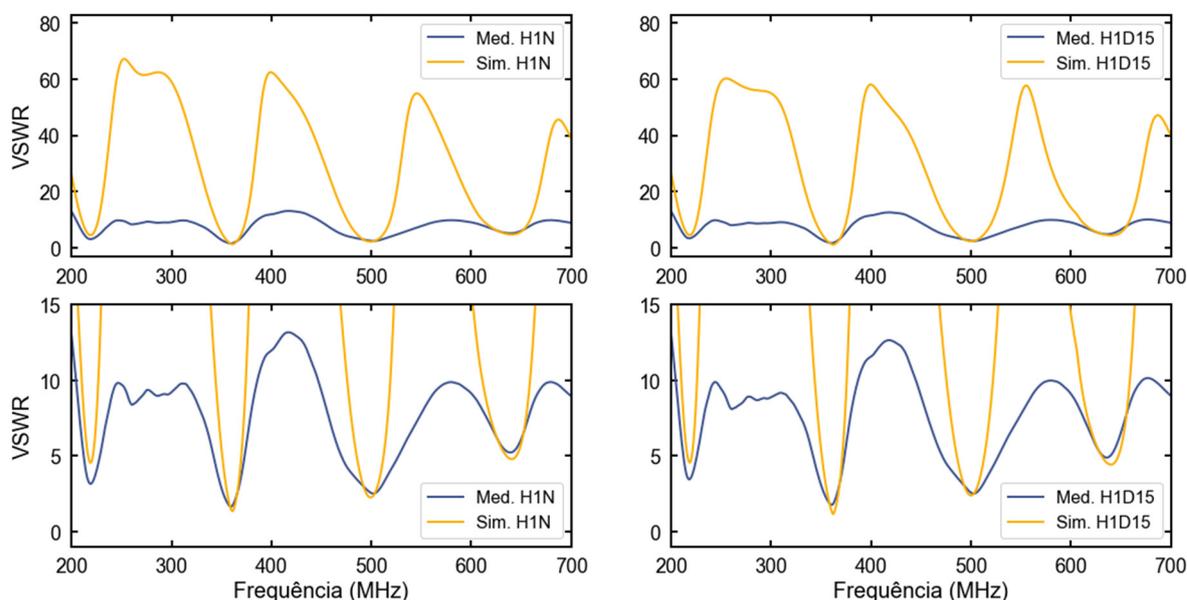
Tabela 11 – Parâmetros de distribuição das frequências de ressonância.

Parâmetro da Distribuição	Ress1	Ress2	Ress3	Ress4
Média da frequência de ressonância (MHz)	220,490	359,929	498,040	637,495
Desvio padrão da frequência de ressonância (MHz)	12,256	4,837	5,844	8,633
Desvio padrão da frequência de ressonância (%)	5,56	1,34	1,17	1,35

Fonte: O autor, 2020.

O parâmetro VSWR é mais sensível aos descasamentos e, por isso, consegue mostrar com maior evidência as diferenças entre os sinais medido e simulado fora da faixa de ressonância. Na Figura 40 são mostradas as respostas obtidas para o VSWR para as hastes H1N e H1D15, com uma visão geral dos sinais na parte superior da figura e uma visão detalhada dos comportamentos em torno dos pontos de ressonância na parte inferior.

Figura 40 – Comparação entre os sinais medido e simulado do VSWR para uma haste normal (H1N) e um exemplo de haste defeituosa (H1D15), com detalhe em torno dos pontos de ressonância.



Fonte: O autor, 2020.

O VSWR tende a 1 na frequência de ressonância principal (Ress2) localizada em torno de 360 MHz. As diferenças observadas entre os sinais medido e simulado nesses sinais são mais pronunciadas ao se distanciar das frequências de ressonância, aumentando gradativamente em função dessa distância. Essas diferenças devem-se a fatores como perdas, imperfeições na

estrutura, acoplamento não-ideal, variações do material e tolerâncias de fabricação nas hastes reais que não são capturadas pelo modelo computacional simulado.

A decomposição dos sinais em suas componentes principais foi realizada visando a redução da dimensionalidade dos dados de entrada da RNA por meio da retirada de informação redundante dos dados. Esse procedimento produz dois ganhos principais: a diminuição do esforço computacional para treinamento do algoritmo, permitindo a avaliação de uma maior quantidade de modelos em um menor intervalo de tempo e a redução da possibilidade de ocorrência de sobreajuste devido à redução de correlações e redundâncias nos sinais de entrada.

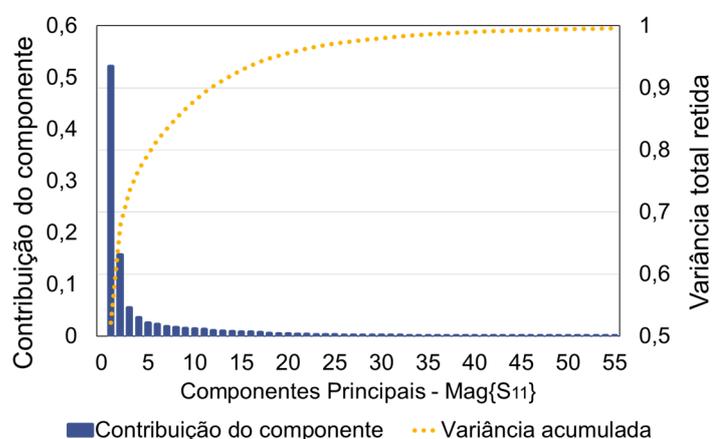
O número de componentes principais utilizado para cada parâmetro foi escolhido de modo a reter 99% da variância do sinal, conforme mostrado na Tabela 12. Essa estratégia possibilita a redução da quantidade de esforço computacional requerido para processamento dos dados, em troca de uma perda de informação tolerável. A distribuição dos componentes principais dos parâmetros magnitude de S_{11} , Fase de S_{11} , $VSWR$, parte real de Z_{in} e parte imaginária de Z_{in} após sua decomposição pelo procedimento de PCA são mostrados nos gráficos da Figura 41 até a Figura 45, nessa ordem.

Tabela 12 – Componentes principais requeridos de cada parâmetro de modo a reter 99% da variância.

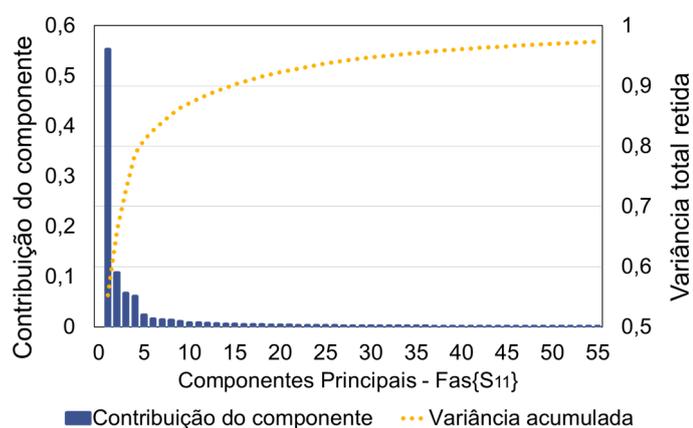
Parâmetro	$Mag\{S_{11}\}$	$Fas\{S_{11}\}$	$VSWR$	$Re\{Z_{in}\}$	$Im\{Z_{in}\}$
Número de Componentes Principais	40	102	44	42	36

Fonte: O autor, 2020.

Figura 41 – Distribuição dos componentes principais da magnitude de S_{11} .

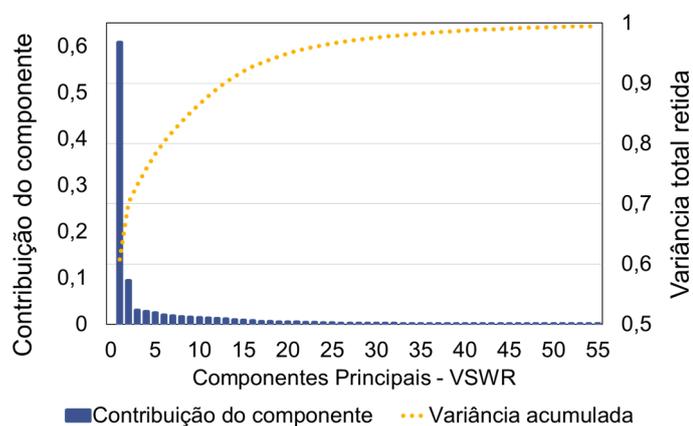


Fonte: O autor, 2020.

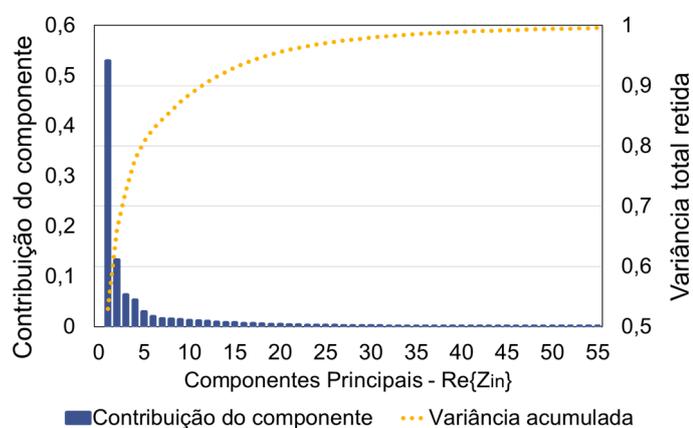
Figura 42 – Distribuição dos componentes principais da fase de S_{11} .

Fonte: O autor, 2020.

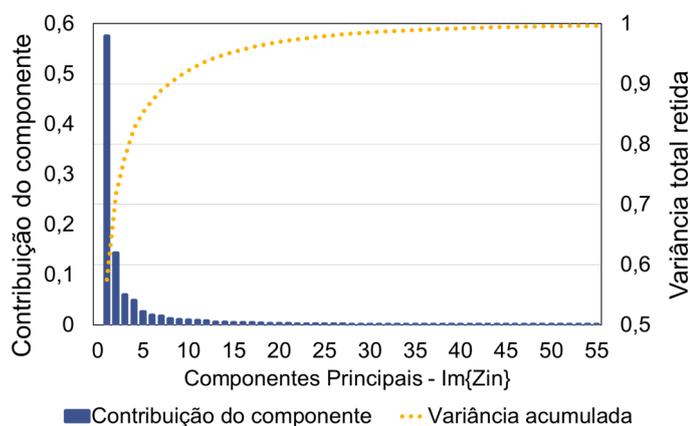
Figura 43 – Distribuição dos componentes principais do VSWR.



Fonte: O autor, 2020.

Figura 44 – Distribuição dos componentes principais da parte real de Z_{in} .

Fonte: O autor, 2020.

Figura 45 – Distribuição dos componentes principais da parte imaginária de Z_{in} .

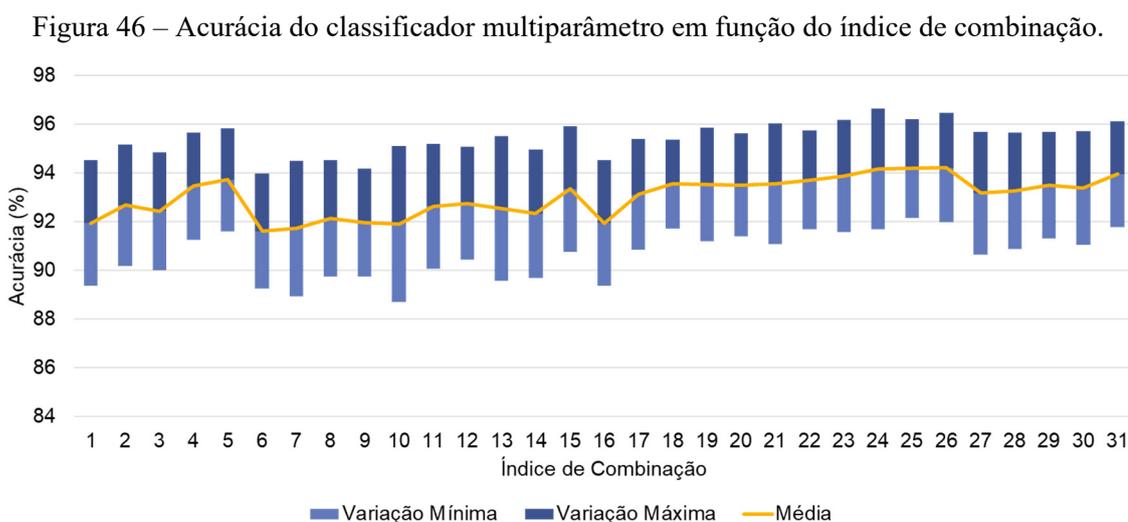
Fonte: O autor, 2020.

Considerando que, após o janelamento dos sinais entre 200 MHz e 700 MHz, cada um dos 05 vetores de entrada relacionados aos parâmetros originalmente obtidos possui dimensão 500, sem o uso do PCA as RNAs teriam que processar um total de 2500 pontos de entrada para a análise do estado de uma única haste. Com o uso do PCA, por outro lado, esse número se reduz a 264 pontos, referente à soma dos componentes principais de cada parâmetro individual mostrado na Tabela 12. Com exceção do sinal $Fas\{S_{11}\}$, que somente atingiu 99% da variância no componente de ordem 102, todos os demais parâmetros conseguiram atingir esse limiar com menos de 45 componentes. O comportamento diferenciado da fase de S_{11} em relação aos outros parâmetros é devido ao seu comportamento ser definido por uma função tangente, possuindo, portanto, menos linearidade e mais pontos de descontinuidade que os demais, o que requer uma maior quantidade de componentes para modelagem de toda sua variação.

Por meio da decomposição em componentes principais, foi selecionada uma quantidade de componentes tal que o somatório de suas contribuições retivesse 99% da variância contida nos dados. Com o uso do PCA, ao custo da perda de apenas 1% da variância contida nos dados é obtida uma diminuição de 89,44% na quantidade de pontos a serem processados, que são reduzidos de 2500 (500 pontos por parâmetro) para 264 (somatório das quantidades da Tabela 12). Os tempos de processamento requeridos pelo algoritmo antes e depois da redução de dimensionalidade com o PCA foram medidos por meio da captura dos tempos inicial e final nestes dois casos e calculando-se o intervalo decorrido em cada cenário. Assim, a diminuição no número de componentes se converte em uma redução no tempo médio para treinamento de uma RNA superior a 40%, passando de 284 s para 168 s, por iteração, permitindo a exploração de uma maior quantidade de arquiteturas distintas em um mesmo intervalo de tempo.

6.2 AVALIAÇÃO DO CLASSIFICADOR MULTIPARÂMETRO

Os desempenhos observados para cada classificador na tarefa de identificação da presença de falhas estruturais nas hastes são aferidos pela acurácia média obtida (quantidade de classificações corretas / quantidade total de amostras avaliadas) e exibidos na Figura 46, que mostra o desempenho médio do sistema para cada combinação de parâmetros dentro de um intervalo bilateral de 95% de confiança (MITCHELL, 1997).



Fonte: O autor, 2020.

Em função do parâmetro adotado, as médias da acurácia do sistema variaram de 91,95% a 93,73%, quando considerado cada parâmetro de forma isolada (índices de combinação de 1 a 5). Com a inserção do estágio combinador final, o sistema conseguiu obter acurácias ligeiramente superiores às dos parâmetros isolados, atingindo uma acurácia máxima de 94,22% para a combinação de índice 26, como apresentado na Tabela 13. Os resultados indicam que, por meio do processamento em um segundo estágio de classificação de análises previamente realizadas sobre os parâmetros de forma isolada, o sistema é capaz de compensar eventuais erros cometidos pelos classificadores individuais em diferentes partes do espaço de entradas, superando, dessa forma, o desempenho dos classificadores quando utilizados individualmente.

A vantagem de utilizar um modelo de aprendizado de máquina é que, apesar da similaridade entre os sinais, o algoritmo consegue automaticamente identificar e selecionar nos dados um conjunto de padrões e características relevantes para a distinção entre as respostas das hastes normal e defeituosa, enquanto as demais diferenças eventualmente observadas entre os sinais são tornadas irrelevantes no processo de classificação.

Tabela 13 – Análise comparativa do desempenho obtidos pelas combinações de parâmetros.

Índice	Parâmetros	Acurácia (%)	Índice	Parâmetros	Acurácia (%)
1	MagS11	91,95±2,57	17	MagS11 + FasS11 + ReZin	93,12±2,27
2	FasS11	92,68±2,49	18	MagS11 + FasS11 + ImZin	93,55±1,83
3	VSWR	92,43±2,42	19	MagS11 + VSWR + ReZin	93,53±2,34
4	ReZin	93,46±2,20	20	MagS11 + VSWR + ImZin	93,51±2,12
5	ImZin	93,73±2,12	21	MagS11 + ReZin + ImZin	93,56±2,47
6	MagS11 + FasS11	91,61±2,37	22	FasS11 + VSWR + ReZin	93,71±2,03
7	MagS11 + VSWR	91,72±2,78	23	FasS11 + VSWR + ImZin	93,87±2,3
8	MagS11 + ReZin	92,15±2,39	24	FasS11 + ReZin + ImZin	94,17±2,47
9	MagS11 + ImZin	91,96±2,21	25	VSWR + ReZin + ImZin	94,19±2,03
10	FasS11 + VSWR	91,91±3,20	26	MagS11 + FasS11 + VSWR + ReZin	94,22±2,24
11	FasS11 + ReZin	92,64±2,56	27	MagS11 + FasS11 + VSWR + ImZin	93,17±2,52
12	FasS11 + ImZin	92,76±2,32	28	MagS11 + FasS11 + ReZin + ImZin	93,27±2,4
13	VSWR + ReZin	92,55±2,97	29	MagS11 + VSWR + ReZin + ImZin	93,50±2,19
14	VSWR + ImZin	92,33±2,63	30	FasS11 + VSWR + ReZin + ImZin	93,38±2,34
15	ReZin + ImZin	93,35±2,58	31	MagS11 + FasS11 + VSWR + ReZin + ImZin	93,96±2,17
16	MagS11 + FasS11 + VSWR	91,95±2,57			

Fonte: O autor, 2020.

Os resultados obtidos indicam que é possível detectar falhas estruturais nas hastes de âncora a partir do processamento de parâmetros eletromagnéticos por classificadores que utilizam inteligência artificial com altos níveis de acurácia. Adicionalmente, o uso de múltiplos parâmetros eletromagnéticos de entrada proporciona às redes neurais uma diversidade maior de representações da estrutura da haste, melhorando ligeiramente o desempenho global do sistema de detecção e tornando-o menos sujeito a sobreajustes nos modelos.

6.3 ANÁLISE COMPARATIVA DOS ALGORITMOS DE CLASSIFICAÇÃO

A Tabela 14 exhibe os resultados comparativos para os algoritmos avaliados. Os modelos apresentaram erros médios relativamente baixos, entre 12,22% e 3,86% na tarefa de detecção da falha com um bom equilíbrio entre falsos positivos e falsos negativos, como mostram os valores de precisão, sensibilidade e f1score apresentados. Como a base de dados utilizada é apenas ligeiramente desbalanceada, os valores da acurácia, definida como $(1 - erro)$, e do f1score são bastante próximos, para todos os casos avaliados.

Tabela 14 – Análise comparativa do desempenho dos algoritmos de aprendizado de máquina.

Algoritmo	Erro (%)	Precisão (%)	Sensibilidade (%)	F1score (%)	Tempo de Treino* (s)
<i>Random Forests</i>	3,86±1,51	96,74±1,28	97,45±2,07	97,09±1,13	43,59±2,93
RNA	6,89±2,01	94,25±2,79	95,62±3,02	94,93±2,30	283,65±49,44
SVM	7,07±2,02	94,91±1,48	95,26±1,22	95,08±1,17	33,26±0,08
Regressão Logística	9,88±2,33	94,00±3,25	91,61±2,18	92,79±2,33	32,67±0,66
K-NN	12,22±2,56	92,05±3,62	88,69±3,87	90,34±2,94	35,15±0,08

* Configuração: Windows 10; 64 bits; Processador Intel I7 Dual Core 2,90 GHz; 6 MB de RAM.

Fonte: O autor, 2020.

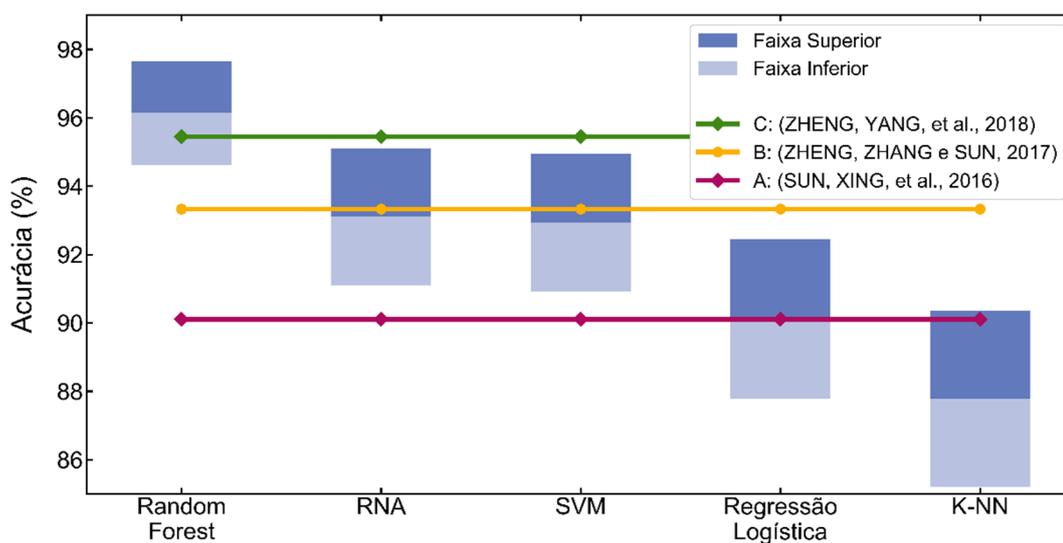
Os resultados apresentados na Tabela 14, foram obtidos utilizando os vetores de entrada originais sem redução de dimensionalidade, pois os algoritmos levaram um tempo aceitável para serem treinados com a base disponível. Esse tempo de treinamento, que variou, em função do algoritmo, da ordem de segundos a poucos minutos, não justificaria nenhuma perda adicional de informação nos dados para redução de dimensionalidade para a aplicação final. Por esse motivo, o PCA foi utilizado apenas na fase de busca de hiperparâmetros devido à quantidade de processamento requerida para treinar e testar uma grande quantidade de configurações de modelo distintas.

A Figura 47 mostra a comparação do desempenho dos algoritmos avaliados com linhas de referência A, B e C, que indicam resultados obtidos por duas metodologias distintas de detecção de falhas em hastes de âncora, reportadas em (SUN, XING, *et al.*, 2016), (ZHENG, ZHANG e SUN, 2017) e (ZHENG, YANG, *et al.*, 2018), respectivamente. Para permitir a comparação entre os métodos com um mesmo indicador, o desempenho dos algoritmos é dado em termos de sua acurácia, e exibida dentro de um intervalo de confiança bilateral de 95%.

A linha de referência A, marca o resultado do classificador apresentado em (SUN, XING, *et al.*, 2016), focado em avaliar a condição da haste em relação a concretagem do seu entorno. O método consistiu em extrair como parâmetros os valores máximos e mínimos não-nulos da distribuição da curtose espectral (ANTONI, 2006) da onda de stress que se propaga na haste após o impacto de um instrumento percussivo. Foram realizadas no *software* ANSYS-DYNA (ANSYS, 2020) 50 simulações de impactos na haste para cada um dos quatro modelos de âncora avaliados, sendo eles: ancoragem normal, defeituosa, excessiva e insuficiente, totalizando um conjunto de 200 sinais. Esses sinais foram agrupados em quatro conjuntos distintos, conforme sua classificação, utilizando uma técnica de agrupamento em dois estágios

com o algoritmo K-Means (MACQUEEN, 1967). Desprezando as amostras referentes aos casos de ancoragem excessiva e insuficiente e considerando apenas os resultados referentes a detecção de ancoragens normais e defeituosas, a acurácia média do método foi de 90,12%.

Figura 47 – Análise comparativa do desempenho do modelo proposto com referências encontradas na literatura em função do algoritmo de aprendizado de máquina utilizado.



Fonte: O autor, 2020.

A linha de referência B marca o resultado do classificador apresentado em (ZHENG, ZHANG e SUN, 2017), que similarmente ao sistema proposto nesta tese, é focado na detecção das falhas na estrutura metálica da haste de âncora em si. O método descrito consistiu em captar, por meio de um acelerômetro, 90 sinais vibracionais para cada uma das três hastes de âncora confeccionadas, sendo uma normal e duas defeituosas, após sofrerem um impacto de um instrumento percussivo. Um conjunto de 17 parâmetros físicos e estatísticos, tais como: valor médio, valor pico-a-pico da onda, desvio padrão e curtose, entre outros, foram extraídos dos sinais medidos nas hastes e processados por meio de PCA para extração de suas componentes principais. Esse conjunto de componentes foi então usado para treinar um classificador do tipo SVM estruturado como uma árvore binária de decisão, por meio de Otimização por Enxame de Partículas (PSO - *Particle Swarm Optimization*) (KHARE e RANGNEKAR, 2013). Considerando o uso de um número superior a 07 componentes principais como entrada do classificador, o método reportou uma acurácia média de 93,33%.

A linha de referência C, marca o resultado do classificador apresentado em (ZHENG, YANG, *et al.*, 2018). Utilizando um modelo híbrido denominado DS-DBM-SVM (*Differential*

Search - Deep Belief Network - Support Vector Machine) os autores propõem um classificador que é capaz de detectar defeitos nas hastes de âncora a partir dos sinais vibracionais coletados por um acelerômetro. Foram utilizadas nos experimentos três hastes de âncora, sendo uma normal e duas defeituosas. Para cada haste foram obtidas 100 leituras dos sinais de vibração, totalizando uma base de dados de 300 amostras com 5 s de duração cada, que foram posteriormente tratadas por análise Wavelet antes de serem aplicadas ao modelo de aprendizado de máquina. O método reportou uma acurácia média de 95,45% na tarefa de diferenciar as três hastes testadas a partir das leituras de vibração.

Pelos resultados apresentados pelo modelo proposto, a utilização uma onda eletromagnética de alta frequência ao invés de uma onda mecânica como sinal de entrada do sistema aumenta potencialmente a capacidade de diagnóstico das falhas em hastes de âncora. Comparando o desempenho do método de diagnóstico proposto com a linha de referência C, que contém o melhor resultado reportado para um sistema de detecção de falhas em hastes de âncora encontrado nas pesquisas do estado da arte (95,45%), se pode observar que pelo uso particular da solução proposta neste trabalho com o modelo *random forests* para a classificação dos dados, é obtido um desempenho superior (96,14%) na tarefa de detecção de falhas estruturais nas hastes. Por sua vez, os classificadores baseados em RNA e SVM apresentaram desempenhos estatisticamente semelhantes entre si, atingindo o mesmo nível de acurácia que o método representado pela linha de referência B. Finalmente, o classificador baseado em regressão logística apresentou desempenho comparável ao representado pela linha de base A, enquanto o que utilizou o algoritmo K-NN não foi capaz de atingir desempenho similar aos das linhas de referência. Vale ressaltar que em nenhuma das referências foi utilizada uma base de dados tão ampla e diversa quanto a gerada neste trabalho, nem feita nenhuma análise estatística da dispersão dos resultados em torno da média, nem estabelecidos intervalos de confiança para os resultados como o que foi criteriosamente realizado e apresentado aqui.

Em comparação às ondas mecânicas, as ondas eletromagnéticas de alta frequência possuem a desvantagem de oferecerem maior dificuldade para sua aplicação na estrutura da haste a ser medida. Tal dificuldade deve-se às restrições para a geração, acoplamento e casamento de impedância de uma onda eletromagnética de alta frequência e que devem ser superadas para que essa onda não seja completamente refletida no momento de sua aplicação. No entanto, ao se propagar em condutores metálicos, as ondas eletromagnéticas de alta frequência apresentam efeito pelicular decorrente da repulsão entre as linhas de corrente elétrica. O efeito pelicular faz com que as linhas de corrente elétrica tendam a se concentrar na

superfície do condutor, fluindo na camada mais externa, como uma película de espessura δ_s . Essa espessura, chamada de profundidade de penetração, é determinada por características do material e pela frequência angular ω da onda, de acordo com a equação (95), em que σ é a condutividade e μ a permeabilidade magnética do meio (POZAR, 2011). Devido ao efeito pelicular, os parâmetros de uma onda eletromagnética de alta frequência são significativamente mais afetados por alterações no material da superfície do que do centro do condutor. Essa característica favorece o uso desse tipo de onda como sensor para detecção de falhas estruturais que ocorram tipicamente no sentido da superfície para o centro do material, como, por exemplo, a corrosão galvânica que atinge as hastes de âncora em campo.

$$\delta_s = \sqrt{\frac{2}{\omega\mu\sigma}} \quad (95)$$

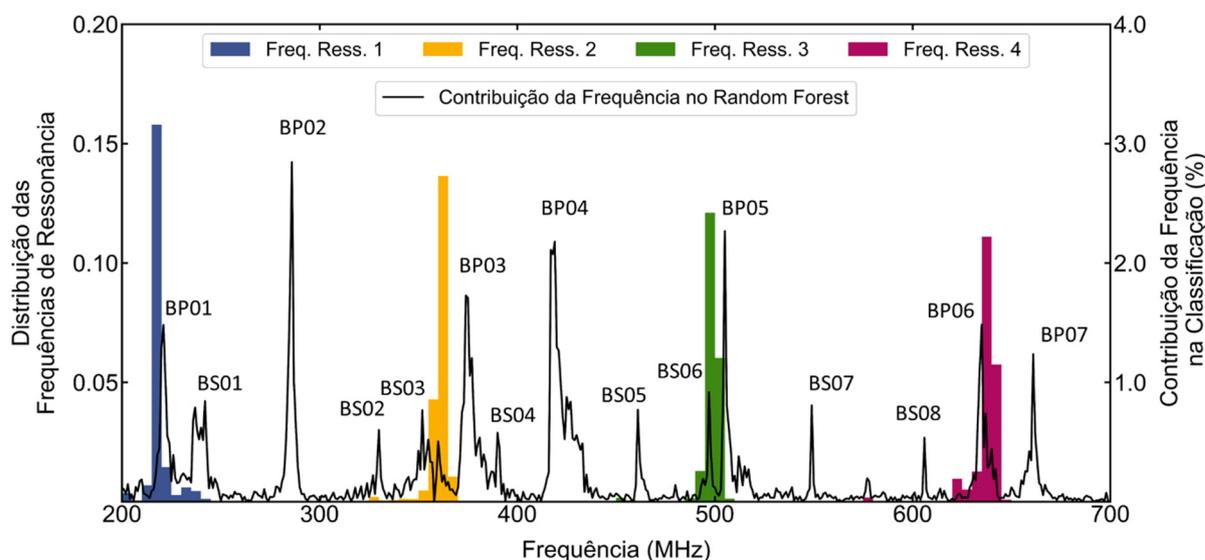
Em relação ao domínio em que os sinais são adquiridos, foi observado que trabalhar com as respostas no domínio da frequência dos sinais eletromagnéticos aplicados nas hastes apresenta vantagens em relação ao domínio do tempo. Os sinais no domínio da frequência são mais estáveis e menos ruidosos que os no domínio do tempo, portanto, requerem menos pré-processamento e, como consequência, podem ser mais facilmente utilizados para o treinamento dos algoritmos de aprendizado de máquina. Adicionalmente, a aquisição de dados no domínio da frequência é mais simples por não requerer a predefinição de nenhum parâmetro interno do analisador de redes vetorial. Devido à complexa geometria do arranjo hastes-conector, a correta representação dos sinais no domínio do tempo dependeria de estimar experimentalmente a permissividade relativa ϵ_r do meio em que o sistema está inserido. Conhecendo o valor da permissividade relativa é possível calcular a velocidade de propagação da onda na estrutura e, somente assim, se é capaz de calibrar adequadamente o analisador de redes para aquisição dos sinais no tempo, um processo complexo e susceptível a erros (ALVES, DE MELO, *et al.*, 2019).

Em relação aos modelos de aprendizado de máquina avaliados, pode-se dizer que o algoritmo *random forests*, conseguiu obter o melhor desempenho devido a sua estratégia de geração artificial de diversidade nos dados por meio da criação aleatória de uma grande quantidade de árvores de decisão. Essa grande quantidade de árvores, no entanto, pode causar uma excessiva lentidão no algoritmo. Contudo, apesar de aspectos relacionados à velocidade

do algoritmo poder serem críticos em aplicações de tempo real, eles não chegam a ser um problema para o diagnóstico de falhas em hastes de âncora.

A possibilidade de interpretação da análise realizada nos dados pelo algoritmo, se revelou uma das maiores vantagens do uso do *random forests* neste estudo. Como os parâmetros eletromagnéticos amostrados em frequência utilizados como vetores de entrada para a análise comparativa dos algoritmos não foram modificados pelo pré-processamento, por meio da análise dos critérios utilizados pelo *random forests* se é capaz de estimar quais componentes de frequência foram considerados mais relevantes para a separação dos dados. A Figura 48 exibe a importância relativa de cada componente de frequência na faixa de 200 MHz a 700 MHz utilizada pelo algoritmo *random forests* na classificação dos dados. Também são mostrados sobrepostos, em cores, para referência, os histogramas com as distribuições de ocorrência das frequências de ressonância.

Figura 48 – Importância relativa das componentes de frequência no processo de detecção das falhas estruturais nas hastes de âncora pelo algoritmo *random forests*.



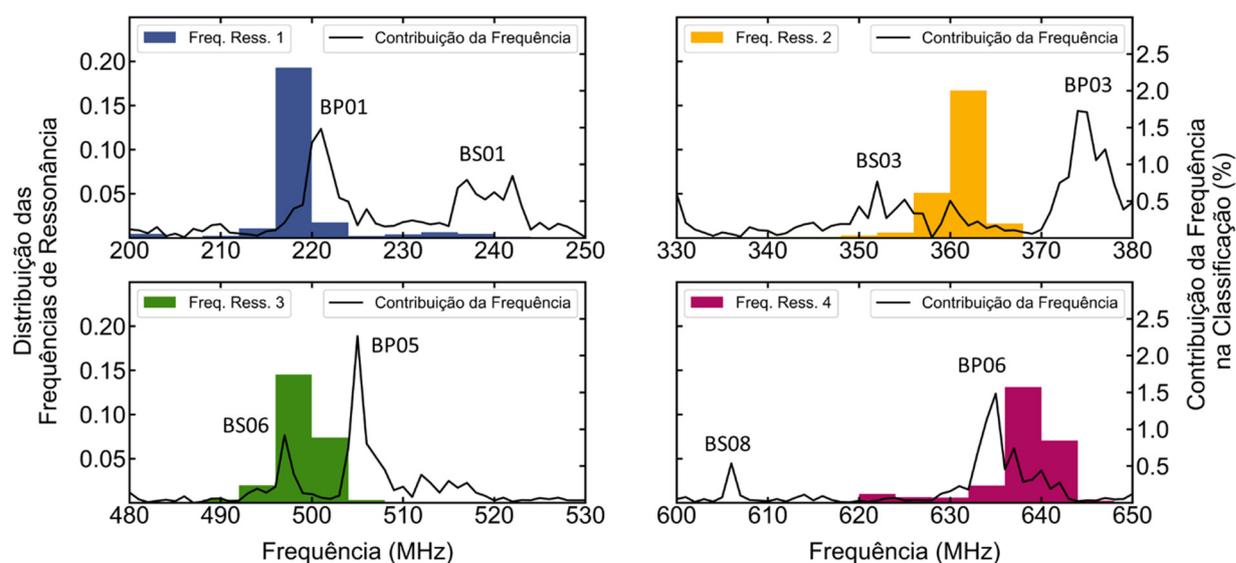
Fonte: O autor, 2020.

Pelos picos apresentados no gráfico, se pode observar que as componentes de frequência utilizadas para detecção das falhas não estão particularmente concentradas em nenhum trecho específico do espectro avaliado, mas ocorrem em bandas distribuídas por toda faixa de frequências, nomeadas no gráfico como bandas primárias (BP) e bandas secundárias (BS), conforme o nível de sua importância. Também é possível observar que dentre as 07 bandas primárias, cujos picos contribuem com mais de 1% de importância na classificação, 04 delas (BP02, BP03, BP04 e BP07) estão significativamente afastadas das frequências de ressonância.

As outras 03 bandas assinaladas como primárias (BP01, BP05 e BP06), embora estejam mais próximas dos pontos de ressonância, sugerem estar dedicadas a detectar variações nas frequências imediatamente adjacentes a esses pontos, e não nas frequências localizadas nos exatos picos de ressonância do sinal.

A Figura 49 exhibe em maiores detalhes as vizinhanças dos pontos de ressonância. Em particular, se pode observar que a mais proeminente frequência de ressonância do sinal, Ress2, localizada em aproximadamente 360 MHz e sinalizada na cor laranja no gráfico, não possui uma importância significativa no que se refere a distinção entre amostras normais e defeituosas de haste. Um comportamento similar se verifica para os demais pontos de ressonância, nos quais o algoritmo mostra creditar maior importância às frequências imediatamente adjacentes do que aos exatos pontos de ressonância em si.

Figura 49 – Detalhamento do grau de importância das bandas de frequência de classificação nas vizinhanças pontos de ressonância.

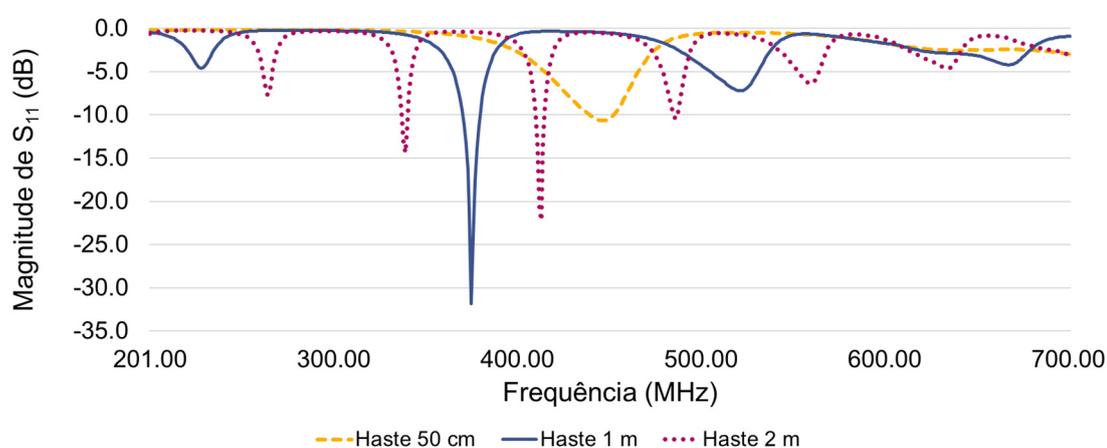


Fonte: O autor, 2020.

O comportamento ressaltado por meio da análise dos dados pelo *random forests* é coerente com os resultados observados nos experimentos e simulações. Em ambos os casos – hastes normais e hastes defeituosas – os sinais coletados possuem o mesmo comportamento geral na faixa de frequências analisada, apresentando quatro pontos de ressonância com pouca variação em sua localização em função da presença de falhas, como mostrado na Figura 37. Esses picos podem ser associados a harmônicos das frequências naturais de excitação da estrutura formada pelo conjunto conector-hastes, cujas quantidades e posições são mais

fortemente dependentes do comprimento da haste do que das características específicas da falha. A Figura 50 apresenta as respostas simuladas para hastes de 50,0 cm, 1,0 m e 2,0 m de comprimento na faixa de 200 MHz a 700 MHz. A escolha desses tamanhos de haste específicos, em que cada haste possui o dobro do comprimento da anterior, permite observar que a quantidade de picos segue um padrão harmônico. Esse padrão é função do comprimento da haste, podendo ser observado que para cada pico da haste de 50,0 cm, ocorrem dois para a haste de 1,0 m e três para a haste de 2,0 m de comprimento, sendo tal comportamento mais evidente na faixa de frequências mais baixas, onde as perdas por atenuação são menores.

Figura 50 – Resposta simulada da $\text{Mag}\{S_{11}\}$ para diferentes comprimentos de hastes sem falhas.



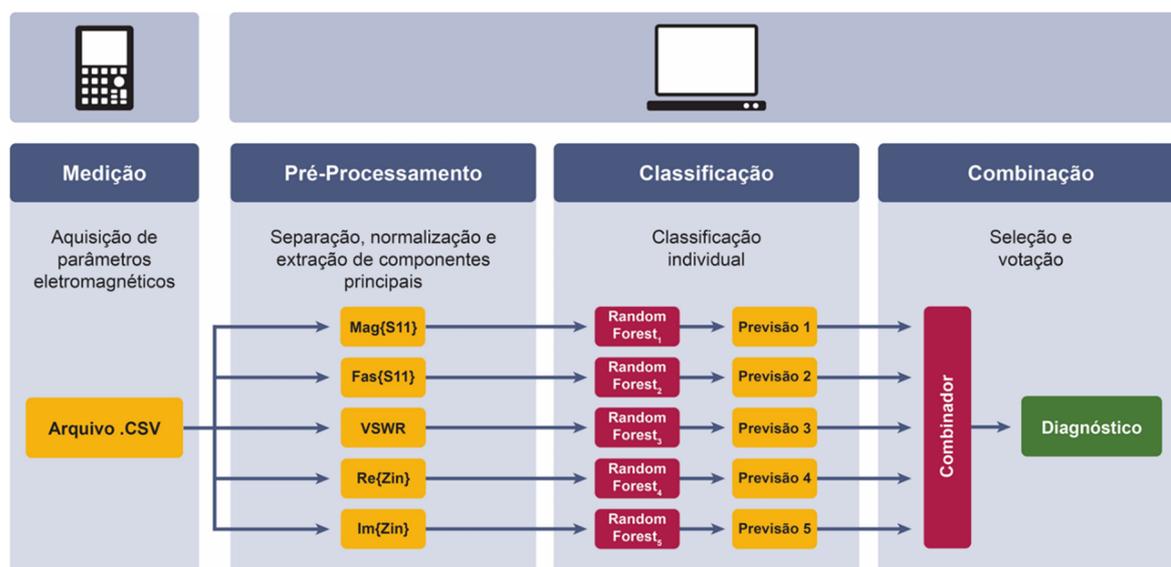
Fonte: O autor, 2020.

Por outro lado, se pode observar que o efeito das falhas é causar sutis distorções no sinal devido a existência de reflexões múltiplas, que são mais evidentes em bandas particulares do espectro, sem relação direta com as frequências de ressonância do conjunto. As análises realizadas indicam, portanto, que as falhas na estrutura da haste geram distúrbios que não afetam em grande escala a envoltória do sinal, mas causam pequenas variações na sua resposta ao longo de toda a faixa de frequências que dificilmente seriam identificadas por observação simples ou filtragem, mas que foram mapeadas com sucesso pela inteligência artificial. Dessa forma, o algoritmo de aprendizado de máquina atua de modo a selecionar e amplificar por meio de ponderação seletiva, aspectos dos sinais que sejam comuns entre amostras de uma mesma classe e diferentes entre elementos de classes distintas.

Uma sequência de novos testes foi realizada com a arquitetura de múltiplos parâmetros de entrada, desta vez utilizando o modelo treinado com o algoritmo *random forests* em

substituição às RNAs nos blocos classificadores, como mostrado na Figura 51. Foram realizadas duas baterias de testes adicionais, uma delas utilizando todas as componentes do vetor de entrada e a outra com as componentes principais após a redução de dimensionalidade.

Figura 51 – Esquema de avaliação da arquitetura de múltiplos parâmetros utilizando o modelo *random forests*.



Fonte: O autor, 2020.

Os resultados são mostrados na Tabela 15, adotando a mesma definição dos índices de combinação da tabela Tabela 7. Os desempenhos médios para os casos com e sem redução de dimensionalidade foram respectivamente de $93,09 \pm 1,56$ % e $95,96 \pm 0,46$ %. Desses resultados se pode observar que, com o *random forests*, há uma diferença significativa de performance entre utilizar todas as componentes do vetor de entrada original ou realizar a redução de dimensionalidade com o PCA. Essa diferença pode ser explicada pelo fato de o algoritmo *random forests* naturalmente construir as árvores de decisão que compõem a floresta aleatória utilizando subconjuntos com um número reduzido de parâmetros de entrada, na ordem de \sqrt{n} . Este fato associado à redução adicional de dimensionalidade feita pelo PCA pode acabar descartando informação relevante do sinal, degradando o desempenho do classificador.

A degradação observada quando aplicada a redução de dimensionalidade pelo PCA pode ser explicada pois, como mostrado na Figura 48 e Figura 49, os aspectos que permitem a detecção das falhas estruturais nas hastes não fazem parte do comportamento principal do sinal, mas estão contidos em variações secundárias. A informação contida nessas sutis variações

secundárias tende a ser subestimada pelo algoritmo PCA, que pressupõe que quanto mais significativa é a variação do sinal, mais importante ela será na modelagem do problema.

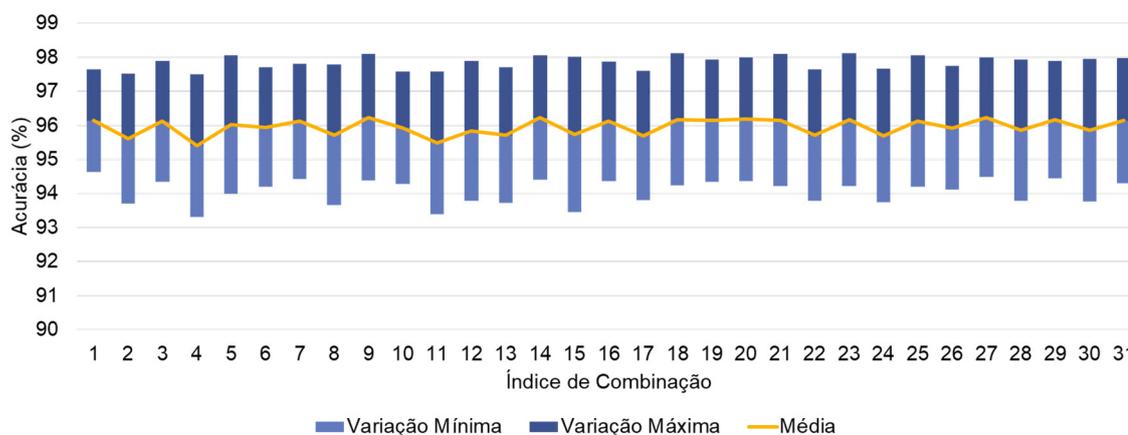
Tabela 15 – Desempenhos obtidos pelas diferentes combinações de parâmetros eletromagnéticos de entrada utilizando o modelo *random forests*.

Índice	Acurácia com dimensão reduzida (%)	Acurácia com sinal completo (%)	Índice	Acurácia com dimensão reduzida (%)	Acurácia com sinal completo (%)
1	91,80±2,57	96,14±1,51	17	93,55±1,83	95,70±1,90
2	92,68±2,49	95,62±1,91	18	93,53±2,34	96,17±1,94
3	92,43±2,42	96,12±1,78	19	93,51±2,12	96,14±1,79
4	93,46±2,20	95,41±2,10	20	93,56±2,47	96,18±1,81
5	93,73±2,12	96,02±2,04	21	93,71±2,03	96,15±1,94
6	91,61±2,37	95,95±1,76	22	93,87±2,30	95,71±1,93
7	91,72±2,78	96,12±1,70	23	94,17±2,47	96,16±1,95
8	92,15±2,39	95,72±2,06	24	94,19±2,03	95,70±1,96
9	91,96±2,21	96,24±1,85	25	94,22±2,24	96,12±1,93
10	91,91±3,20	95,93±1,65	26	93,17±2,52	95,93±1,81
11	92,64±2,56	95,49±2,10	27	93,27±2,40	96,24±1,75
12	92,76±2,32	95,84±2,06	28	93,50±2,19	95,86±2,08
13	92,55±2,97	95,71±1,99	29	93,38±2,34	96,17±1,72
14	92,33±2,63	96,23±1,82	30	93,96±2,17	95,86±2,09
15	93,35±2,58	95,74±2,28	31	93,94±2,10	96,14±1,84
16	93,12±2,27	96,12±1,76	Média	93,09±1,56	95,96±0,46

Fonte: O autor, 2020.

Por outro lado, como mostrado na Figura 52, não é percebida diferença significativa entre o desempenho de qualquer parâmetro particular ou de alguma combinação no que se refere a acurácia do diagnóstico. Os resultados sugerem que na tarefa de classificação das hastes a partir dos parâmetros eletromagnéticos medidos, quando utilizados vetores de entrada com todas as componentes do sinal original, a utilização do algoritmo *random forests*, consegue extrair dos dados quase a totalidade da informação disponível para a separação das classes. Desse modo, as pequenas variações de desempenho observadas entre os diferentes índices de combinação para esse caso, podem ser associadas a mera aleatoriedade na seleção dos dados de treinamento e teste no processo de avaliação dos modelos.

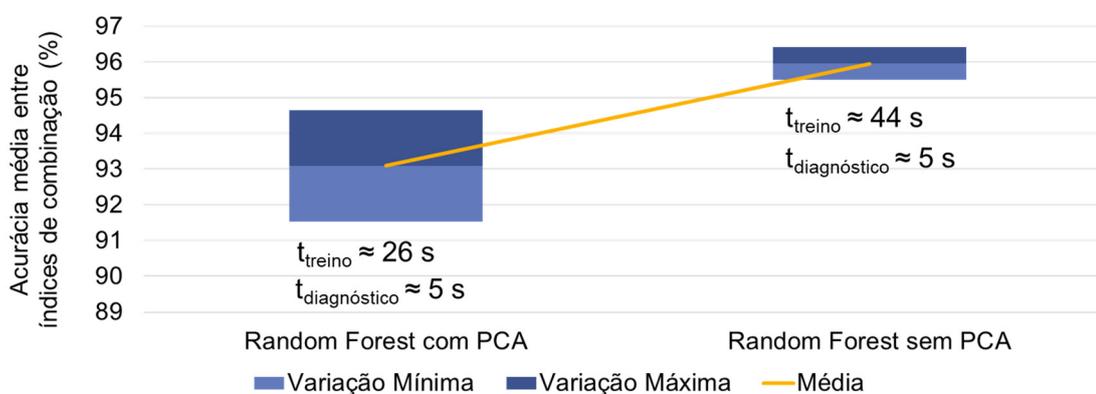
Figura 52 – Acurácia obtida pela arquitetura de múltiplos parâmetros utilizando o modelo *random forests*.



Fonte: O autor, 2020.

O tempo médio requerido para treinamento do modelo *random forests* nos testes se manteve em torno de 44 s, quando não utilizada redução de dimensionalidade por meio do PCA. Por outro lado, quando a redução de dimensionalidade com o PCA foi utilizada, esse tempo médio foi reduzido para aproximadamente 26 s, ao custo de uma diminuição na acurácia média de cerca de 3%, como apresentado na Figura 53. No entanto, em ambos os casos, uma vez que o modelo é treinado, o diagnóstico é obtido em menos de 5 s, o que, apesar de relativamente lento, é suficientemente satisfatório para a aplicação específica de detecção de falhas estruturais nas hastes de âncora.

Figura 53 – Efeito da redução de dimensionalidade na acurácia e tempo de processamento observados pelo modelo *random forests*.



Fonte: O autor, 2020.

7 CONCLUSÕES

Processos corrosivos podem gerar falhas estruturais ou até mesmo a completa ruptura das hastes que ancoram as torres estaiadas que são amplamente utilizadas nas linhas de transmissão em diversos países. Essas falhas podem levar, em casos extremos, à queda da torre e a interrupção do fornecimento de energia elétrica, causando danos financeiros e de imagem significativos às companhias envolvidas e impactos nas vidas das pessoas. Devido às graves consequências desses eventos, a avaliação constante da integridade das hastes de âncora das torres estaiadas é fundamental para garantir a confiabilidade e a segurança operacionais das linhas de transmissão de energia elétrica, evitando acidentes e seus altos custos associados.

Atualmente, o método convencional de detecção de falhas estruturais nas hastes de âncora é realizado por meio de um processo de escavação seguida de inspeção visual. Esse procedimento é lento, invasivo, destrutivo e de alto custo devido às dificuldades logísticas associadas para mobilização de recursos nos pontos onde as torres geralmente estão instaladas. Diante dessa questão, a realização da detecção de falhas estruturais de modo não-invasivo e não-destrutivo por meio da medição a partir da superfície dos parâmetros eletromagnéticos de um sinal de alta frequência aplicado na haste, se apresenta como uma alternativa atrativa e viável. No entanto, por sua complexidade, os sinais adquiridos por esse procedimento requerem algum tipo de processamento adicional para que as sutis características que indiquem a presença das falhas sejam devidamente mapeadas e extraídas dos dados.

Para solucionar tais problemas, esta tese propôs um sistema de diagnóstico utilizando inteligência artificial que recebe como entradas um conjunto de parâmetros eletromagnéticos medidos da haste com um analisador de redes vetorial e produz como saída o seu estado, classificando-a como normal ou defeituosa. Os resultados obtidos mostram que, é possível realizar a identificação de falhas estruturais nas hastes a partir de medições dos parâmetros eletromagnéticos feitas na superfície sem que haja a necessidade de escavação e destruição provisória da ancoragem da torre.

Diante da inerente dificuldade de realizar a medição e classificação de uma grande quantidade de hastes em campo, foi adotada a estratégia de construir uma extensa base de dados para treinamento dos algoritmos com modelos computacionais simulados. Adicionalmente foi confeccionado um conjunto de hastes físicas para medição em laboratório, com os objetivos de enriquecer a base de dados com parâmetros reais e validar os resultados das simulações. Os modelos desenvolvidos apresentaram alta fidelidade aos sinais medidos e permitiram a definição da faixa de frequências entre 200 MHz e 700 MHz como aquela que melhor atende

aos critérios de similaridade entre os sinais medidos e simulados e quantidade de informação nos dados para a detecção das falhas.

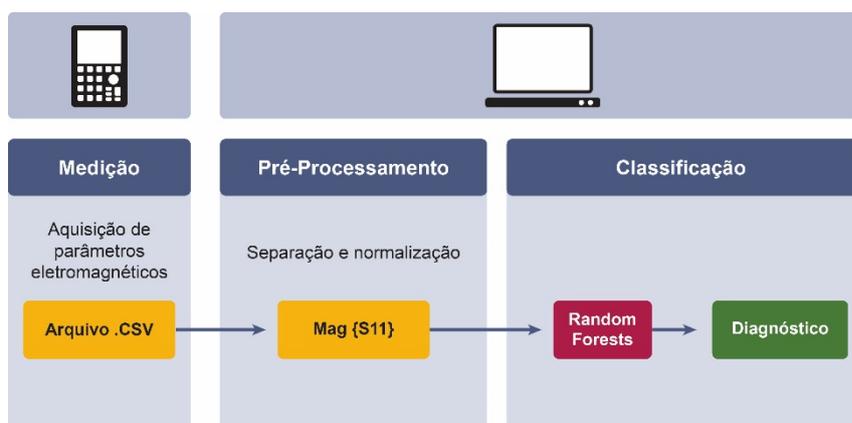
Utilizando classificadores baseados em redes neurais artificiais para a detecção das falhas foram obtidos resultados de acurácia média que variaram entre 91,95% para o parâmetro $\text{Mag}\{S_{11}\}$ a 93,73% para o parâmetro $\text{Im}\{Z_{in}\}$, quando observados isoladamente. Com a utilização do comitê de classificadores (*ensemble*), que permite a análise conjunta da combinação dos resultados individuais obtidos com os classificadores do estágio anterior, é atingida uma acurácia de 94,22% para a associação dos parâmetros $\text{Mag}\{S_{11}\} + \text{Fas}\{S_{11}\} + \text{VSWR} + \text{Re}\{Z_{in}\}$. Esses resultados indicam que, por meio da combinação dos resultados de classificação obtidos pelos parâmetros individuais, o sistema é capaz de compensar eventuais erros cometidos por esses classificadores em diferentes partes do espaço de entradas, obtendo desempenhos ligeiramente superiores.

A análise comparativa de algoritmos de aprendizado de máquina mostrou que o classificador proposto com o algoritmo *random forests* se destacou em termos de desempenho, superando métodos descritos na literatura que utilizam os algoritmos SVM e K-Means para processar parâmetros coletados de ondas mecânicas. A acurácia média do modelo baseado no algoritmo *random forests* proposto nesta tese foi de $95,96 \pm 0,46\%$, não variando significativamente em relação a qual combinação de parâmetros eletromagnéticos é utilizada como entrada, embora tenha se mostrado negativamente sensível à redução de dimensionalidade pelo PCA. No entanto, de acordo com os resultados obtidos, a redução de dimensionalidade para este modelo não é obrigatória e pode ser ignorada na etapa de pré-processamento do sistema, pois o tempo médio de análise dos dados requerido pelo o algoritmo, na ordem de segundos, é plenamente aceitável para a aplicação.

O classificador utilizando *random forests* com todas as componentes do vetor de entrada obteve desempenhos igualmente altos independentemente do parâmetro utilizado, superando significativamente os resultados de métodos anteriormente descritos na literatura. O nível de desempenho obtido sugere que é possível utilizar no módulo de aquisição de dados do sistema um analisador de redes portátil mais simples e de baixo custo, que opere na faixa de 200 MHz a 700 MHz e que, por exemplo, seja capaz de medir apenas a magnitude de S_{11} , aumentando a viabilidade da solução integrada final. Dessa forma, é possível utilizar um *hardware* de menor complexidade e custo no crítico módulo de medição, com a contrapartida de aumentar ligeiramente a carga de processamento do módulo de diagnóstico, o que é plenamente atendido pelos processadores comerciais atuais.

Diante dos resultados observados, recomenda-se para o diagnóstico das falhas estruturais em hastes de âncora, a arquitetura descrita na Figura 54, que é a que melhor atende aos critérios de desempenho, simplicidade e custo de implementação. Essa solução apresentou durante os testes um desempenho médio superior a 96%, ao mesmo tempo que exigiu menor quantidade de pré-processamento e utilizou como entrada o parâmetro mais facilmente encontrado em analisadores de rede de menor complexidade e custo.

Figura 54 – Arquitetura proposta para o sistema de diagnóstico.



Fonte: O autor, 2020.

Desconsiderando o valor do investimento para pesquisa e desenvolvimento da solução, o custo estimado de produção do kit do sistema de diagnóstico proposto, contemplando um analisador de rede vetorial de baixo custo, processamento embarcado em um computador arquitetura *Raspberry Pi*, confecção de um par de conectores para altas frequências, cabos, demais acessórios e uma maleta de transporte é estimado na ordem de US\$ 150.000,00, ou, aproximadamente, R\$ 750.000,00, com a premissa da cotação do dólar em torno de R\$ 5,00. Tomando o custo médio de cerca de R\$ 14.000,00 por âncora inspecionada reportado pelas operadoras de linhas de transmissão para execução da inspeção pelo método convencional por escavação da haste, o retorno financeiro com o uso do sistema de diagnóstico inteligente proposto se daria após cerca de 53 inspeções, ou seja, pouco mais de 13 torres, sem considerar eventuais ganhos pela não realização de escavações desnecessárias, o que torna o método bastante atrativo também do ponto de vista financeiro.

O sistema proposto possui duas principais limitações, que devem ser observadas como pontos de atenção para garantia de seu correto funcionamento. A primeira delas é que, como o sinal de entrada é uma onda eletromagnética de alta frequência sujeita a atenuação e reflexões

indesejadas devido ao descasamento de impedância do sistema de medição, a correta aquisição dos parâmetros eletromagnéticos de entrada pelo método proposto depende de um acoplamento adequado entre o analisador de redes vetorial e a haste. A segunda limitação é que, por utilizar um método estatístico para avaliação dos dados, deve ser inicialmente fornecido ao sistema, em seu estágio de treinamento, um significativo conjunto de amostras previamente classificadas dos sinais que serão analisados, que possua tamanho e distribuição de probabilidade compatíveis com aqueles que serão processados pelo sistema em seu modo de operação.

Observadas as limitações mencionadas, o sistema de diagnóstico proposto possui as vantagens de apresentar uma alta acurácia, ser portátil, não-destrutivo, não-invasivo e de resposta imediata. Além disso, é possível reduzir custos logísticos operacionais bem como minimizar a exposição de pessoas ao risco de acidentes durante a inspeção das hastes de âncora. Devido à sua baixa complexidade computacional, o sistema pode ser implementado e embarcado em *hardware* simples e, por operar com sinais no domínio da frequência, não são necessárias nenhuma predefinição ou parametrização prévia no analisador de redes vetorial utilizado para aquisição dos dados. O diagnóstico automático com saída digital disponibilizado pelo sistema possibilita sua utilização em aplicações de monitoramento remoto e uma integração com baixo esforço a aplicações de IoT. Finalmente, por ser uma aplicação de aprendizado de máquina, o sistema proposto pode ser realimentado com novos dados adquiridos de sua utilização em campo, melhorando continuamente sua capacidade de diagnóstico.

Os resultados obtidos indicam que a metodologia de diagnóstico proposta é confiável, apresenta um alto desempenho e permite o desenvolvimento de um sistema alternativo ao método convencional por escavação e inspeção visual para a detecção de falhas nas hastes de âncora das torres estaiadas das linhas de transmissão de energia elétrica. Utilizando o sistema proposto, o diagnóstico da haste pode ser realizado de forma automática e transparente para o operador, não necessitando, por parte desse usuário, de nenhum conhecimento formal prévio das tecnologias utilizadas para a sua utilização em campo.

As inovadoras possibilidades de aplicação avançada do sistema proposto são inúmeras. A partir da análise dos dados e tendências gerados pelo sistema por meio de técnicas de *Data Science and Analytics*, é possível estimar com precisão quais estruturas específicas apresentam degradação e quais os seus diferentes níveis de severidade, permitindo à equipe de engenharia planejar, priorizar intervenções e alocar recursos de forma eficiente. Essa característica impulsiona a evolução do paradigma da manutenção preventiva para o da manutenção preditiva, reduzindo-se o custo de operação das linhas de transmissão por meio da redução da quantidade

de intervenções desnecessárias e, ao mesmo tempo, aumentando-se a confiabilidade, disponibilidade e segurança do sistema elétrico.

7.1 TRABALHOS FUTUROS: EVOLUÇÃO DO MÉTODO

É sempre válido ressaltar que, como os algoritmos de aprendizado de máquina possuem uma grande flexibilidade de parametrização, é difícil estabelecer que a busca pelo melhor modelo para o problema abordado esteja concluída de forma inequívoca. Por esse motivo, quando um problema é abordado por estratégias de aprendizado de máquina, são tipicamente adotados modelos que, apesar de provavelmente apresentarem um desempenho subótimo, são soluções consideradas suficientemente satisfatórias para o negócio específico.

Assim, devido à grande variedade de soluções que são continuamente desenvolvidas na área de inteligência artificial, existe ainda margem para maiores investigações buscando uma configuração, estratégia ou arquitetura específica que aumente a capacidade de detecção das falhas nas hastes e que não tenha sido abordada nesta tese. Essas possibilidades são bastante promissoras, em especial no campo das RNAs, onde ainda podem ser exploradas variações na quantidade de camadas (*Deep Learning – DL*), outros tipos de funções de ativação e utilização de parâmetros concatenados como dados de entrada dos classificadores. Adicionalmente, podem ser avaliados diferentes esquemas a serem utilizados pelo comitê, tais como: combinação linear, votação majoritária simples ou votação ponderada, onde a ponderação dos parâmetros pode ser definida por algoritmos genéticos ou PSO (ESCOVEDO, KOSHIYAMA, *et al.*, 2020). Além disso, podem ser avaliadas outras arquiteturas tais como redes neurais convolucionais (*Convolutional Neural Networks – CNN*) (MASSA, MARCANTONIO, *et al.*, 2019), redes neurais recorrentes (*Recurrent Neural Networks – RNN*) (LIU, NA, *et al.*, 2016) e redes adversárias generativas (*Generative Adversarial Networks – GAN*) (CRESWELL, WHITE, *et al.*, 2018), que não foram avaliadas neste trabalho, mas têm apresentado resultados animadores em diversas aplicações.

Embora não tenha sido escopo desta tese propor modificações na configuração física dos conectores utilizados, os resultados obtidos pelo sistema de diagnóstico podem ser utilizados como referência para direcionar melhorias em componentes do sistema de medição, adequando conectores, cabos e demais equipamentos ao uso nas condições reais de campo. Estudos específicos voltados para a análise do desempenho dos conectores e avaliação de diferentes estratégias para utilização do arranjo de medição em campo já estão sendo conduzidos em linhas de transmissão operacionais, como apresentado na Figura 55.

Figura 55 – Medições de campo realizadas em Linha de Transmissão no município de Sobral – CE.



Fonte: O autor, 2020.

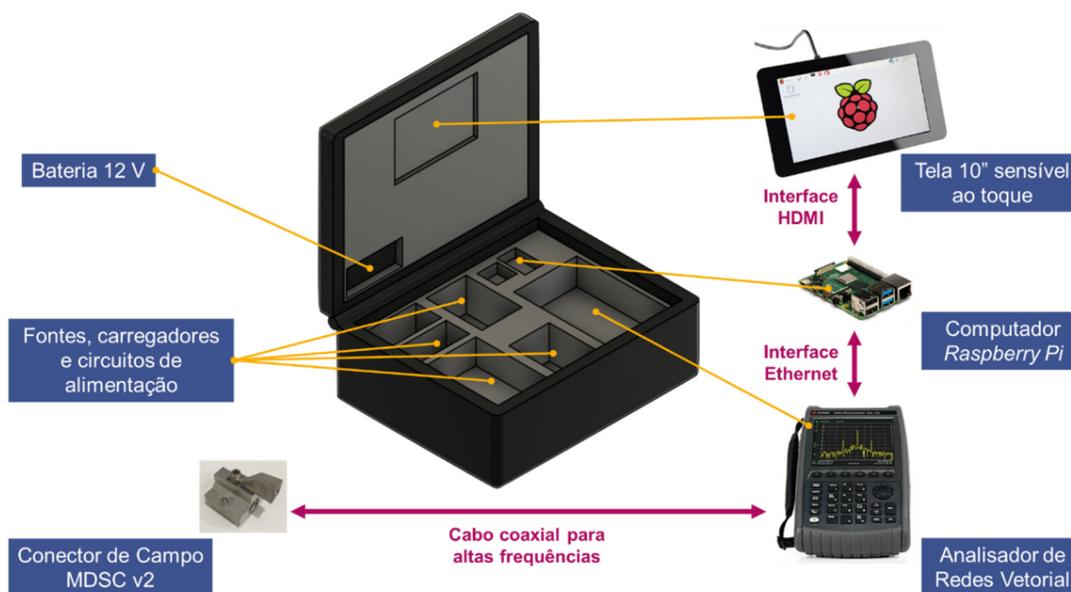
No que se refere aos classificadores, além da otimização dos modelos, trabalhos futuros incluem seu melhoramento por meio da adição de medições realizadas em campo e de novas simulações na base de dados, bem como pela busca de outros parâmetros de entrada que sejam capazes de aumentar a acurácia da detecção das falhas. Adaptações na metodologia apresentada podem ser realizadas para possibilitar o desenvolvimento de modelos de regressão e a investigação de formas de obtenção de informações quantitativas referentes ao nível de criticidade da falha, como localização, profundidade e/ou extensão.

Uma evolução adicional do sistema se encontra em fase de desenvolvimento para implementação futura e se refere à sua integração em uma plataforma embarcada, portátil e com uma interface amigável, para que a aquisição, processamento e análise dos dados ocorra de forma transparente ao usuário. Os *softwares* destinados ao comando remoto do ARV, o algoritmo de aprendizado de máquina treinado e a interface gráfica com o usuário que ocorrerá por meio de uma tela sensível ao toque (*touchscreen*), serão embarcados em um computador de placa única modelo *Raspberry Pi*.

No estágio atual, os dados das medições são inicialmente armazenados localmente na memória do ARV e posteriormente exportados para um computador em separado, o qual possui as aplicações desenvolvidas para análise de sinais e diagnóstico. Em um futuro próximo, com a conclusão do desenvolvimento e integração dos componentes do sistema embarcado, a realização do diagnóstico ocorrerá *in loco* pelo *Raspberry Pi*, dispensando a necessidade de importação dos dados para outro computador. Todos os equipamentos e componentes

eletrônicos serão acomodados em uma maleta compacta de transporte, permitindo o fácil manuseio do sistema de diagnóstico em campo, conforme protótipo apresentado na Figura 56.

Figura 56 – Protótipo conceitual do sistema embarcado de diagnóstico com a disposição futura dos equipamentos na maleta compacta de transporte (fora de escala).



Fonte: O autor, 2020.

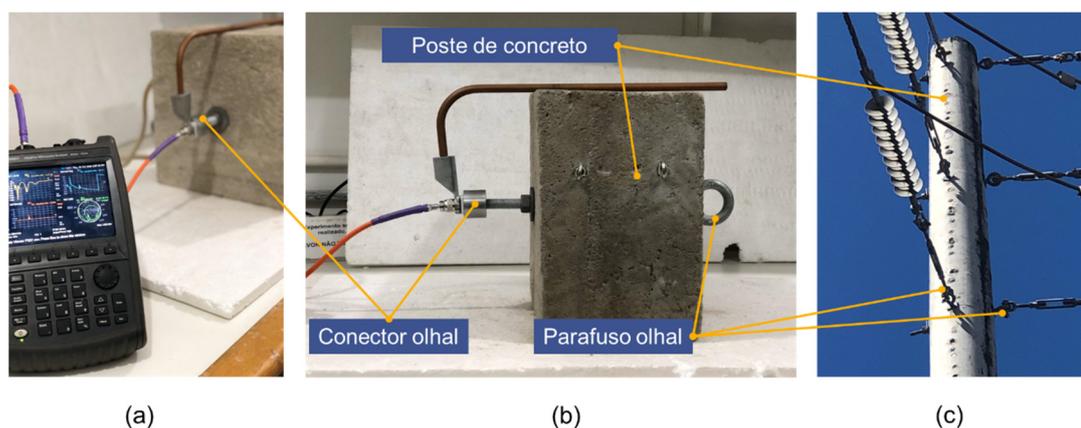
7.2 TRABALHOS FUTUROS: EXTENSÕES

Diante da grande aplicabilidade dos modelos de aprendizado de máquina nos mais distintos campos do eletromagnetismo, diversas utilizações dos conceitos trabalhados nessa tese em diferentes aplicações de engenharia surgiram das interações com outros pesquisadores do Laboratório de Micro-ondas da UFPE. Os estudos voltados para tais novas aplicações se encontram atualmente em andamento, tendo alguns deles já apresentado resultados.

O primeiro deles é uma extensão direta da metodologia proposta para detecção de falhas estruturais em parafusos olhais de fixação de postes de concreto. Tais parafusos sustentam as cadeias de isoladores e cabos e são elementos críticos dos sistemas de distribuição de energia elétrica. Por meio de um conjunto de adaptações físicas no conector, é possível se utilizar a metodologia descrita nesta tese para detectar falhas estruturais nesses parafusos, mesmo quando este defeito se encontra oculto no trecho interno da estrutura de concreto do poste, como apresentado na Figura 57. A técnica de diagnóstico pretendida possui a vantagem de poder ser realizada “a quente”, ou seja, sem a necessidade de desligamento da rede, aumentando a

eficiência, confiabilidade e custos de manutenção do sistema de distribuição de energia elétrica. Esta metodologia já apresentou seus primeiros resultados, que foram publicados em (SILVA FILHO, BARBOSA, *et al.*, 2020).

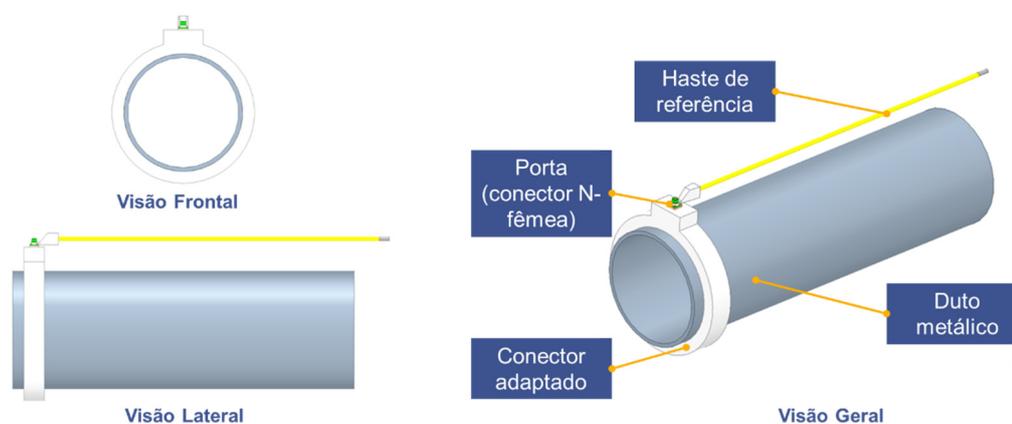
Figura 57 – Adaptação da metodologia proposta para detecção de falhas em parafusos olhais de postes de concreto. Medição em bancada (a) e (b). Parafusos olhais em serviço (c).



Fonte: Adaptado de Silva Filho, Barbosa *et al.*, 2020.

Uma segunda extensão direta do método pode ter sua viabilidade avaliada para a detecção de corrosão em dutos metálicos de transporte de óleo e gás, por meio de adaptações no conector que permitam o devido acoplamento mecânico e eletromagnético dos equipamentos de medição ao duto, como no exemplo mostrado na Figura 58. Essa aplicação se encontra com os estudos iniciados, em fase de desenvolvimento do projeto conceitual do conector.

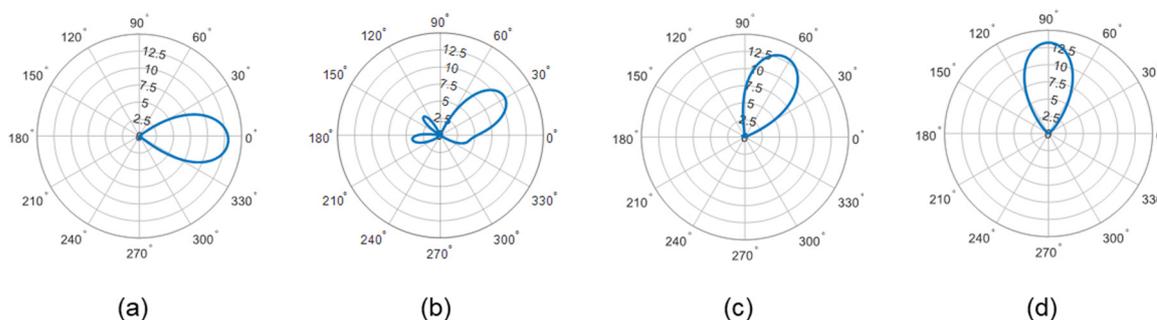
Figura 58 – Visões do modelo em desenvolvimento para aplicação do método proposto na detecção de corrosão em dutos de transporte de óleo e gás.



Fonte: O autor, 2020.

Outra utilização possível de técnicas de otimização de inteligência artificial em aplicações de eletromagnetismo diz respeito à determinação de parâmetros ótimos. Por meio de algoritmos como o PSO, por exemplo, é possível determinar automaticamente os valores ótimos dos coeficientes e fases requeridos para excitação de arranjos de antenas inteligentes de modo que o feixe resultante da interação entre as irradiações dos elementos individuais direcionem a máxima potência para um ângulo específico desejado, como apresentado na Figura 59.

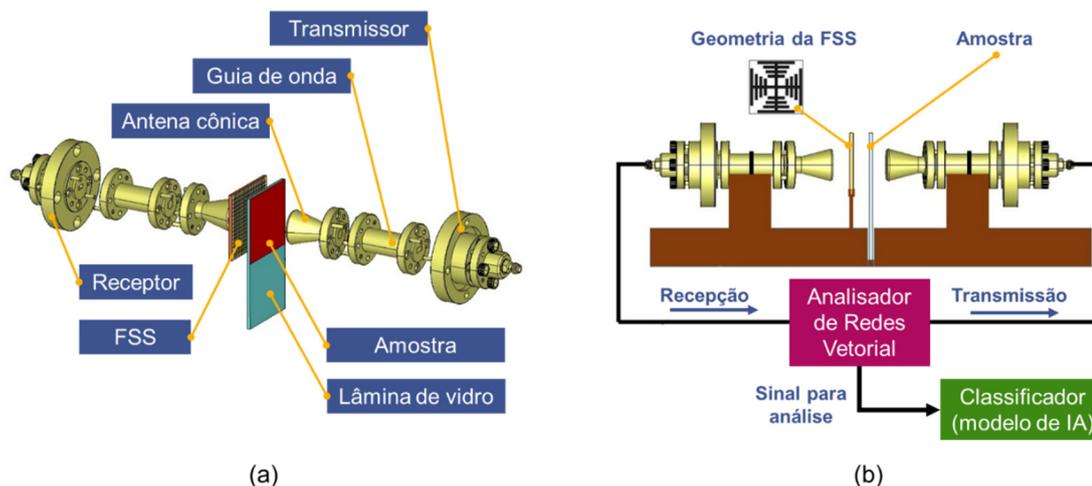
Figura 59 – Diagramas de radiação de arranjo de antenas Yagi-Uda obtidos com otimização por PSO. 0° – 12.98 dBi (a). 30° – 10.89 dBi (b). 60° – 12.98 dBi (c). 90° – 13.26 dBi (d).



Fonte: O autor, 2020.

Na área de biosensores, algoritmos de aprendizado de máquina podem ser treinados para o desenvolvimento de classificadores especializados em detectar substâncias ou patógenos em amostras biológicas por meio da extração de parâmetros diferenciadores, bem como da classificação automática das leituras do sensor por um modelo de inteligência artificial. O princípio de funcionamento desses biosensores se baseia na irradiação da amostra (sangue ou soro) com ondas eletromagnéticas em altas frequências e na premissa de que haja variação da capacidade de absorção dessas ondas pela amostra devido a alterações de suas propriedades eletromagnéticas, como condutividade elétrica, resistividade ou constante dielétrica, devido à presença de uma determinada substância ou da existência de um patógeno específico, seus anticorpos e / ou seus antígenos. As sutis diferenças causadas pelas variações entre os padrões positivos e negativos podem ser captadas por uma superfície seletora em frequência (*FSS – Frequency Selective Surface*). Os sinais captados pela FSS são analisados por um modelo de aprendizado de máquina para detectar com precisão a presença ou ausência de um determinado reagente na amostra biológica. Como exemplo, o protótipo apresentado na Figura 60 está sendo proposto para detecção rápida de amostras de vírus, como o SARS-COV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*) ou coronavírus, causador da pandemia COVID-19.

Figura 60 – Protótipo de sistema para detecção rápida de vírus por absorção eletromagnética. Visão explodida do sensor (a). Esquema básico de funcionamento (b).



Fonte: O autor, 2020.

7.3 CONTRIBUIÇÕES E PRODUÇÃO CIENTÍFICA

Como contribuições geradas pelos estudos realizados durante a realização deste doutorado, os resultados apresentados nesta tese produziram, diretamente ou por meio de colaborações desenvolvidas no Laboratório de Micro-ondas da UFPE, o conjunto de publicações científicas listado a seguir. Essa produção técnico-científica inclui um artigo publicado em periódico de alto fator de impacto e diversos trabalhos aceitos em importantes conferências internacionais.

7.3.1 Trabalhos publicados em periódicos

- a) BARBOSA, D. C. P.; DE MEDEIROS, L. H. A.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; COUTINHO, M. S.; ALVES, M. M.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L. Machine learning approach to detect faults in anchor rods of power transmission lines. **IEEE Antennas and Wireless Propagation Letters**, [S.I.], v. 18, n. 11, p. 2335-2339, 2019.

7.3.2 Trabalhos publicados em anais de conferências

- a) BARBOSA, D. C. P.; DE MEDEIROS, L. H. A.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; COUTINHO, M. S.; ALVES, M. M.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. An electromagnetic multi-parameter strategy to detect faults in anchor rods using neural networks. *In:*

SBMO/IEEE MTTs INTERNATIONAL MICROWAVE AND OPTOELECTRONICS CONFERENCE (IMOC), 18., 2019, Aveiro. **Anais [...]**. Portugal, 2019. p. 1-4.

- b) COUTINHO, M. S.; LOURENÇO NOVO, L. R. G. S.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P.; DE MEDEIROS, L. H. A.; DE MELO, M. T.; BARBOSA, D. C. P.; ALVES, M. M.; TARRAGÔ, V. L.; DOS SANTOS, R. G. M. A novel methodology to detect faults on anchor rods using reflectometry and machine learning. *In: SBMO/IEEE MTTs INTERNATIONAL MICROWAVE AND OPTOELECTRONICS CONFERENCE (IMOC), 18., 2019, Aveiro. Anais [...]*. Portugal, 2019. p. 1-4.
- c) ALVES, M. M.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; COUTINHO, M. S.; BARBOSA, D. C. P.; DE MEDEIROS, L. H. A.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. A non-destructive inspection of anchor rods based on frequency domain reflectometry. *In: SBMO/IEEE MTTs INTERNATIONAL MICROWAVE AND OPTOELECTRONICS CONFERENCE (IMOC), 18., 2019, Aveiro. Anais [...]*. Portugal, 2019. p. 1-4.
- d) BARBOSA, D. C. P.; DE MEDEIROS, L. H. A.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; COUTINHO, M. S.; ALVES, M. M.; MARTINS, T. G. S.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. Arquitetura de rede neural artificial para detecção de falhas estruturais em hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica. *In: ENCONTRO REGIONAL ÍBERO-AMERICANO DO CIGRÉ (ERAC), 18., 2019, Foz do Iguaçu. Anais [...]*. Brasil, 2019. p. 1-8.
- e) COUTINHO, M. S.; LOURENÇO NOVO, L. R. G. S.; DE MELO, M. T.; DE MEDEIROS, L. H. A.; BARBOSA, D. C. P.; ALVES, M. M.; MARTINS, T. G. S.; TARRAGÔ, V. L.; DOS SANTOS, R. G. M.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. Sistema diagnóstico da integridade de hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica. *In: ENCONTRO REGIONAL ÍBERO-AMERICANO DO CIGRÉ (ERAC), 18., 2019, Foz do Iguaçu. Anais [...]*. Brasil, 2019. p. 1-8.

7.3.3 Trabalhos aceitos em conferências aguardando publicação

- a) BARBOSA, D. C. P.; DE MEDEIROS, L. H. A.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; COUTINHO, M. S.; ALVES, M. M.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. Artificial neural network-based system for location of structural faults on anchor rods using input impedance response. *In: IEEE CONFERENCE ON ELECTROMAGNETIC FIELD COMPUTATION (CEFC 2020), 19.; 2020, Pisa, Itália. Data Prevista da realização da Conferência: 16 a 19/11/2020.*

- b) TARRAGÔ, V. L.; LOURENÇO NOVO, L. R. G. S.; DE MELO, M. T.; DE MEDEIROS, L. H. A.; COUTINHO, M. S.; BARBOSA, D. C. P.; ALVES, M. M.; DOS SANTOS, R. G. M.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. An equivalent model to optimize the electromagnetic simulation time for metallic rods used to anchor power transmission towers. *In: IEEE CONFERENCE ON ELECTROMAGNETIC FIELD COMPUTATION (CEFC 2020)*, 19.; 2020, Pisa, Itália. Data Prevista da realização da Conferência: 16 a 19/11/2020.
- c) ALVES, M. M.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; DE MEDEIROS, L. H. A.; COUTINHO, M. S.; BARBOSA, D. C. P.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; GAMA, P. H. R. P.; LOTT NETO, H. B. D. T. Optimization based on PSO and Ansys-HFSS to enhance reflectometry analysis on anchor rod fault localization. *In: IEEE CONFERENCE ON ELECTROMAGNETIC FIELD COMPUTATION (CEFC 2020)*, 19.; 2020, Pisa, Itália. Data Prevista da realização da Conferência: 16 a 19/11/2020.
- d) SILVA FILHO, H. V. H.; BARBOSA, D. C. P.; COUTINHO, M. S.; DE MELO, M. T.; DOS SANTOS, R. G. M.; LLAMAS-GARRO, I. Reliable structural failure detection in eye bolts using reflectometry signals. *In: EUROPEAN MICROWAVE WEEK (EUMW 2020)*, 2020, Utrecht, Holanda. Data Prevista da realização da Conferência: 10 a 15/01/2021.
- e) TARRAGÔ, V. L.; LOURENÇO NOVO, L. R. G. S.; DE MELO, M. T.; DE MEDEIROS, L. H. A.; COUTINHO, M. S.; BARBOSA, D. C. P.; ALVES, M. M.; DOS SANTOS, R. G. M.; PONTES, L. P.; GAMA, P. H. R. P.; LOTT NETO, H. B. D. T. Modelagem eletromagnética da ancoragem de torres transmissoras de energia elétrica. *In: SBMO – SIMPÓSIO BRASILEIRO DE MICRO-ONDAS E OPTOELETRÔNICA*, 19.; CBMAG – CONGRESSO BRASILEIRO DE ELETROMAGNETISMO (MOMAG 2020), 14., 2020, Niterói - RJ, Brasil. Data Prevista da realização da Conferência: 08 a 12/11/2020.
- f) ALVES, M. M.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; DE MEDEIROS, L. H. A.; COUTINHO, M. S.; BARBOSA, D. C. P.; DOS SANTOS, R. G. M.; TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. Otimização da técnica de reflectometria pelo algoritmo PSO na localização de falhas em haste. *In: SBMO – SIMPÓSIO BRASILEIRO DE MICRO-ONDAS E OPTOELETRÔNICA*, 19.; CBMAG – CONGRESSO BRASILEIRO DE ELETROMAGNETISMO (MOMAG 2020), 14., 2020, Niterói - RJ, Brasil. Data Prevista da realização da Conferência: 08 a 12/11/2020.
- g) ALVES, M. M.; DE MELO, M. T.; LOURENÇO NOVO, L. R. G. S.; DE MEDEIROS, L. H. A.; COUTINHO, M. S.; BARBOSA, D. C. P.; DOS SANTOS, R. G. M.;

TARRAGÔ, V. L.; LOTT NETO, H. B. D. T.; GAMA, P. H. R. P. A novel iterative method to estimate the soil complex permittivity from measurement and simulation modeling. *In: IEEE RADIO & WIRELESS WEEK (RWW 2021)*, 2021, San Diego - CA, USA. Data Prevista da realização da Conferência: 17 a 20/01/2021.

REFERÊNCIAS

- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Resolução Normativa nº 846**, Rio de Janeiro: ANEEL, 2019.
- AGGARWAL, C. C. **Neural Networks and Deep Learning**. [S.l.]: Springer, 2018.
- ALHAJRI, M. I.; ALI, N. T.; SHUBAIR, R. M. Classification of indoor environments for IoT applications: A machine learning approach. **IEEE Antennas and Wireless Propagation Letters**, v. 17, n. 12, p. 2164-2168, 2018.
- ALI, J. et al. Random forests and decision trees. **International Journal of Computer Science Issues (IJCSI)**, v. 9, n. 5, p. 272, 2012.
- ALVES, M. M. et al. **A non-destructive inspection of anchor rods based on frequency domain reflectometry**. 18th SBMO/IEEE MTTs International Microwave and Optoelectronics Conference (IMOC). Aveiro: [s.n.]. 2019.
- AMBASANA, N.; MUTNURY, B.; GOPE, D. **Intelligent rapid investigation of S-parameters (IRIS)**. IEEE 24th Electrical Performance of Electronic Packaging and Systems (EPEPS). San Jose: [s.n.]. 2015.
- ANSYS. ANSYS Workbench LS-DYNA, 2020. Disponível em: <<https://www.ansys.com/products/structures/ansys-ls-dyna/lstdyna-capabilities>>. Acesso em: 08 Julho 2020.
- ANTONI, J. The spectral kurtosis: a useful tool for characterising nonstationary signals. **Mechanical systems and signal processing**, v. 20, n. 2, p. 282-307, 2006.
- ASANUMA, J.; DOI, S.; IGARASHI, H. Transfer Learning Through Deep Learning: Application to Topology Optimization of Electric Motor. **IEEE Transactions on Magnetics**, v. 56, n. 3, p. 1-4, 2020.
- AWOYEMI, J. O.; ADETUNMBI, A. O.; OLUWADARE, S. A. **Credit card fraud detection using machine learning techniques: A comparative analysis**. 2017 International Conference on Computing Networking and Informatics (ICCNI). Ota: [s.n.]. 2017.
- AYALA, B. R.; CHEN, J. **A machine learning approach to evaluating translation quality**. ACM/IEEE Joint Conference on Digital Libraries (JCDL). Toronto: [s.n.]. 2017.
- AYESTARÁN, R. G. Fast near-field multifocusing of antenna arrays including element coupling using neural networks. **IEEE Antennas and Wireless Propagation Letters**, v. 17, n. 7, p. 1233-1237, 2018.
- BALDI, P.; HORNIK, K. Neural networks and principal component analysis: Learning from examples without local minima. **Neural Networks**, v. 2, n. 1, p. 53-58, 1989.
- BARMADA, S. et al. Deep Learning and Reduced Models for Fast Optimization in Electromagnetics. **IEEE Transactions on Magnetics**, v. 56, n. 3, p. 1-4, 2020.

BELLMAN, R. E. **An Introduction to Artificial Intelligence: Can Computers Think?** [S.l.]: Boyd & Fraser Publishing Company, 1978.

BENTES, J. L. **Análise dinâmica da ruptura de cabos em torres autoportantes e estaiadas de linhas de transmissão.** 2013. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul. Porto Alegre. 2013.

BERTENS, P. et al. **A machine-learning item recommendation system for video games.** 2018 IEEE Conference on Computational Intelligence and Games (CIG). Maastricht: [s.n.]. 2018.

BERTSEKAS, D. P. **Nonlinear Programming.** Belmont, MA: Athenas Scientific, 1995.

BEZERRA, J. M. et al. **Localization and Diagnosis of Stay Rod of V Guyed Towers Corrosion.** ICHVE International Conference on High Voltage Engineering and Application. Poznan, Poland: [s.n.]. 2014.

BISHOP, C. M. **Pattern Recognition and Machine Learning.** [S.l.]: Springer, 2006.

BLAUS, B., 2013. Disponível em:
<<https://commons.wikimedia.org/w/index.php?curid=28761830>>. Acesso em: 14 Agosto 2020.

BRAMER, M. **Priciples of Data Mining.** [S.l.]: Springer, 2016.

BREIMAN, L. Random decision forests. **Machine Learning**, Montréal, v. 45, n. 1, p. 5-32, 2001.

BREIMAN, L. et al. **Classification and Regression Trees.** Belmont: CRC Press, 1984.

BRESNAHAN, D. et al. Monitoring human head and neck-based motions from around-neck creeping wave propagations. **IEEE Antennas and Wireless Propagation Letters**, v. 17, n. 7, p. 1199-1203, 2017.

BURRASCANO, P.; FIORI, S.; MONGIARDO, M. A review of artificial neural networks applications in microwave computer-aided design (invited article). **International Journal of RF and Microwave Computer-Aided Engineering**, v. 9, n. 3, p. 158-174, 1999.

CAU, F. et al. A signal-processing tool for non-destructive testing of inaccessible pipes. **Engineering Applications of Artificial Intelligence**, v. 19, n. 7, p. 753-760, 2006.

CENDES, Z. **The development of HFSS.** USNC-URSI Radio Science Meeting. Las Croabas, Puerto Rico: [s.n.]. 2016.

CHARNIAK, E.; MCDERMOTT, D. **Introduction to Artificial Intelligence.** [S.l.]: Addison-Wesley, 1985.

CHEN, M. et al. Disease prediction by machine learning over big data from healthcare communities. **IEEE Access**, v. 5, p. 8869-8879, 2017.

CHEN, Y. Q.; FINK, O.; SANSAVINI, G. Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction. **IEEE Transactions on Industrial Electronics**, v. 65, n. 1, p. 561-569, 2018.

CHENG, D. K. **Field and Wave Electromagnetics**. [S.l.]: Pearson Education Asia Limited and Tsinghua University Press, 1989.

CHOU, J.-S.; NGO, N.-T.; CHONG, W. K. The use of artificial intelligence combiners for modeling steel pitting risk and corrosion rate. **Engineering Applications of Artificial Intelligence**, v. 65, p. 471-483, 2017.

COUTINHO, M. S. et al. **Sistema diagnóstico da integridade de hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica**. XVIII Encontro Regional Ibero-Americano do CIGRÉ (ERAC). Foz do Iguaçu, Brasil: [s.n.]. 2019.

CRESWELL, A. et al. Generative adversarial networks: An overview. **IEEE Signal Processing Magazine**, [D.I.], v. 35, n. 1, p. 53-65, 2018.

DE MASI, G. et al. **Machine Learning approach to corrosion assessment in subsea pipelines**. MTS/IEEE Oceans'15. Genova, Italy: [s.n.]. 2015.

DHANKHAD, S.; MOHAMMED, E.; FAR, B. **Supervised machine learning algorithms for credit card fraudulent transaction detection: a comparative study**. 2018 IEEE International Conference on Information Reuse and Integration (IRI). Salt Lake City: [s.n.]. 2018.

DOI, S.; SASAKI, H.; IGARASHI, H. Multi-objective topology optimization of rotating machines using deep learning. **IEEE transactions on magnetics**, v. 55, n. 6, p. 1-5, 2019.

DREISEITL, S.; OHNO-MACHADO, L. Logistic regression and artificial neural network classification models: a methodology review. **Journal of Biomedical Informatics**, v. 35, n. 6-6, p. 352-359, 2002.

DUMIN, O.; KHMARA, S.; SHYROKORAD, D. **Artificial neural networks in time domain electromagnetics**. XI International Conference on Antenna Theory and Techniques (ICATT). Kiev: [s.n.]. 2017.

DYWIDAG-SYSTEMS-INTERNATIONAL. Sistemas DYWIDAG para Linhas de Transmissão, São Paulo, 06 Fevereiro 2018. Disponível em: <<https://www.dywidag.com.br/downloads/catalogos-protendidos-dywidag/>>. Acesso em: 07 Julho 2020.

ESCOVEDO, T. et al. Neuroevolutionary learning in nonstationary environments. **Applied Intelligence**, [D.I.], v. 50, p. 1590-1608, 2020.

FAHAD, M. et al. **Corrosion Detection in Industrial Pipes Using Guided Acoustics and Radial Basis Function Neural Network**. International Conference on Robotics and Automation Sciences. Hong Kong: [s.n.]. 2017.

FAN, P.; ZHOU, R.; CHANG, Z. Novel neural network modeling method and applications. **International Journal of RF and Microwave Computer-Aided Engineering**, v. 25, n. 9, p. 769-779, 2015.

FENG, J. et al. Fast reconstruction of defect profiles from magnetic flux leakage measurements using a RBFNN based error adjustment methodology. **IET Science, Measurement & Technology**, v. 11, n. 3, p. 262-269, 2016.

FOUNDATION, R. The R Project for Statistical Computing, 2020. Disponível em: <<https://www.r-project.org/>>. Acesso em: 14 Agosto 2020.

FUJIYOSHI, H.; HIRAKAWA, T.; YAMASHITA, T. Deep learning-based image recognition for autonomous driving. **IATSS research**, v. 43, n. 4, p. 244-252, 2019.

GLOROT, X.; BORDES, A.; BENGIO, Y. **Deep sparse rectifier neural networks**. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS). Fort Lauderdale: [s.n.]. 2011.

GOLIK, P.; DOETSCH, P.; NEY, H. Cross-entropy vs. squared error training: a theoretical and experimental comparison. **Interspeech**, v. 13, p. 1756-1760, 2013.

GÓMEZ-PÉREZ, P. et al. Retrieving vegetation reradiation patterns by means of artificial neural networks. **IEEE Antennas and Wireless Propagation Letters**, v. 15, p. 1097-1100, 2015.

GONZÁLEZ, H. et al. **Determination of the Descriptors for the Design of a Classifier that Allows the Detection of Loss of Material in Metal Sheets Based on Signals of non-Destructive Tests**. International Conference on Artificial Intelligence and Big Data (ICAIBD). Chengdu, China: [s.n.]. 2018.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016.

GOSAL, G. et al. Transmitarray antenna design using forward and inverse neural network modeling. **IEEE Antennas and Wireless Propagation Letters**, v. 15, p. 1483-1486, 2015.

GOUTTE, C.; GAUSSIER, E. **A probabilistic interpretation of precision, recall and F-score, with implication for evaluation**. European Conference on Information Retrieval. Santiago de Compostela: [s.n.]. 2005.

GUILIZZONI, R.; FINCH, G.; HARMON, S. Subsurface corrosion detection in industrial steel structures. **IEEE Magnetics Letters**, v. 10, p. 1-5, 2019.

GUO-DONG, L. **Corrosion Evaluation Model of Reinforcement in Concrete Based on ANN**. 8th International Conference on Intelligent Computation Technology and Automation. Nanchang, China: [s.n.]. 2015.

HARIKESHA, R. et al. **ANN model for predicting the intergranular corrosion susceptibility of friction stir processed aluminium alloy AA5083**. 2nd International

Conference on Communication and Electronics Systems (ICCES 2017). Coimbatore, India: [s.n.]. 2017.

HARKOUSS, Y.; SHRAIM, H.; HUSSEIN, B. Direction of arrival estimation for smart antenna in multipath environment using convolutional neural network. **International Journal of RF and Microwave Computer-Aided Engineering**, v. 28, n. 6, p. e21282, 2018.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. [S.l.]: Springer, 2009.

HATHOUT, I. et al. **Digital Image Expert System for Corrosion Analysis of Steel Transmission Structures**. IEEE Power & Energy Society General Meeting. Chicago, USA: [s.n.]. 2017.

HAUGELAND, J. **Artificial Intelligence: The Very Idea**. [S.l.]: MIT Press, 1985.

HAYKIN, S. **Neural networks: A comprehensive foundation**. 2a. ed. Singapore: Pearson Prentice Hall, 1999.

HAYKIN, S. **Neural Networks and Learning Machines**. New York, NY, USA: Pearson Prentice Hall, 2009.

HE, K. et al. **Delving deep into rectifiers: Surpassing human-level performance on imagenet classification**. Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Washington: [s.n.]. 2015.

HEUMANN, C.; SHALABH, M. S. **Introduction to Statistics and Data Analysis**. [S.l.]: Springer, 2016.

HO, T. K. **Random decision forests**. HO, Tin Kam. RaProceedings of 3rd international conference on document analysis and recognition. Montréal: [s.n.]. 1995.

HOLMSTROM, M.; LIU, D.; VO, C. Machine learning applied to weather forecasting. **Stanford University**, p. 1-5, 2016.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p. 359-366, 1989.

HOY, M. B. Alexa, Siri, Cortana, and more: an introduction to voice assistants. **Medical reference services quarterly**, v. 37, n. 1, p. 81-88, 2018.

IBRAHIM, I. et al. Surrogate-Based Acoustic Noise Prediction of Electric Motors. **IEEE Transactions on Magnetics**, v. 56, n. 2, p. 1-4, 2020.

JAMES, G. et al. **An Introduction to Statistical Learning with applications in R**. [S.l.]: Springer, 2013.

JAMSHIDI, M. (. et al. A novel neural-based approach for design of microstrip filters. **AEU-International Journal of Electronics and Communications**, v. 110, p. 152847, 2019.

JAVA. JAVA, 2020. Disponível em: <https://www.java.com/pt_BR/>. Acesso em: 14 Agosto 2020.

JAZDZEWSKA, A. et al. **Complex Corrosion Monitoring System for Crude Distillation Unit in Form of Neural Network**. Global Smart Industry Conference (GloSIC). Chelyabinsk, Russia: [s.n.]. 2018.

JIMÉNEZ, L. A. E.; BENALCÁZAR, M. E.; SOTOMAYOR, N. **Gesture recognition and machine learning applied to sign language translation**. VII Latin American Congress on Biomedical Engineering CLAIB. Bucamaranga: [s.n.]. 2017.

KAMINSKI JR, J. **Incertezas de modelo na análise de torres metálicas treliçadas de linhas de transmissão**. 2007. Tese (Doutorado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul. Porto Alegre. 2007.

KAO, C. Y.; HUNG, S.-L. Detection of structural damage via free vibration responses generated by approximating artificial neural networks. **Computers & Structures**, v. 81, n. 28-29, p. 2631-2644, 2003.

KAPETANAKIS, T. N. et al. Applying neuro-fuzzy soft computing techniques to the circular loop antenna radiation problem. **IEEE Antennas and Wireless Propagation Letters**, v. 17, n. 9, p. 1673-1676, 2018.

KARLIK, B.; OLGAC, A. V. Performance analysis of various activation functions in generalized MLP architectures of neural networks. **International Journal of Artificial Intelligence and Expert Systems**, v. 1, n. 4, p. 111-122, 2011.

KARSOLIYA, S. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. **International Journal of Engineering Trends and Technology**, v. 3, n. 6, p. 714-717, 2012.

KERAS, 2016. Disponível em: <<https://keras.io/>>.

KHADOM, A. A.; KARIM, A. M. A.; FARHAN, S. N. **Mathematical and Artificial Neural Network Models for Corrosion of High Temperature – High Pressure Boiler Pipes in Presence of Oxygen Scavengers**. 1st International Scientific Conference of Engineering Sciences - 3rd Scientific Conference of Engineering Science (ISCES). Diyala, Iraq: [s.n.]. 2018.

KHAN, A.; , G. V.; LOWTHER, D. Deep learning for magnetic field estimation. **IEEE Transactions on Magnetics**, v. 55, n. 6, p. 1-4, 2019.

KHANAL, S. S. et al. A systematic review: machine learning based recommendation systems for e-learning. **Education and Information Technologies**, v. 25, p. 2635-2664, 2019.

KHARE, A.; RANGNEKAR, S. Review of Particle Swarm Optimization and applications in solar photovoltaic system. **Applied Soft Computing**, v. 13, n. 5, p. 2997-3006, 2013.

KNIGMA, D. P.; BA, J. L. **ADAM**: A method for stochastic optimization. International Conference on Learning Representations (ICLR). San Diego: [s.n.]. 2015.

KOCH, G. Cost of Corrosion. In: _____ **Trends in Oil and Gas Corrosion Research and Technologies**. Ohio, USA: Elsevier - Woodhead Publishing, 2017. p. 3-30.

KOCH, G. H. et al. **Corrosion costs and preventive strategies in the United States**. United States. 2002.

KOHAVI, R. **A study of cross-validation and bootstrap for accuracy estimation and model selection**. International Joint Conference on Artificial Intelligence. Montréal, QC - Canada: [s.n.]. 1995.

KUMARAVEL, R.; TAMILVANAN, K. Design of 132 kV transmission line tower with guy wire support and design comparison of self supporting tower. **International Journal of Innovative Research in Science, Engineering and Technology**, v. 5, n. 4, p. 5248-5253, 2016.

KURZWEIL, R. **The Age of Intelligent Machines**. [S.l.]: MIT Press, 1990.

LECUN, Y. A. et al. Efficient Backprop. In: _____ **Neural Networks: Tricks of the Trade**. Berlin, Germany: Springer, 2012. p. 9-48.

LEI, M. et al. An automatic extraction algorithm for measurement of installed rock bolt length based on stress wave reflection. **Measurement**, v. 122, p. 563-572, 2018.

LI, J.; LIANG, G. **Petrochemical Equipment Corrosion Prediction Based on BP Artificial Neural Network**. IEEE International Conference on Mechatronics and Automation (ICMA). Beijing, China: [s.n.]. 2015.

LI, M.; SOLTANOLKOTABI, M.; OYMAK, S. **Gradient descent with early stopping is provably robust to label noise for overparameterized neural networks**. International Conference on Artificial Intelligence and Statistics. Palermo: [s.n.]. 2020. p. 4313-4324.

LIN, L. **FNN-Based Reinforced Concrete Bridge Evaluation Model**. International Conference on Smart Grid and Electrical Automation (ICSGEA). Changsha, China: [s.n.]. 2017.

LIU, L. et al. **AI-facilitated Coating Corrosion Assessment System for Productivity Enhancement**. 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). Wuhan, China: [s.n.]. 2018.

LIU, W. et al. **A review of neural network based techniques for nonlinear microwave device modeling**. IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO). Beijing: [s.n.]. 2016. p. 1-2.

LIU, X.; QIN, L.; HUANG, B. Ultrasonic guided wave for monitoring corrosion of steel bar. **IOP Conference Series: Earth and Environmental Science**, v. 108, n. 2, p. 022071, 2018.

LIU, Y.-Y. et al. Structure damage diagnosis using neural network and feature fusion. **Engineering applications of artificial intelligence**, v. 24, n. 1, p. 87-92, 2011.

LO, C. M.; CHIU, Y. P.; LIN, M. D. **Predicting Atmospheric Corrosion Rates of Copper in Taiwan Industrial Zones Using Artificial Neural Network**. IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Singapore: [s.n.]. 2017.

LONGFEI, T.; ZHIHONG, X.; VENKATESH, B. Contactor modeling technology based on an artificial neural network. **IEEE Transactions on Magnetics**, v. 54, n. 2, p. 1-8, 2017.

LORENA, A. C.; DE CARVALHO, A. C. P. L. F. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43-67, 2007.

LOURENÇO NOVO, L. R. G. D. S. **Desenvolvimento de método para detecção de desgaste ou corrosão em hastes de âncora de torres estaiadas de linhas de transmissão de energia elétrica**. 2015. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Pernambuco. Recife. 2015.

LOURENÇO NOVO, L. R. G. S. et al. Design of connector for measurements in high frequency on anchor rods. **International Journal of Applied Electromagnetics and Mechanics**, v. 45, n. 1-4, p. 457–464, 2014.

LU, S. et al. An estimation method of defect size from MFL image using visual transformation convolutional neural network. **IEEE Transactions on Industrial Informatics**, v. 15, n. 1, p. 213-224, 2018.

LYASOTA, D.; MOROZOV, V.; MAGRO, V. **Application of neural networks to detect metal bodies**. IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO). Kiev: [s.n.]. 2015.

MACQUEEN, J. **Some methods for classification and analysis of multivariate observations**. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Berkeley: [s.n.]. 1967.

MARTIN, S.; CHOI, C. T. M. A new divide-and-conquer method for 3-d electrical impedance tomography. **IEEE Transactions on Magnetics**, v. 54, n. 3, p. 1-4, 2017.

MASSA, A. et al. DNNs as applied to electromagnetics, antennas, and propagation—A review. **IEEE Antennas and Wireless Propagation Letters**, [D.I.], v. 18, n. 11, p. 2225-2229, 2019.

MATLAB. Matlab, 2020. Disponível em: <<https://www.mathworks.com/products/matlab.html>>. Acesso em: 14 Agosto 2020.

MATOS, T. S. et al. Effect of chloride-induced corrosion on the mechanical properties of galvanized carbon steel anchor rods of guyed power transmission line. **Espaço Energia**, v. 28, p. 20-25, 2018.

MEDFOUNI, E. M. B. I. et al. Assessment of corroded rock bolts with pulse echo tests. **Journal of Infrastructure Systems**, v. 23, n. 4, p. 04017007, 2017.

MISHRA, S.; YADAV, R. N.; SINGH, R. P. Directivity estimations for short dipole antenna arrays using radial basis function neural networks. **IEEE Antennas and Wireless Propagation Letters**, v. 14, p. 1219-1222, 2015.

MITCHELL, T. M. **Machine Learning**. New York, NY, USA: McGraw-Hill, 1997.

MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of Machine Learning**. Cambridge, Massachusetts, USA: The MIT Press, 2012.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos de Aprendizado de Máquina. In: _____ **Sistemas Inteligentes - Fundamentos e Aplicações**. [S.l.]: Manole, 2003. p. 89-114.

MOORKENS, J. What to expect from Neural Machine Translation: a practical in-class translation evaluation exercise. **The Interpreter and Translator Trainer**, v. 12, n. 4, p. 375-387, 2018.

NILSSON, N. J. **Artificial Intelligence: A New Synthesis**. [S.l.]: Morgan Kaufmann, 1998.

NUMPY , 2020. Disponível em: <<https://numpy.org/>>. Acesso em: 14 Agosto 2020.

OCTAVE. GNU Octave: Scientific Programming Language, 2020. Disponível em: <<https://www.gnu.org/software/octave/>>. Acesso em: 14 Agosto 2020.

OLIPHANT, T. E. Python for Scientific Computing. **Computing in Science and Engineering**, 2007. 10-20.

OOMMEN, T. et al. An objective analysis of support vector machine based classification for remote sensing. **Mathematical geosciences**, v. 40, n. 4, p. 409-424, 2008.

PACHECO, M. P.; DANZIGER, F. A. B.; PEREIRA PINTO, C. Design of shallow foundations under tensile loading for transmission line towers: An overview. **Engineering Geology**, v. 101, n. 3-4, p. 226-235, 2008.

PAK, M.; KIM, S. **A review of deep learning in image recognition**. 4th international conference on computer applications and information processing technology (CAIPT). Kuta: [s.n.]. 2017.

PANOV, P.; DZEROSKI, S. **Combining bagging and random subspaces to create better ensembles**. International Symposium on Intelligent Data Analysis. Heidelberg: [s.n.]. 2007.

PEDREGOSA, F. et al. Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825-2830, 2011.

PEREIRA, J. M.; BASTO, M.; DA SILVA, A. F. The logistic lasso and ridge regression in predicting corporate failure. **Procedia Economics and Finance**, [D.I.], v. 39, p. 634-641, 2016.

POLYAKOV, E. V. et al. **Investigation and development of the intelligent voice assistant for the Internet of Things using machine learning**. Moscow Workshop on Electronic and Networking Technologies (MWENT). Moscou: [s.n.]. 2018.

- POOLE, D. Probabilistic Horn abduction and Bayesian networks. **Artificial intelligence**, v. 64, n. 1, p. 81-129, 1993.
- POZAR, D. M. **Microwave Engineering**. 4^a. ed. New York, NY, USA: John Wiley & Sons, 2011.
- PRABHAKAR, E. et al. Mining Better Advertisement Tool for Government Schemes Using Machine Learning. **International Journal of Psychosocial Rehabilitation**, v. 23, n. 4, p. 1122-1135, 2019.
- PYTHON, 2020. Disponível em: <<https://www.python.org/>>. Acesso em: 14 Agosto 2020.
- QIN, Y.; ZHU, J.-G. Deep Neural Network: Data Detection Channel for Hard Disk Drives by Learning. **IEEE Transactions on Magnetics**, v. 56, n. 2, p. 1-8, 2020.
- RAY, P.; MISHRA, D. P.; BUDUMURU, G. K. **Hybrid technique for underground cable to locate the fault**. International Conference on Information Technology. Fiji: [s.n.]. 2016.
- REGIS JR, O. et al. **LT de emergência de 500 kV com um condutor por fase: estudos, medições e constatações**. XIX Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE). Rio de Janeiro: [s.n.]. 2007.
- REN, K. et al. Bidding machine: Learning to bid for directly optimizing profits in display advertising. **IEEE Transactions on Knowledge and Data Engineering**, v. 30, n. 4, p. 645-659, 2017.
- REZAIE, A. et al. **Corrosion Risk Assessment at Anchor Shafts of Telecommunication Towers**. CORROSION 2019. Nashville, TN, USA: [s.n.]. 2019.
- RIBEIRO, P.; NETO, M.; CARDOSO, S. Strategy for determining a magnet position in a 2-D space using 1-D sensors. **IEEE Transactions on Magnetics**, v. 54, n. 11, p. 1-5, 2018.
- RICH, E.; KNIGHT, K. **Artificial Intelligence**. [S.l.]: McGraw-Hill, 1991.
- RUPPERT, D.; MATTERSON, D. S. **Statistics and Data Analysis for Financial Engineering**. [S.l.]: Springer, 2015.
- RUSSEL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. [S.l.]: Prentice Hall, 2010.
- SÁNCHEZ, J. S. et al. Analysis of new techniques to obtain quality training sets. **Pattern Recognition Letters**, v. 24, n. 7, p. 1015-1022, 2003.
- SANTOS, E. J. P.; SILVA, L. B. M. Calculation of scattering parameters in multiple-interface transmission-line transducers. **Measurement**, v. 47, p. 248-254, 2014.
- SASAKI, H.; IGARASHI, H. Topology optimization accelerated by deep learning. **IEEE Transactions on Magnetics**, v. 55, n. 6, p. 1-5, 2019.
- SCHER, S.; MESSORI, G. Predicting weather forecast uncertainty with machine learning. **Quarterly Journal of the Royal Meteorological Society**, v. 144, n. 717, p. 2830-2841, 2018.

SCHOLKOPF, B.; SMOLA, A. J. **Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.** [S.l.]: MIT Press, 2001.

SENS, M. A.; SEBRÃO, M. Z. **Suscetibilidade à corrosão em âncoras de estaiamento de torres de transmissão de energia elétrica.** INTERCORR 2012. Salvador: [s.n.]. 2012.

SHEELA, K. G.; DEEPA, S. N. Review on methods to fix number of hidden neurons in neural networks. **Mathematical Problems in Engineering**, v. 2013, p. 1-11, 2013.

SHI, X. et al. An Automated Machine Learning (AutoML) Method of Risk Prediction for Decision-Making of Autonomous Vehicles. **IEEE Transactions on Intelligent Transportation Systems**, p. 1-10, 2020.

SHOJI, M.; HIRATA, A. **Ultrasonic guided wave testing of anchor rods embedded in soil.** IEEE International Ultrasonics Symposium (IUS). Tours, France: [s.n.]. 2016.

SHOTTON, J. et al. **Real-time human pose recognition in parts from single depth images.** IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs: [s.n.]. 2011.

SILVA FILHO, H. V. H. et al. **Reliable structural failure detection in eye bolts using reflectometry signals.** European Microwave Week (EuMW 2020). Utrecht: [s.n.]. 2020.

SMOLA, A. J.; SCHÖLKOPF. A tutorial on support vector regression. **Statistics and Computing**, [D.I.], v. 14, p. 199–222, 2004.

SOLA, J.; SEVILLA, J. Importance of input data normalization for the application of neural networks to complex industrial problems. **IEEE Transactions in Nuclear Science**, v. 44, p. 1464-1468, 1997.

SOWAH, R. A. et al. **Design of power distribution network fault data collector for fault detection, location and classification using machine learning.** IEEE 7th International Conference on Adaptive Science & Technology (ICAST). Acra: [s.n.]. 2018.

SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, n. 1, p. 1929-1958, 2014.

SUN, X. et al. Pruning Elman neural network and its application in bolt defects classification. **International Journal of Machine Learning and Cybernetics**, v. 10, n. 7, p. 1847-1862, 2019.

SUN, X.-Y. et al. **Classification of anchor bolts based on spectral kurtosis and K-means clustering algorithm.** International Conference on Machine Learning and Cybernetics (ICMLC). Jeju Island, South Korea: [s.n.]. 2016.

SUPONENKOV, A. et al. **Application of image recognition and machine learning technologies for payment data processing review and challenges.** 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE). Riga: [s.n.]. 2017.

- TAHERI, P. et al. **Inspection and mitigation of underground corrosion at anchor shafts of telecommunication towers**. Proceedings of NACE INTERNATIONAL Corrosion Conference & Expo. New Orleans, USA: [s.n.]. 2017. p. 404-418.
- TAK, J. et al. A 3-d-printedw-band slotted waveguide array antenna optimized using machine learning. **IEEE Antennas and Wireless Propagation Letters**, v. 17, n. 11, p. 2008-2012, 2018.
- TAN, W. C. et al. **Learning with Corrosion Feature: For Automated Quantitative Risk Analysis of Corrosion Mechanism**. IEEE 14th International Conference on Automation Science and Engineering (CASE). Munich, Germany: [s.n.]. 2018.
- UNAL, F.; EKICI, S. **A Fault Location Technique for HVDC Transmission Lines using Extreme Learning Machines**. 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG). Istanbul: [s.n.]. 2017.
- URUCHURTU-CHAVARIN, J.; MALO-TAMAYO, J. M.; HERNÁNDEZ-PEREZ, J. A. Artificial intelligence for the assessment on the corrosion conditions diagnosis of transmission line tower foundations. **Recent Patents on Corrosion Science**, v. 2, n. 2, p. 98-111, 2012.
- VAPNIK, V. N. **The Nature of Statistical Learning Theory**. New York: Springer, 1995.
- VERMA, S. et al. **Corrosion current density modelling using artificial neural network**. 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). Noida, Uttar Pradesh, India: [s.n.]. 2018.
- WANG, B.; SANIIE, J. **Ultrasonic flaw detection based on temporal and spectral signals applied to neural network**. IEEE International Ultrasonics Symposium (IUS). Washington, D.C., USA: [s.n.]. 2017.
- WASSERMAN, P. D. **Neural Computing: Theory and Practice**. [S.l.]: Coriolis Group, 1989.
- WIENS, J.; SHENOY, E. S. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. **Clinical Infectious Diseases**, v. 66, n. 1, p. 149-153, 2018.
- WINSTON, P. H. **Artificial Intelligence**. [S.l.]: Addison-Wesley, 1992.
- WOOTTON, A. J.; DAY, C. R.; HAYCOCK, P. W. **Fault Detection in Steel-Reinforced Concrete Using Echo State Networks**. International Joint Conference on Neural Networks (IJCNN). Rio de Janeiro, Brazil: [s.n.]. 2018.
- WU, X. et al. Top 10 algorithms in data mining. **Knowledge and Information Systems**, v. 14, n. 1, p. 1-37, 2008.
- WU, X. et al. **Prediction of the Laws of Carbon Steel Erosion Corrosion in Sour Water System based on Decision Tree and Two kinds of Artificial Neural Network Model**. IEEE Chinese Control and Decision Conference (CCDC). Yinchuan, China: [s.n.]. 2016.
- YAO, H. M.; JIANG, L. Machine-Learning-Based PML for the FDTD Method. **IEEE Antennas and Wireless Propagation Letters**, v. 18, n. 1, p. 192-196, 2018.

YUEN, K. V.; LAM, H. F. On the complexity of artificial neural networks for smart structures monitoring. **Engineering Structures**, v. 28, n. 7, p. 977-984, 2006.

ZAMPIRON, I. **Avaliação das características e desempenho de estruturas para telecomunicações visando o projeto de novo modelo de torre autoportante**. 2008. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul. Porto Alegre. 2008.

ZHAI, S. et al. **Deepintent**: Learning attentions for online advertising with recurrent neural networks. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. San Francisco: [s.n.]. 2016.

ZHANG, B. et al. Detection and identification of object based on a magnetostrictive tactile sensing system. **IEEE Transactions on Magnetics**, v. 54, n. 11, p. 1-5, 2018.

ZHANG, W. et al. Detection of minor damage in structures with guided wave signals and nonlinear oscillator. **Measurement**, v. 122, p. 532-544, 2018.

ZHENG, H. Q. et al. **Nondestructive Detection of Anchorage Quality of Rock Bolt Based on DS-DBN-SVM**. International Conference on Machine Learning and Cybernetics (ICMLC). Adelaide: [s.n.]. 2018.

ZHENG, H.; ZHANG, S.; SUN, X. **Classification recognition of anchor rod based on PSO-SVM**. 29th Chinese Control And Decision Conference (CCDC). Chongqing, China: [s.n.]. 2017.

ZHENG, W. et al. Bio-Inspired Magnetostrictive Tactile Sensor for Surface Material Recognition. **IEEE Transactions on Magnetics**, v. 55, n. 7, p. 1-7, 2019.