UNIVERSIDADE DE SÃO PAULO–USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Narco Afonso Ravazzoli Maciejewski

**Detecção e Diagnóstico de Defeitos no
Regime Transitório de Motores de
Indução Baseado em Sistemas
Inteligentes**

Narco Afonso Ravazzoli Maciejewski

**Detecção e Diagnóstico de Defeitos no
Regime Transitório de Motores de
Indução Baseado em Sistemas
Inteligentes**

Tese de doutorado apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo como parte dos requisitos para a obtenção do título de Doutor em Ciências, Programa de Engenharia Elétrica.

Área de concentração: Sistemas Dinâmicos

Orientador: Rogério Andrade Flauzino

São Carlos

2021

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

M152d Maciejewski, Narco Afonso Ravazzoli
Detecção e Diagnóstico de Defeitos no Regime
Transitório de Motores de Indução Baseado em Sistemas
Inteligentes / Narco Afonso Ravazzoli Maciejewski;
orientador Rogério Andrade Flauzino. São Carlos, 2021.

Tese (Doutorado) - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em Sistemas
Dinâmicos -- Escola de Engenharia de São Carlos da
Universidade de São Paulo, 2021.

1. Defeitos em motores. 2. Sistema especialista. 3.
Inteligência artificial. 4. Reconhecimento de padrões.
5. Análise de séries temporais. 6. Análise de robustez.
I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **NARCO AFONSO RAVAZZOLI MACIEJEWSKI.**

Título da tese: "Detecção e diagnóstico de defeitos no regime transitório de motores de indução baseado em sistemas inteligentes".

Data da defesa: 16/04/2021.

Comissão Julgadora

Resultado

Prof. Associado **Rogério Andrade Flauzino**
(Orientador)

(Escola de Engenharia de São Carlos – EESC/USP)

Aprovado

Prof. Dr. **Renato Machado Monaro**

(Escola Politécnica/EP-USP)

Aprovado

Prof. Dr. **Geraldo Carvalho Brito Júnior**

(Universidade Estadual do Oeste do Paraná/UNIOESTE)

Aprovado

Prof. Dr. **Jefferson Tales Oliva**

(Universidade Tecnológica Federal do Paraná/UTFPR)

Aprovado

Prof. Dr. **Guilherme de Oliveira Kunz**

(Universidade Estadual do Oeste do Paraná/UNIOESTE)

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:

Prof. Dr. **João Bosco Augusto London Junior**

Presidente da Comissão de Pós-Graduação:

Prof. Titular **Murilo Araujo Romero**

“Só se pode alcançar um grande êxito quando nos mantemos fiéis a nós mesmos.”

(Friedrich Nietzsche)

Aos meus pais, Afonso e Cleunice, que muitas vezes se doaram e renunciaram aos seus sonhos, para que eu pudesse realizar os meus. E graças à união de todos, os obstáculos foram ultrapassados, vitórias foram conquistadas e alegrias compartilhadas. Agradeço pela paciência e compreensão com minha ausência durante essa longa jornada.

Muito obrigado!

Ao meu irmão, Afonso Júnior, pela amizade, carinho e companheirismo de sempre; por estar sempre torcendo pelas minhas conquistas. Pelo apoio e incentivo incondicional.

Obrigado!

À Universidade de São Paulo.

Agradecimentos

Desafio tão grande quanto escrever esta tese de doutorado, foi utilizar algumas poucas frases para agradecer as pessoas que fizeram parte desta minha trajetória.

Na minha primeira viagem a Foz do Iguaçu no ano de 2012, na procura de um apartamento, que conheci meu grande amigo, companheiro e colega de república Jefferson Tales Oliva, que além de nossos pais serem grandes amigos, ainda não o conhecia. Agradeço pela amizade, parceria e conversas que me abriram vários horizontes.

Agradeço aos meus amigos de infância, Lucas Fávero e Matheus Boaretto. Agradeço pela amizade, carinho e pelos momentos de descontração que até hoje são presentes.

Aos meus estimados amigos de minha cidade Nova Aurora, Jorge Corrêa, Kelven Venturini, Cleiton Venturini, Guilherme Oliveira, Flávia Borges, Andrey Armatiuk, Raul Pezenti, Lucas Fávero, Matheus Boaretto, Gustavo Neves, Higor Prado, Bruna Emília, Fabiana Marques e Jhonny Rosseti. Agradeço por toda amizade, carinho e por não permitirem a distância influenciar em nossa relação.

Aos meus grandes amigos dos anos acadêmicos, Jeison Picosso, Vinícius Bonini, Samuel Dianin, Mateus Martini, Arthur Sanches, Matheus Telles, Matheus Zanardini, Rodrigo Nepomuceno, Anderson Kobs, Tiago Giorgetti, Mariana Iuliano, Maike Henrique, Yago Zanatta, Jefferson Tales Oliva, Moacir Fontequê Júnior, Jediel Levi Ribas Teixeira e Thomás Gnoatto. Agradeço por todo o tempo passado juntos, pelo carinho e pela amizade nos momentos de estudo e descontração.

Aos meus colegas e amigos de trabalho, Katiani Pereira, Claudia Bozz, Fernando Gaiotto, Ricardo Menon, Welisson Vogado e Breno Carneiro. Agradeço por todos os conselhos e por me fazerem parte da equipe.

A professora Aline Elly Trembl, que conheci na graduação na disciplina de processos de fabricação e só após três anos voltaríamos a nos falar com o propósito da continuidade da pesquisa deste trabalho. Agradeço por toda a disponibilidade, apoio, e, acima de tudo, confiança. Palavras não são suficiente para agradecer.

Ao professor Rogério Andrade Flauzino por me acolher no doutorado, por todas as críticas construtivas, paciência, por todas as oportunidades compartilhadas e por fazer parte de minha vida. Foram mais de 30000km percorridos de Foz do Iguaçu para São

Carlos, e valeu a pena cada minuto. Só tenho a agradecer, obrigado.

Um agradecimento especial para minha namorada Renata Pessi. Agradeço pela amizade, carinho, compreensão e por todo o tempo desprendido a me ajudar. Obrigado também pelo companheirismo acima de tudo.

Ao meu grande irmão, Afonso Maciejewski Júnior, pelo companheirismo, amizade e discussões, principalmente nestes últimos anos. Sinto não estar presente, mas minha amizade, carinho e conselhos nunca mudaram.

Meu agradecimento mais profundo só poderia ser dedicado a duas pessoas: meus pais, Afonso Maciejewski e Cleunice Ravazzoli. Todos esses 26 anos ao meu lado, incondicionalmente. Nos momentos mais difíceis, que não foram raros neste último ano, sempre me fazendo acreditar que chegaria ao final nesta tão importante e gratificante etapa. Este período me mostrou o que significa um relacionamento em família. Sou grato por cada gesto carinhoso, cada conselho, cada conversa e pelo imenso amor, dedicação, educação e apoio incondicional em cada escolha da minha vida. Minha mente e coração estão sempre com vocês.

A todos que direta ou indiretamente contribuíram com o desenvolvimento deste trabalho.

“Como o tecido do universo é o mais perfeito e fruto do trabalho do mais sábio criador, nada acontece no universo sem que alguma lei de máximo e mínimo apareça.”
(Leonhard Paul Euler)

Resumo

Maciejewski, Narco Afonso Ravazzoli **Detecção e Diagnóstico de Defeitos no Regime Transitório de Motores de Indução Baseado em Sistemas Inteligentes.** 151 p. Tese de doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2021.

Os motores de indução trifásicos são os principais elementos de conversão de energia elétrica em mecânica para o setor produtivo, visto que apresentam baixo custo e longa durabilidade. Além disso, a redução de custos de manutenção e a implantação de técnicas preditivas torna-se um incentivo para o desenvolvimento de sistemas capazes de identificar defeitos intrínsecos à máquina. Frente a isso, essa pesquisa visa constituir uma metodologia robusta baseada em sistemas inteligentes para o diagnóstico e a detecção de defeitos de barras quebradas em motores de indução trifásicos ainda no regime transitório de operação. Para isso, inicialmente foi realizada uma busca sistematizada da literatura com o objetivo de alcançar o estado da arte da temática de detecção e de diagnóstico de defeitos em máquinas elétricas rotativas. O sistema inteligente proposto foi validado e testado com dados experimentais obtidos em condições normais e defeituosas do motor. Dessa maneira, foi proposta uma configuração baseada inteligência computacional para a identificação e supervisão de defeitos nesses elementos, e essas arquiteturas foram avaliadas e comparadas por meios de uma medida multicritério, inédita para essa finalidade, para atestar a relevância da metodologia proposta em relação as já existentes. Os resultados mostraram uma metodologia ágil e robusta para a recomendação de sistemas de detecção e de diagnóstico de defeitos em motores de indução com base nas características dos algoritmos de seleção de atributos e de aprendizado de máquinas.

Palavras-chave: Defeitos em motores. Sistema especialista. Inteligência artificial. Reconhecimento de padrões. Análise de séries temporais. Análise de robustez.

Abstract

Maciejewski, Narco Afonso Ravazzoli **Defect Detection and Diagnosis in the Transient Regime of Induction Motors Based on Intelligent Systems.** 151 p. Ph.D. Thesis – São Carlos School of Engineering, University of São Paulo, 2021.

Three-phase induction motors are the main elements for converting electrical energy into mechanics for the productive sector since they have low cost and long durability. Also, the reduction of maintenance costs and the implementation of predictive techniques becomes an incentive for developing systems capable of identifying defects intrinsic to the machine. Given this, this research aims to constitute a robust methodology based on intelligent systems for diagnosing and detecting broken bar defects in three-phase induction motors still in the transient regime of operation. To this end, a systematic search of the literature was initially carried out to achieve the state of art of detecting and diagnosing defects in rotating electrical machines. The proposed intelligent system was validated and tested with experimental data obtained under normal and defective engine conditions. In this way, a configuration based on computational intelligence was proposed to identify and supervise defects in these elements. These architectures were evaluated and compared employing a multicriteria measure, unprecedented for this purpose, to attest to the proposed methodology's relevance concerning the already existing ones. The results showed an agile and robust methodology for recommending defect detection and diagnosis systems in induction motors based on the characteristics of the attribute selection and machine learning algorithms.

Keywords: Engine defects. Expert system. Artificial intelligence. Pattern Recognition. Time series analysis. Robustness analysis.

Lista de ilustrações

Figura 1	Etapas da revisão sistemática utilizada no trabalho	34
Figura 2	Distribuição do número de trabalhos por ano de publicação.	36
Figura 3	Distribuição dos trabalhos selecionados nas principais bases de dados de interesse.	36
Figura 4	Esquema geral do laboratório de ensaios de máquinas elétricas	52
Figura 5	Excitação do enrolamento de campo do gerador de corrente contínua (GCC)	53
Figura 6	Sensores Hall de corrente e tensão	53
Figura 7	Procedimento para inserção de defeitos de barra em rotores	54
Figura 8	Representação dos pontos de coletas das variáveis de vibração mecânica	55
Figura 9	Fluxograma com a sequência lógica da metodologia aplicada neste trabalho	58
Figura 10	Exemplo de identificação da existência de dados incoerentes para o sinal de corrente em função do tempo (a) com relação a dados que representam o comportamento real esperado (b) do sinal de corrente na fase B	61
Figura 11	Diagrama representativo do método para identificação do regime transitório implementado neste trabalho	63
Figura 12	Medida de desempenho multicritério	70
Figura 13	Fluxograma referente à metodologia empregada para a verificação da robustez dos sistemas propostos	71
Figura 14	Fluxograma referente à metodologia empregada para a análise de robustez dos sistemas propostos	72
Figura 15	Sistema de detecção e de diagnóstico proposto	76
Figura 16	Ilustração da medida multicritério para avaliação de desempenho de classificadores, para cada algoritmo de seleção de atributos (SA), para o caso de classificação binária	86

Figura 17	Áreas do hexágono multicritério para cada configuração de algoritmos de seleção de atributos (SA) e de aprendizado de máquina (AM) para o caso de classificação binária	87
Figura 18	Ilustração da medida multicritério para avaliação de desempenho de classificadores, para cada algoritmo de seleção de atributos (SA), para o caso de classificação multiclasse	90
Figura 19	Áreas do hexágono multicritério para cada configuração de algoritmos de seleção de atributos (SA) e de aprendizado de máquina para o caso de classificação multiclasse	91
Figura 20	Porcentagem de atributos selecionados por cada algoritmo de seleção de atributos (SA) para o caso da classificação binária, em função da porcentagem de ruído branco agregado ao sinal de entrada para a extração dos atributos	94
Figura 21	Taxa de falsos positivos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária	94
Figura 22	Taxa de falsos negativos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária	95
Figura 23	Taxa de erros em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária	95
Figura 24	Área complementar à curva ROC em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária	96
Figura 25	Área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação binária	96
Figura 26	Porcentagem de atributos selecionados por cada algoritmo de seleção de atributos (SA) para o caso da classificação multiclasse, em função da porcentagem de ruído branco agregado ao sinal de entrada para a extração dos atributos	101
Figura 27	Taxa de falsos positivos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse	101
Figura 28	Taxa de falsos negativos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse	102

Figura 29	Taxa de erros em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse	102
Figura 30	Área complementar à curva ROC em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse	103
Figura 31	Área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação multiclasse	103
Figura 32	Análise da robustez por meio da área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação binária	107
Figura 33	Análise da robustez por meio da área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação multiclasse	109
Figura 34	Avaliação do desempenho preditivo dos atributos oriundos de cada tipo de sinal experimental avaliado utilizando o seletor de atributos CBF e o classificador SVM	113
Figura 35	Avaliação do desempenho preditivo dos atributos oriundos de cada tipo de sinal experimental avaliado utilizando o seletor de atributos CBF e o classificador RD	115
Figura 36	Categorias genéricas de curtose considerando a distribuição gaussiana .	138
Figura 37	Resultados para a análise de robustez segundo o parâmetro taxa de falsos positivos para o caso da classificação binária	147
Figura 38	Resultados para a análise de robustez segundo o parâmetro taxa de falsos negativos para o caso da classificação binária	148
Figura 39	Resultados para a análise de robustez segundo o parâmetro taxa de erros para o caso da classificação binária	148
Figura 40	Resultados para a análise de robustez segundo o parâmetro área complementar a curva ROC para o caso da classificação binária	149
Figura 41	Resultados para a análise de robustez segundo o parâmetro taxa de falsos positivos para o caso da classificação multiclasse	149
Figura 42	Resultados para a análise de robustez segundo o parâmetro taxa de falsos negativos para o caso da classificação multiclasse	150

Figura 43	Resultados para a análise de robustez segundo o parâmetro taxa de erros para o caso da classificação multiclasse	150
Figura 44	Resultados para a análise de robustez segundo o parâmetro área complementar a curva ROC para o caso da classificação multiclasse	151

Lista de tabelas

Tabela 1	Combinações de estratégias e direções de busca	47
Tabela 2	Maneira com a qual os dados foram agrupados e organizados para as análises posteriores	60
Tabela 3	Fatores de multiplicação para a normalização dos dados em função do carregamento imposto ao eixo dos motores de indução trifásicos (MITs)	62
Tabela 4	Categorias e medidas extraídas como descritores dos conjuntos de dados	65
Tabela 5	Organização das configurações de algoritmos de seleção de atributos (SA) e de aprendizado de máquinas (AM) dos sistemas inteligentes para diagnóstico por meio de códigos identificadores (IDs)	78
Tabela 6	Atributos selecionados por cada um dos algoritmos de seleção de atributos (SA) para o caso de classificação binária	81
Tabela 7	Atributos selecionados por cada um dos algoritmos de seleção de atributos (SA) para o caso de classificação multiclasse	82
Tabela 8	Análise da distribuição normal pelos atributos extraídos em função de cada categoria	84
Tabela 9	Porcentagem de atributos selecionados em função do tipo de sinal experimental usado para a extração de atributos para cada algoritmo de seleção de atributos (SA) para os dois cenários experimentais	85
Tabela 10	<i>Ranking</i> para seleção dos sistemas de detecção para o cenário de classificação binária	88
Tabela 11	Comparação dos modelos induzidos em termos de média e desvio padrão, entre parênteses, para medidas de importância para a classificação binária	89
Tabela 12	<i>Ranking</i> para seleção dos sistemas de diagnóstico para o cenário de classificação multiclasse	91
Tabela 13	Comparação dos modelos induzidos em termos de média e desvio padrão, entre parênteses, para medidas de importância para a classificação binária	92

Tabela 14	<i>Ranking</i> para identificação dos sistemas de detecção para o cenário de classificação binária em função da verificação da robustez	99
Tabela 15	<i>Ranking</i> para identificação dos sistemas de diagnóstico para o cenário de classificação multiclasse em função da verificação da robustez	106
Tabela 16	<i>Ranking</i> para identificação dos sistemas de detecção para o cenário de classificação binária em função da análise de robustez	108
Tabela 17	<i>Ranking</i> para identificação dos sistemas de diagnóstico para o cenário de classificação multiclasse em função da análise de robustez	110
Tabela 18	Relação dos atributos selecionados e do desempenho do sistema de inferência composto pelo seletor de atributos CBF e pelo classificador SVM quando utilizado apenas um dos tipos dos sinais experimentais coletados para o sistema de detecção de defeitos, classificação binária .	113
Tabela 19	Relação dos atributos selecionados e do desempenho do sistema de inferência composto pelo seletor de atributos CBF e pelo classificador RD quando utilizado apenas um dos tipos dos sinais experimentais coletados para o sistema de diagnóstico da severidade de defeitos, classificação multiclasse	116
Tabela 20	Formato atributo-valor	137

Lista de siglas

ACROC Área Complementar a Área sob a curva *Receiver operating characteristics*

AG Algoritmos Genéticos

AM Aprendizado de Máquina

ARIMA Modelo Auto-regressivo Integrado de Médias Móveis

CBF Filtro Baseado em Consistência

CFS Seleção de Atributos Baseada em Correlação

DCO Decomposição em Componentes Ortogonais

EESC Escola de Engenharia de São Carlos

FFT *Fast Fourier Transform*

GCC Gerador de Corrente Contínua

HW Holt-Winters

ID Código Identificador

kNN k-Vizinhos mais Próximos

LACIME Laboratório de Controle Inteligente de Máquinas Elétricas

LAIPS Laboratório de Automação Inteligente de Processos e Sistemas

MA Médias Móveis

MAE Erro Absoluto Médio

MAPE Erro Percentual Absoluto Médio

MASE Erro Médio Absoluto em Escala

MCC Máquina de Corrente Contínua

MCSA *Motor Current Signature Analysis*

MIT Motor de Indução Trifásico

MLP Multilayer Perceptron

MSE Erro Quadrático Médio

NB Naive Bayes

NRMSE Raiz do Erro Quadrático Médio Normalizado

O() Notação *Big-O* de complexidade computacional

PAS Porcentagem de Atributos Seleccionados

POCID *Prediction of Change in Direction*

PSO *Particle Swarm Optimazation*

R Coeficiente de Correlação

R^2 Coeficiente de Determinação

RAE Erro Absoluto Relativo

RBF Rede com Função de Ativação de Base

RD *Random Forest*

RMSE Raiz do Erro Quadrático Médio

RNA Redes Neurais Artificiais

ROC *Receiver Operating Characteristics*

SA Seleção de Atributos

SARIMA Modelo Autorregressivo Integrado de Médias Móveis Sazonal

SES Suavização Exponencial Simples

SETAR *Self-exciting Threshold Autoregressive*

SIF Sistema de Inferência *Fuzzy*

SMAPE Erro Percentual Absoluto Médio Simétrico

SVM *Support Vector Machine*

TA Tempo de Aprendizado

TE Taxa de Erro

TFN Taxa de Falsos Negativos

TFP Taxa de Falsos Positivos

USP Universidade de São Paulo

WT *Wavelet Transform*

Sumário

1	Introdução	27
1.1	Objetivos	30
1.2	Assertivas	31
1.3	Hipótese	31
1.4	Organização do Trabalho	31
1.5	Principais Contribuições	32
2	Revisão Sistemática da Literatura	33
2.1	Método da Revisão Sistemática	33
2.2	Resultados da Revisão Sistemática e Meta Análise	35
2.3	Estado da Arte	39
2.3.1	Defeitos em Motores de Indução	39
2.3.2	Técnicas de Detecção e de Diagnóstico de Defeitos	40
3	Aspectos da Bancada Experimental e Delineamento Experimental	51
3.1	Disposição Geral da Bancada de Experimentos	51
3.2	Inserção de Defeitos nas Barras dos Rotores	54
3.3	Delineamento Experimental	55
4	Metodologia	57
4.1	Pré-Processamento e Extração de Atributos	57
4.1.1	Dados Incoerentes	58
4.1.2	Normalização dos Dados em Função da Variação dos Carregamentos	61
4.1.3	Método para Identificação do Regime Transitório	62
4.1.4	Atributos Extraídos	63
4.2	Seleção dos Atributos Significativos	64
4.3	Modelos de Aprendizado de Máquina	67
4.4	Avaliação de Desempenho	69
4.5	Análise e Verificação da Robustez dos Sistemas Propostos	70

4.6	Sistema de Diagnóstico e Monitoramento de Defeitos em MITs	75
5	Resultados e Discussões	79
5.1	Banco de Dados Experimental	80
5.2	Análise dos Atributos Seleccionados	80
5.3	Comparação dos Modelos Induzidos	86
5.4	Verificação da Robustez dos Modelos Induzidos	93
5.4.1	Verificação da Robustez dos Modelos Induzidos para a Classifica- ção Binária	93
5.4.2	Verificação da Robustez dos Modelos Induzidos para a Classifica- ção Multiclasse	100
5.5	Análise da Robustez dos Modelos Induzidos	106
5.6	Recomendação de Sistema de Detecção e Diagnóstico	111
6	Conclusão	119
	Referências	123
	Apêndices	135
APÊNDICE A	Descrição dos Atributos Extraídos	137
A.1	Medidas Estatísticas	137
A.2	Medidas Baseadas em Informação	139
A.3	Medida Baseada em Complexidade	141
A.4	Medidas Baseadas na Análise de Sistemas	142
APÊNDICE B	Resultados Complementares da Análise de Robustez	147

Capítulo 1

Introdução

A necessidade crescente de busca por processos mais racionais e econômicos no que tange à utilização eficiente da energia elétrica tem implicado na realização de estudos detalhados para a otimização das técnicas de manufatura, os quais visam também a adequada conservação da energia elétrica. Dentro deste contexto, os motores elétricos são equipamentos indispensáveis na maioria dos processos que envolvem os setores produtivos, pois os mesmos se destacam como principais elementos de conversão de energia elétrica em mecânica. Como consequência, essas máquinas elétricas passaram por melhorias que atendem às necessidades emergentes dos processos produtivos por sua importância energética dentro do cenário nacional (PUCCINI et al., 2015; ROMERO-TRANCOSO, 2011).

Grande parte das aplicações industriais que necessitam de conversão de energia elétrica em mecânica utilizam máquinas elétricas rotativas, como os motores de indução, pois o baixo custo associado à longa durabilidade refletem suas características mais atrativas. Além disso, a crescente necessidade dos setores produtivos frente à manutenção preventiva e preditiva, que visam à redução de custo de manutenção, minimização de transtornos gerados por desligamentos não planejados de linhas de produção, e até mesmo inibição de riscos de acidentes de trabalho, torna-se relevante e motivador o desenvolvimento de sistemas capazes de identificar os defeitos incipientes inerentes à máquina (ZHANG et al., 2014).

Os motores elétricos estão expostos a uma variedade de condições ambientais, operando sob diversos tipos de cargas, os quais em conjunto com o envelhecimento natural podem ocasionar falhas inerentes às próprias máquinas ou associadas às condições severas. Neste contexto, o monitoramento, a fim de detectar e diagnosticar os defeitos em seus estágios iniciais, é de suma importância para tratá-los adequadamente, evitando falhas e, consequentemente, reduzindo o custo de manutenção, o custo de estoque de peças e o desligamento do setor industrial no qual as máquinas estão contidas (BINDU; THOMAS, 2014).

As falhas em motores elétricos podem ser, no geral, de característica elétrica, como

defeitos no estator e no rotor, e mecânicas, como falhas de excentricidade e nos componentes mecânicos. Dessa maneira, a aplicação de técnicas para a predição e prevenção de falhas mecânicas e elétricas aumentam a disponibilidade e a confiabilidade das máquinas elétricas rotativas.

Os métodos e protocolos empregados para o diagnóstico e proteção contra falhas são tão antigos quanto as próprias máquinas. Quanto às aplicações em Motores de Indução Trifásicos (MITs), inicialmente, os fabricantes contavam com proteções simples, como sobrecorrente e sobretensão, com o intuito de garantir segurança e confiabilidade na operação. No entanto, como as funções das máquinas rotativas foram se tornando cada vez mais complexas, as melhorias nas áreas de diagnóstico e de monitoramento de defeitos são cada vez mais necessárias. Frente a isso, além do monitoramento da saúde estrutural das máquinas elétricas rotativas, é fundamental o acompanhamento dos defeitos desde o início, visando reduzir os custos e otimizando as funções de manutenção e evitando perdas financeiras (LIBONI, 2017; COSTA; KASHIWAGI; MATHIAS, 2015).

Há diversos defeitos que podem evoluir para uma falha, impedindo o desempenho da função requerida de máquinas elétricas rotativas, sendo os principais listados (VAS, 1993):

- ❑ falhas no estator resultando na abertura ou curto-circuito de um ou mais enrolamentos de fase do estator;
- ❑ conexões anormais dos enrolamentos do estator;
- ❑ barra do rotor quebrada ou terminais do rotor trincados;
- ❑ irregularidades estáticas ou dinâmicas da abertura de ar;
- ❑ atrito entre o estator e o rotor, excentricidade, causando problemas ao núcleo do estator e aos enrolamentos;
- ❑ enrolamento em campo do rotor em curto;
- ❑ falhas nos mancais de rolamentos e na caixa de engrenagens.

Das falhas apresentadas, as hegemônicas estão presentes no rolamento, no estator ou na armadura, nas barras do rotor e as relacionadas à excentricidade. Quanto aos defeitos nas barras do rotor, apresentam diversos sintomas, sendo prevalecentes as tensões cabo máquina desequilibradas e correntes de linha, aumento do torque intermitente, diminuição do torque médio, redução de eficiência e aquecimento excessivo (VAS, 1993; KLIMAN; STEIN, 1990). Dessa maneira, essa pesquisa foi direcionada inteiramente ao defeito de barras quebradas do rotor de modo isolado.

Quando uma barra do rotor é interrompida, nenhuma corrente flui por ela. Como resultado, nenhum fluxo magnético é gerado ao redor da barra. Assim, é gerada assimetria no campo magnético do rotor, gerando um campo de rotação para trás diferente de zero

que gira na frequência de deslizamento do rotor (MEHRJOU, 2011). Esta ocorrência leva à indução da componente de falha harmônica na corrente do estator. Esses harmônicos sobrepostos são usados como uma assinatura de falha de barra rompida por técnicas comumente usadas para diagnóstico, como a análise de assinatura da corrente do motor. No entanto, no regime transitório de operação, há ruído e componentes aleatórios dos sinais que podem gerar falsos positivos quando analisados por técnicas em regime de operação estacionário (SINGH; NAIKAN, 2018).

Atualmente, a maioria das técnicas desenvolvidas para detecção e diagnóstico de defeitos de barras quebradas do rotor em MIT são aplicadas em sinais no regime permanente, ou estacionário (ZAMUDIO-RAMIREZ et al., 2020; KANG et al., 2018; LOPES et al., 2015; PALÁCIOS et al., 2015; SERONATO; NETO, 2015; DIAO et al., 2013; ROMERO-TRONCOSO et al., 2011; EREN; DEVANEY, 2004). No entanto, para algumas aplicações, é interessante o desbravamento dos defeitos e indícios de ocorrências ainda no estágio inicial de operação, ou no regime transitório. Essas aplicações são baseadas no acionamento intermitente de máquinas e também para testes de comissionamento iniciais após a instalação (PONS-LLINARES et al., 2016; BINDU; THOMAS, 2014).

Ainda assim, o interesse em detectar e diagnosticar defeitos em regime transiente é motivado pelo excelente desempenho das técnicas em condições em que os métodos em regime permanente apresentam baixo desempenho, como na detecção de barras quebradas em máquinas com baixo escorregamento nominal, quebras de barra externa na máquina de indução de gaiola dupla, quebras de barra em máquinas com dutos de ar axiais, assimetrias e excentricidades de rotores em máquinas com oscilações de carga contínua (PONS-LLINARES et al., 2016; YANG et al., 2014; ANTONINO-DAVIU et al., 2012; PONS-LLINARES et al., 2011). Nestes casos, as falhas dificilmente são detectadas no estado estacionário de operação, pois estão acopladas a componente fundamental ou ruído, ou ainda podem levar a um diagnóstico de falso positivo ou erro tipo I (PONS-LLINARES et al., 2016).

Caso contrário, as máquinas de indução que ocasionalmente operam em um estado estacionário restrito estão ganhando importância no ambiente industrial. Exemplos relevantes de tais aplicações são a geração de energia eólica, veículos elétricos ou motores movidos por conversores eletrônicos que operam sistemas controlados em circuito fechado. Nessas aplicações, os motores de indução geralmente passam por mudanças contínuas e aleatórias nas condições de carga, velocidade ou alimentação. A aplicação de técnicas convencionais de estado estacionário como a *Fast Fourier Transform* (FFT), que calcula a transformada de Fourier discreta, leva a espectros borrados onde os componentes da falha se espalham e se tornam não identificáveis (HENAO et al., 2014). O uso de modernas técnicas diagnósticas não estacionárias, como as propostas por Vedreño-Santos et al. (2014), Gritli et al. (2011), permite o diagnóstico de defeitos neste tipo de condição.

O diagnóstico em condições transitórias baseia-se na análise das grandezas diagnósticas em uma ampla gama de condições de trabalho, ao contrário do diagnóstico em regime

permanente, que avalia o comportamento da máquina em um dado regime para diversas condições de trabalho. Consequentemente, o diagnóstico em condições transitórias é conceitualmente mais confiável quanto às aplicações do que um diagnóstico em estado estacionário.

A maioria dos métodos convencionais de detecção e diagnóstico de estado estacionário é baseada na identificação de componentes de falha distintos em um espectro de corrente. Da mesma forma, os métodos de diagnóstico no estado estacionário geralmente são baseados em sistemas inteligentes, combinando técnicas de seleção de atributos e aprendizado de máquina para identificar a evolução dos componentes de falha. Para tanto, os atributos extraídos dos sinais constituem a etapa fundamental do diagnóstico de defeitos no regime transiente de operação (PONS-LLINARES et al., 2016; MACIEJEWSKI; TREML; FLAUZINO, 2020). Como a detecção e o diagnóstico de defeitos ainda no regime transiente de operação estão sujeitos a diversos fatores aleatórios, como ruído, as técnicas utilizadas devem ser robustas ao ruído, reduzindo a taxa de erros e falsos alarmes (ROCHA et al., 2020; MALEK; ABDELSALAM; HASSAN, 2017).

Com a finalidade de monitorar e detectar defeitos em máquinas elétricas rotativas, muitos métodos foram desenvolvidos até o momento, nos mais diversos campos da ciência. Desta maneira, este trabalho apresenta o desenvolvimento de uma metodologia acerca do tema de detecção e diagnóstico de defeitos em máquinas elétricas rotativas, ainda no regime transitório de operação, para a recomendação de técnicas inteligentes robustas e de acordo com as condições impostas. O campo dessa pesquisa é amplo, pois existem técnicas que utilizam de sinais temporais e em frequência, de temperatura, de ruídos e de vibrações, de sinais de tensão e de corrente elétrica e os baseados em sistemas inteligentes. Assim, nesta pesquisa foram utilizados sinais de tensão e corrente elétrica e de velocidade de vibração mecânica para alcançar os objetivos propostos.

1.1 Objetivos

O objetivo geral é desenvolver uma metodologia para a detecção e o diagnóstico da severidade de defeitos de barras quebradas de rotores em motores de indução trifásicos no regime transitório utilizando técnicas de sistemas inteligentes. Nesse contexto, para o alcance do objetivo geral, os seguintes objetivos específicos foram estabelecidos:

1. Realizar uma busca sistemática das técnicas utilizadas para identificação e monitoramento de defeitos de barras quebradas em motores de indução trifásicos;
2. Utilizar técnicas de pré-processamento para limpeza e adequação dos dados obtidos por meio de bancada experimental;

3. Selecionar e aplicar algoritmos de seleção de atributos e ferramentas de aprendizagem de máquinas que sejam capazes de identificar o defeito e caracterizar a sua severidade por meio de atributos extraídos;
4. Avaliar o desempenho e a robustez das técnicas utilizadas e criar um processo para recomendação de configurações de processos para identificar e caracterizar os defeitos em máquinas elétricas de indução trifásicas.

1.2 Assertivas

Assertiva 1: Defeitos de barra quebrada do rotor afetam os sinais de corrente, de tensão elétrica e de vibração mecânica em máquinas de indução trifásicas.

Assertiva 2: Os sinais elétricos e mecânicos são passíveis de discretização nos regimes transitório e permanente.

Assertiva 3: Características de sinais simplificam a análise quando se necessita de comparação.

1.3 Hipótese

A presença de defeitos de barra quebrada em rotores de motores de indução trifásicos afetam os sinais de corrente, de tensão elétrica e de vibração mecânica e é possível identificar a presença e a severidade de defeitos por meio de aprendizado de máquinas utilizando características extraídas destes sinais.

1.4 Organização do Trabalho

No Capítulo 2, Revisão Sistemática da Literatura, são descritos os protocolos utilizados na realização de uma busca sistematizada da literatura e os resultados da meta-análise dos trabalhos encontrados. Não obstante, são apresentadas as inovações publicadas, patenteadas e registradas no tema de detecção e diagnóstico de falhas em MITs. E, para os trabalhos que apresentaram objetivos similares à questão estabelecida neste trabalho, são realizadas as comparações pertinentes.

No Capítulo 3, Aspectos da Bancada Experimental e Delineamento Experimental, é apresentado o enfoque construtivo da bancada experimental e o modo de aquisição de dados. Ainda, são descritos os procedimentos para a construção do banco de dados que foram utilizados nos capítulos seguintes deste trabalho.

No Capítulo 4, Metodologia, são descritos os protocolos do pré-processamento dos dados, da seleção de atributos (SA), do aprendizado de máquina (AM) e do método de avaliação do desempenho e da robustez dos sistemas de detecção e de diagnóstico

construídos. Já no Capítulo 5, Resultados e Discussões, são apresentados e discutidos os resultados obtidos nos procedimentos experimentais para o estudo de caso de detecção e de diagnóstico de defeitos em MITs para o banco de dados concebido.

Ao final, no Capítulo 6, Conclusão, é apresentada a conclusão obtida com relação aos objetivos enunciados neste trabalho.

1.5 Principais Contribuições

As principais contribuições deste trabalho de pesquisa para a comunidade científica são listadas a seguir:

- ❑ Publicação da revisão sistemática da literatura desenvolvida na 2020 *FORTEI-International Conference on Electrical Engineering* (FORTEI-ICEE) em Bandung, Indonésia, que foi indexada pela base *IEEEExplore*;
- ❑ Publicação da base de dados e dos procedimentos experimentais utilizados neste trabalho na base de dados *IEEE Dataport* em 2020;
- ❑ Submissão de artigo científico na revista *Mechanical Systems and Signal Processing*, em 2021, acerca da análise de robustez na detecção e diagnóstico de defeitos no regime permanente de operação de motores de indução;
- ❑ Submissão de artigo científico na revista *Mechanical Systems and Signal Processing*, em 2021, com a temática de detecção e diagnóstico de defeitos em regime transitório de operação de motores de indução trifásicos avaliando a robustez e o desempenho preditivo dos atributos.

Capítulo 2

Revisão Sistemática da Literatura

Neste trabalho, inicialmente, foi realizada uma revisão sistemática da literatura, incluindo patentes e registros de *software*, com o objetivo de elucidar o estado da arte referente ao tema proposto. Essas informações compiladas serviram de base para a proposição de um método para detecção e diagnóstico de defeitos em máquinas elétricas rotativas.

2.1 Método da Revisão Sistemática

O método sistemático de revisão de literatura fornece um rigoroso processo replicável para se atingir as evidências relevantes para uma questão de pesquisa específica (KITCHENHAM; CHARTERS, 2007). Embora esse método tenha surgido em áreas como medicina, atualmente existem diretrizes e aplicações em outras áreas. Nas ciências exatas, várias aplicações podem ser encontradas, incluindo uma revisão sistemática que pesquisa outras revisões sistemáticas. A revisão sistemática da literatura, neste trabalho, foi realizada de acordo com as etapas e os aspectos inseridos na Figura 1.

A revisão sistemática teve o período de busca entre Janeiro de 2013 e Dezembro de 2020 para abordar o lastro das pesquisas concebidas pelo grupo no qual esse projeto faz parte. Esse processo de busca, de seleção e de análise dos artigos pode ser realizado de modo cronológico com possibilidades de realimentação a cada nova interação das etapas.

Sob a concepção metodológica da revisão sistemática da literatura, para o primeiro passo do procedimento, foi proposto, com base nos objetivos desse trabalho, como questão de busca, Etapa 1: “Quais métodos, aplicativos e ferramentas têm sido utilizadas para analisar, simular, descrever e detectar o comportamento mecânico e elétrico de defeitos em máquinas elétricas rotativas?”.

A identificação das bases de dados que delinearam as buscas, Etapa 2, foi realizada reunindo as principais bases utilizadas em pesquisas na área das engenharias, especificamente, e em outras áreas pertencentes às ciências exatas. Com isso, as bases selecionadas

Figura 1 – Etapas da revisão sistemática utilizada no trabalho (Modificado de Nunes (2017)).

foram representadas pela *IEEEExplore*, *Scopus*, *Web of Science*, *SciELO*, *Wiley* e *CiteSeer* (PUCCINI et al., 2015; CAMPOS, 2013).

A Etapa 3 foi definida como a etapa da determinação dos critérios de inclusão e de exclusão de artigos para posterior análise e para formarem a meta-análise da literatura relacionada. Os critérios de seleção de artigos foram construídos em modalidades de perguntas para afunilar a escolha dos trabalhos de interesse, e estão evidenciados a seguir:

- ☐ O trabalho apresenta relação com a questão de pesquisa?
- ☐ O trabalho possui relação com implementação de sistemas, aplicativos ou ferramentas para análise e o monitoramento de falhas em máquinas elétricas rotativas?
- ☐ O método do trabalho é realizado com base em experimentos, simulações ou análises matemáticas?
- ☐ O trabalho é acessível institucionalmente?
- ☐ O trabalho não é repetido nas bases de dados buscadas?
- ☐ O trabalho é redigido na língua portuguesa, espanhola ou inglesa?

Vale ressaltar que esses critérios são de inclusão de artigos para a análise crítica, e, se porventura algum desses critérios não forem atendidos, o artigo deve ser excluído.

Assim, finalizando a parte de coleta e de seleção de artigos, iniciou-se a busca e a seleção dos artigos, Etapa 4, de acordo com os critérios determinados na Etapa 3. Para isso, foram definidas palavras-chave ou palavras similares e sinônimos relacionadas ao tema para construção de uma *string* de busca. Com isso, a *string* identificada é então apresentada com os operadores lógicos para a inserção nas bases de busca: “(*“electrical machines”OR “rotary machines”OR “motors”OR “induction machine”*) AND (*“machine*

learning"OR "fault diagnosis"OR "attribute selection"OR "artificial intelligence"OR "prediction"OR "forecasting"OR "pattern recognition") AND ("timeseries"OR "time-series"OR "time series"OR "frequency spectrum"OR "dynamic systems")".

Depois da construção da *string* e dos critérios de seleção, iniciou-se a busca em cada uma das bases de dados selecionadas e os atributos adquiridos da busca eram inseridos em uma planilha eletrônica. Após o período de busca, no primeiro momento é realizada a segunda parte da revisão, ou seja, a seleção e a organização dos trabalhos coletados, apresentada na Seção 2.2.

2.2 Resultados da Revisão Sistemática e Meta Análise

Com a aplicação do processo de busca sistemático da literatura, foram identificadas 3388 publicações, tendo esse número reduzido para 734 após a aplicação dos critérios de inclusão. Em seguida, foi realizada a leitura do título e resumo de todos os trabalhos apurados para efetivar a seleção. Quando a seleção não pode ser concretizada incontestavelmente, considerou-se a leitura do trabalho completo, até que não restassem dúvidas sobre a relevância do trabalho. Ao final desse processo, obtiveram-se 169 publicações das quais foram extraídas informações para responder à questão de pesquisa e compor as Etapas 5, 6 e 7 do método da revisão sistemática.

Na Figura 2 é exibido um gráfico de barras contendo a distribuição, por ano de publicação, do número de trabalhos selecionados ao término da revisão sistemática da literatura. O gráfico mostra que entre os anos de 2013 e 2016, a média de trabalhos publicados era de 9,5 por ano, enquanto que nos anos seguintes até o presente, a média cresceu para aproximadamente 33 trabalhos publicados por ano. Tal fato pode ser explicado devido à consolidação do interesse pelo tema pesquisado, principalmente nas aplicações de sistemas inteligentes para a detecção de defeitos em máquinas elétricas.

Figura 2 – Distribuição do número de trabalhos por ano de publicação

Complementarmente, na Figura 3 é apresentado um gráfico com a distribuição de trabalhos selecionados para cada uma das bases de dados buscadas. Quanto a esse quesito, as bases de dados *IEEEExplore* e *Scopus* se sobressaem em quantidade de trabalhos publicados com relação às outras bases, sendo excluídos os trabalhos duplicados. Esse fato é nítido pelo caso dessas bases de dados serem as mais procuradas por pesquisadores das áreas de engenharias e de ciências exatas.

Figura 3 – Distribuição dos trabalhos selecionados nas principais bases de dados de interesse

A análise dos 169 trabalhos selecionados na revisão sistemática permitiu implementar a seguinte meta-análise, seguindo as Etapas 5, 6 e 7 do método da revisão sistemática:

Defeitos dos Motores de Indução: 40,83% dos trabalhos estudaram os defeitos dos mancais, 8,85% os defeitos no estator, 20,71% os defeitos da barra quebrada no

rotor e 17,75% os defeitos de excentricidade. Alguns estudos analisaram defeitos em conjunto, sendo 11,83%, discretizando ou não os defeitos;

Sinais Adquiridos: 57,39% dos estudos analisaram sinais elétricos, 29,59% sinais mecânicos e 13,02% analisaram sinais híbridos, contendo sinais elétricos e térmicos ou mecânicos e térmicos;

Regime de Análise: os estudos que implementaram os métodos de análise em regime estacionário foram de 79,88%, em regime transitório de 2,96% e em regime híbrido de 17,16%;

Domínio de Análise: 55,03% dos artigos utilizaram métodos no domínio do tempo, 23,67% no domínio da frequência, 8,88% no domínio da tempo-frequência e apenas 21 artigos aplicaram o domínio híbrido. Vale ressaltar que, no domínio da frequência, 55,03% dos artigos utilizaram a técnica *Motor Current Signature Analysis* (MCSA);

Métodos de Predição: 55,03% dos trabalhos empregaram a técnica MCSA, 47,93% empregaram Redes Neurais Artificiais (RNA), 37,27% construíram modelos Autorregressivos Integrados de Médias Móveis (ARIMA) ou Autorregressivos Integrados de Médias Móveis Sazonais (SARIMA), 13,02% usaram modelos híbridos, 21,89% aplicaram Máquinas de Suporte Vetorial (SVM), 11,83% executaram variações do algoritmo k-Vizinhos mais Próximos (kNN), 8,88% usufruíram dos métodos de Médias Móveis (MA) e de Suavização Exponencial Simples (SES), 4,14% abordaram técnicas bayesianas e 4,14% recorreram aos modelos de Holt-Winters (HW);

Abordagens para Predição de Séries Temporais: Aplicações reais foram tema de 57 trabalhos, dos quais 59,65% utilizaram algoritmos decorrentes da abordagem não-paramétrica, 19,30% empregaram métodos da abordagem paramétrica e 21,05% construíram modelos híbridos baseados em ambas as abordagens;

Seleção de Atributos: Os artigos que apresentaram uma seleção de atributos foram 66,27%. Desses, 75,89% utilizaram a abordagem de filtro, 17,86% a abordagem *wrapper* e apenas 6,25% a abordagem embutida;

Medidas de Avaliação: 24,85% dos trabalhos adotaram o Erro Absoluto Médio (MAE), 28,99% avaliaram a Raiz do Erro Quadrático Médio (RMSE), 30,18% verificaram o Erro Percentual Absoluto Médio (MAPE), 31,95% analisaram o Erro Percentual Absoluto Médio Simétrico (SMAPE), 20,71% utilizaram o Erro Quadrático Médio (MSE), 12,42% abordaram o Erro Médio Absoluto em Escala (MASE), 6,51% recorreram ao Coeficiente de Correlação (R) e 4,14% aplicaram o Erro Absoluto Relativo (RAE). As medidas U de Theil (U-Theil), Prediction of Change in Direction (POCID), Raiz do Erro Quadrático Médio Normalizado (NRMSE) e o Coeficiente de Determinação (R²) apareceram em 3,55% das publicações;

Configuração dos Parâmetros: 65,68% dos trabalhos selecionados aplicaram alguma técnica de busca capaz de realizar a escolha dos parâmetros do modelo da melhor maneira possível para ferramentas de aprendizado de máquinas. Entre essas técnicas, encontram-se a de treino e teste ou validação *holdout* e a de validação cruzada;

Método *Baseline*: Avaliações empíricas que contemplam algoritmos estatísticos e de aprendizado de máquina foram tema de 33 trabalhos;

Aprendizado de Máquinas: Dos 33 trabalhos que apresentaram modelos de aprendizado de máquinas, 33,33% utilizaram o algoritmo Naive Bayes, 21,21% o algoritmo de árvore de decisão C4.5, 24,24% RNA e 21,21% utilizaram *ensembles* de árvores de decisão;

Conjuntos de Dados Reais: O uso de conjuntos de dados reais esteve presente em 71 publicações, das quais 71,83% utilizaram dados produzidos em instituições, empresas ou indústrias e 28,17% investigaram dados mantidos em repositórios;

Conjuntos de Dados Sintéticos: 7,69% dos trabalhos selecionados construíram conjuntos de dados sintéticos para avaliar o desempenho de algoritmos de predição. Esses conjuntos de dados foram gerados computacionalmente para conter propriedades específicas, como a presença ou ausência de tendência e/ou sazonalidade;

Estudo de Caso: 78,70% realizaram estudos de casos reais, permitindo auxílio na tomada de decisão. 69,17% desses trabalhos tinham a finalidade de atuar nas engenharias, 21,80% tinha atuação nas áreas de informática e ciência da computação e 9,02% possuía aplicações nas áreas sócio-econômicas e saúde.

Com a meta-análise dos trabalhos selecionados, observa-se que a maioria dos métodos empregados para detecção de padrões em dados temporais vem utilizando sistemas inteligentes ou híbridos. Em específico, o uso de MCSA e RNA vem crescendo como integrador desses sistemas de detecção, com o auxílio de técnicas clássicas, como a FFT para calcular a transformada discreta de Fourier. Desse modo, pode-se prever que os próximos trabalhos que contemplem o estado da arte desse tema envolvam essas técnicas de aprendizado de máquinas. Assim, pode-se propor e testar técnicas híbridas com relação às expostas nesse trabalho para a construção de um método completo e robusto.

As técnicas de processamento estão relacionadas com o tipo de sinal. Recentemente, uma revisão sistemática sobre métodos de detecção e diagnóstico de defeitos típicos de motores de indução foi publicada, com o intuito de definir o estado da arte desta tecnologia. Segundo os autores, os dois principais sinais utilizados para essa finalidade são os elétricos e mecânicos, e em 79,88% das aplicações são em regime estacionário (MACIEJEWSKI; TREML; FLAUZINO, 2020). Dessa forma, foram analisadas técnicas de processamento de sinais de vibração mecânica e sinais elétricos em condição de regime permanente.

Durante a execução da meta-análise verificou-se alguns trabalhos, do total de 169, que envolviam avaliações e métodos mais populares e com resultados mais adequados com relação à temática da pesquisa. Frente a isso, esses trabalhos são fundamentais para compor o estado da arte do tema dessa pesquisa e possibilitam um caminho para a proposta de um novo método para monitoramento e detecção de defeitos em máquinas elétricas rotativas, baseado em sistemas inteligentes.

Considerando os resultados apresentados, o defeito dos rolamentos apresentou maior nível de precisão na detecção e diagnóstico, com percentual médio de 96,33% ao utilizar as técnicas de SVM aplicadas com atributos estatísticos, como curtose, média e desvio padrão nos ciclos de vibração mecânica. Portanto, entre os principais defeitos típicos dos motores de indução, o defeito nos mancais de rolamento é o que apresenta melhor representação física do defeito. Quanto aos defeitos do estator, foi obtida uma precisão de 87,22%. As falhas das barras quebradas do rotor, no entanto, apresentaram uma precisão média de 92,56% para os trabalhos analisados.

Quando foi comentado sobre defeitos híbridos, ou defeitos sobrepostos, os métodos de detecção e diagnóstico ainda não apresentam precisão suficiente para comparar com os métodos aplicados para apenas um defeito. Isso ocorre porque existem muitos fatores aleatórios envolvidos e a dificuldade de obter dados experimentais classificados.

Ainda assim, poucos estudos apresentaram métodos capazes de identificar defeitos em conjunto, ou híbridos, e quando aplicados, a precisão foi baixa. Isso é explicado pelo fato de ser um desafio monitorar mais de um defeito ao mesmo tempo, mas essa é uma realidade comum nos motores de indução, e um defeito pode levar ao aparecimento de outros ou se tornar uma falha. Com a aplicação do método deste trabalho, foi possível verificar a deficiência dos métodos atuais na detecção de mais de um defeito ao mesmo tempo. Mesmo assim, para fins de última geração, os defeitos no estator e na barra quebrada, mesmo que tenham sido frequentemente estudadas, precisam ser implementadas por métodos mais confiáveis e ágeis para realizar o monitoramento em tempo real, principalmente em análises no regime transitório de operação.

2.3 Estado da Arte

2.3.1 Defeitos em Motores de Indução

Na maioria das aplicações de MITs, as condições ambientais e o pouco conhecimento técnico das cargas acopladas ao eixo podem ocasionar diminuição em seu rendimento, sobreaquecimento e consequente redução em sua vida útil, má utilização da energia elétrica sob o aspecto do baixo fator de potência, defeitos mecânicos e elétricos, entre outros. Desta forma, observa-se na literatura a crescente necessidade da detecção e monitoramento de defeitos, permitindo identificação suficientemente rápida e evitando-se assim o

desligamento da produção e reduzindo custos de manutenção corretiva (GHANBARI, 2016; DRIF; CARDOSO, 2014; SANTOS, 2013).

As possíveis falhas em motores de indução podem ser classificadas, primariamente em elétricas e mecânicas. As falhas mecânicas se desdobram em falhas de excentricidade e falhas no rolamento, enquanto que as falhas elétricas em falhas no estator e falhas no rotor. Ainda, as falhas no estator podem ser no enrolamento ou externas, como no acionamento, e as falhas no rotor podem ser quebras do rotor ou quebras dos anéis terminais.

As falhas mecânicas podem resultar do desalinhamento do centro do eixo do rotor com o centro do estator, denominadas de excentricidade, e podem causar empenamento do eixo e danos nos mancais de rolamento. Geralmente são ocasionadas por montagens incorretas ou erros de fabricação. Os desgastes dos mancais de rolamentos causam comprometimento do desempenho da máquina, visto que o atrito e a excentricidade comprometem o regime de trabalho da máquina (BINDU; THOMAS, 2014; SEERA, 2014; MEHRJOU, 2011; NANDI; TOLIYAT; LI, 2005).

Os defeitos nos enrolamento ocorrem devido às deformidades na isolação estatórica, podendo ocasionar curto-circuito entre os enrolamentos do motor. Umidade, valores dielétricos intensos, esforços mecânicos, contaminação por óleo e folga na laminação do núcleo magnético são causas de defeitos de isolação. Esses defeitos ocasionam altas correntes elétricas influenciando diretamente a condição de operação do motor e causando sua degradação já que são falhas acumulativas (BINDU; THOMAS, 2014; DRIF; CARDOSO, 2014; MEHRJOU, 2011; NANDI; TOLIYAT; LI, 2005).

As falhas que ocorrem nos dispositivos de acionamento dos MITs são consideradas falhas externas e influenciam diretamente o desempenho do motor, tanto com relação ao torque do motor quanto à velocidade. As anormalidades nos acionamentos advêm de desbalanceamento de tensão, sobretensão e subtensão (SANTOS, 2013).

Por fim, defeitos no rotor, na grande maioria dos casos, devem-se às quebras das barras da gaiola de esquilo. Essas quebras são consequência de vibrações excessivas, oscilações de torque do motor e mudanças bruscas na dinâmica de partida (LIBONI, 2017; COSTA; KASHIWAGI; MATHIAS, 2015; GHORBANIAN; FAIZ, 2015; BINDU; THOMAS, 2014; GU, 2015; SEERA, 2014; MEHRJOU, 2011; NANDI; TOLIYAT; LI, 2005). Esses defeitos causam redução do tempo de serviço, ou disponibilidade do equipamento, diminuem o torque médio do motor e produzem vibrações. Além disso, esse tipo de defeito é acumulativo e destrutivo, pois barras adjacentes também se deterioram devido ao estresse mecânico e térmico (MATIC, 2012). A rachadura dos anéis da gaiola de esquilo também contribui para a composição dos defeitos em MITs, porém são menos incidentes (SANTOS, 2013).

2.3.2 Técnicas de Detecção e de Diagnóstico de Defeitos

Frente aos defeitos expostos que podem ocasionar falhas, percebe-se que o desenvolvimento de técnicas de diagnóstico de defeitos é de fundamental importância diante da

necessidade de confiabilidade nos processos e mitigação de consequências das falhas, de forma a garantir maior segurança e disponibilidade nas operações.

Os sistemas de diagnóstico de defeitos podem ser classificados em sistemas baseados em modelo ou em sinais. Os sistemas baseados em modelo utilizam modelos computacionais do sistema dinâmico em análise para se obter referências dos pontos de operação desejados. Porém, esses sistemas de diagnóstico possuem restrições como incertezas nos modelos e complexidade computacional. Consequentemente, sistemas baseados em sinais têm sido amplamente abordados pela literatura específica ao tema nos últimos anos, principalmente devido à complexidade de se modelar sistemas eletromecânicos (GHANBARI, 2016).

Os métodos de diagnóstico de falhas baseados em sinais consistem na leitura de variáveis do motor e em técnicas de processamento de sinais com o objetivo de extrair características que possam apontar o acontecimento de um evento anômalo. Por meio dessas características, ou atributos, estáticos ou inteligentes, índices são ajustados para acusar o acontecimento de um defeito, assim como diagnosticar sua severidade de forma automática. Desta maneira, o estado da arte de sistemas de diagnóstico surge em junção do processamento de sinais e dos sistemas de inferência de conhecimento (LIBONI, 2017).

Um sistema de diagnóstico adequado para a aplicação implica em uma arquitetura de variáveis de monitoramento apropriadas, técnicas de processamento de sinais e análise de compromisso entre custo e aplicabilidade de recursos. Segundo Ostojic (2013), o sistema de identificação deve ainda ser dedicado a encontrar a severidade de defeitos, de forma a auxiliar o tratamento das ocorrências. Variáveis como tensão e corrente elétricas, vibração mecânicas e temperatura são alguns exemplos de sinais monitorados. Uma análise cuidadosa da relação dessas variáveis com os possíveis defeitos de um MIT foi explicitada por Santos, Silva e Suetake (2012), por meio de uma compilação da literatura.

Uma das tarefas mais importantes de um sistema de detecção e de diagnóstico de defeitos reside no processamento dos sinais monitorados. Este processamento é primordial para um sistema de identificação confiável e preciso, pois é o responsável por extrair dos sinais características correlacionadas ao acontecimento de um defeito, de modo a permitir que classificadores inteligentes ou estatísticos consigam realizar o diagnóstico de maneira automática. Atualmente, métodos de processamento desses sinais se dão por ferramentas aplicadas no domínio do tempo ou da frequência, e no domínio tempo-frequência, como por exemplo, *Fast Fourier Transform* (FFT) para o cálculo da transformada de Fourier discreta e *Wavelet Transform* (WT) (YANG, 2016; COSTA; KASHIWAGI; MATHIAS, 2015; EBRAHIMI; FAIZ; AKIN, 2012; MEHRJOU, 2011).

A maioria dos sistemas de detecção e de diagnóstico possui suas análises de sinais baseadas em transformadas no domínio da frequência, mais especificamente no processamento de sinais advindos da assinatura espectral da corrente elétrica do motor, (MCSA) (YANG, 2016; GU, 2015; RANGEL-MAGDALENO, 2014; GARCIA-ESCUDEIRO, 2011; GARCIA-PEREZ, 2011). Análises da potência elétrica instantânea também são encontrados em Drif e Car-

doso (2014) e Bossio (2009). Uma das vantagens em se utilizar sinais elétricos para o monitoramento dos MITs se dá pelo fato desses sinais serem facilmente adquiridos de forma não invasiva e por meio de transdutores simples que atualmente são incorporados na operação do motor. No entanto, essa medida, MCSA, é geralmente aplicada para experimentos em motores em regime permanente, sendo a região transitória de funcionamento pouco explorada.

Por meio da busca de trabalhos relacionados, evidencia-se que problemas associados principalmente com a extração de características dos sinais monitorados ainda estão em aberto (YANG, 2016; MEHRJOU, 2011). Pode-se citar dificuldades de extração de características devido ao espalhamento espectral em algoritmos de processamento no domínio da frequência, à necessidade de largos períodos de aquisição de sinais, à falta de imunidade a ruídos dos algoritmos, principalmente quando motores são alimentados por inversores de frequência, à falta de implementações nos períodos transitórios de funcionamento de motores, entre outros. Em algoritmos de processamento no domínio do tempo, dificuldades de extração de características ocorrem em motores em baixo torque. Deste modo, novos estudos na área de identificação e diagnóstico de defeitos, principalmente na área de processamento de sinais, são necessários de forma a perpassar problemas existentes.

Os atributos provenientes do processamento de variáveis monitoradas devem revelar informações importantes para a manutenção e diagnóstico dos motores de indução, e dessa maneira, auxiliar a tarefa dos sistemas de inferência. Em estudos de prognóstico de degradação e vida útil de motores de indução, definem-se métricas para analisar se um índice resultante do processamento de sinais monitorados é um bom indicador de degradação do equipamento (COBLE; HINES, 2009). Um bom índice de prognóstico deve ser monotônico com a severidade do defeito, ou seja, deve ser estritamente crescente ou decrescente com a severidade. Além disso, o equipamento deve ser diagnosticado com defeito dentro de um intervalo pequeno de valores desse índice, ou seja, apresentar baixa variabilidade. Outro importante aspecto advindo dos conceitos de prognóstico é que as características indicativas da degradação devem ter a mesma tendência, ou seja, a mesma orientação funcional durante o tempo. Essas três características são denominadas de monotonicidade, variabilidade e trendabilidade.

Pode-se estender esses conceitos para a aplicação no diagnóstico de defeitos. A monotonicidade de um índice indicativo de defeito é importante, pois apresenta relação com a severidade da ocorrência. A variabilidade também deve ser tida como característica importante, pois mostra uma medida de variância do valor do índice significativo de falha dentre uma população de equipamentos. Uma grande variância pode resultar em dificuldades para sistemas de identificação e proteção. Já a trendabilidade mostra que os índices de degradação de uma população possuem a mesma tendência funcional.

Os sistemas de classificação e tomada de decisão para indicar a ocorrência do defeito também são importantes. Devem aprender a identificar características específicas do

defeito nos índices resultantes do processamento de sinais. Análises estatísticas e classificadores advindos da aprendizagem de máquinas são algumas das ferramentas utilizadas como sistemas de inferência.

Para detecção de barras quebradas, tema central desta pesquisa, vários trabalhos utilizam diversos sistemas inteligentes e outras ferramentas de aprendizagem de máquinas para a tomada de decisão, como as redes neurais artificiais (RNA) (GODOY, 2016; PAWLAK; KOWALSKI, 2010; SADEGHIAN; YE; WU, 2009), sistemas de inferência *Fuzzy* (SIFs) (SEERA, 2014; ROMERO-TRANCOSO, 2011), *Support Vector Machines* (SVMs) (GODOY, 2016; MATIC, 2012; BACCARINI, 2011), *Particle Swarm Optimazation* (PSO) (SOUALHI; CLEC; RAZIC, 2013), Algoritmos Genéticos (AG) (RAZIK; CORREA; SILVA, 2009), classificadores Bayesianos (SILVA; POVINELLI; DEMERDASH, 2013) e árvores de decisão (GARCIA-ESCUDEIRO, 2011; DY; BRODLEY, 2000; HALL, 1999). Nesses trabalhos, não só a detecção de defeitos é realizada, mas também um diagnóstico de sua severidade, a qual é representada pela quantidade de barras quebradas no rotor.

Até recentemente, a ferramenta de detecção de defeito prevalecente era a técnica MCSA. As referências Benbouzid e Hachemi (2000), Benbouzid e Kilman (2003) fornecem uma revisão útil desses métodos baseados em técnicas. Além das técnicas baseadas nos cálculos da transformada discreta de Fourier aplicadas à corrente da linha, defeitos de barras quebradas, defeitos do estator e defeitos de excentricidade dinâmica também podem ser detectadas usando um espectro de ordem superior. Em particular, bispectro e trispectro, a partir de uma única medição de transdutor de vibração radial. O estudo foi desenvolvido para detectar e diagnosticar defeitos em motores de indução alimentados por inversor e alimentados diretamente pela rede.

O trabalho de Ristanoski, Liu e Bailey (2013) destacou que integrar elementos do tempo no processo de aprendizagem constitui o maior desafio no uso de SVM para predição de séries temporais, pois elas são suscetíveis a erros quando mudanças de distribuição ocorrem com frequência ao longo da série. Para auxiliar nessa questão, foi investigada a distribuição de erros nas predições obtidas pelo algoritmo SVM. Uma vez identificadas as amostras que produziam os maiores erros, observou-se a sua correlação com as mudanças que ocorriam na distribuição da série histórica. O entendimento desse comportamento motivou os autores a propor uma função de perda dependente do tempo, a qual viabiliza a inclusão de informações sobre mudanças de distribuição da série diretamente no processo de aprendizagem. Os experimentos foram conduzidos a partir de dados reais, e o método proposto foi comparado com sua versão alternativa, que adota média quadrática, e com outros seis algoritmos de predição: RNA, kNN, SVM, redes com função de ativação de base (RBF), *robust regression* e modelo autorregressivo integrado de médias móveis sazonal (SARIMA). Os resultados, expressos de acordo com a medida erro quadrático médio, sugeriram que o uso de uma função de perda dependente do tempo pode reduzir a variância global dos erros e, portanto, acarretar em predições mais precisas.

Em Claveria e Torra (2014) foi investigado o desempenho preditivo dos algoritmos RNA, modelo auto-regressivo integrado de médias móveis (ARIMA) e *self-exciting threshold autoregressive* (SETAR). Dados mensais pré-processados de chegadas de turistas internacionais para Catalunha, entre os anos de 2001 a 2009, foram adotados no trabalho como indicadores oficiais da demanda turística. Ao comparar o desempenho dos métodos sobre distintos horizontes de predição, o modelo ARIMA superou os algoritmos RNA e SETAR, principalmente para horizontes mais curtos. Os resultados obtidos via RNA apontaram um *trade-off* entre o grau de pré-processamento e a qualidade das predições, as quais foram mais precisas na presença de não-linearidade nos dados. Devido aos diferentes padrões de conduta do consumidor em turismo, verificou-se que as predições de chegadas foram mais precisas do que as de dormidas de estrangeiros.

No trabalho de Zhang et al. (2014) foi concebido um estudo empírico que abrangeu quatro métodos de predição de séries temporais: regressão linear, suavização exponencial simples (SES), SARIMA e SVM. Empregaram-se nessa análise nove conjuntos de dados acerca de doenças infecciosas, os quais foram coletados por meio de um sistema nacional de vigilância em saúde pública na China. Os resultados obtidos, avaliados de acordo com as medidas baseadas puramente nos erros, demonstraram que, embora o algoritmo SVM tenha superado, na maioria dos casos, o desempenho dos modelos estatísticos, nenhum dos métodos averiguados foi significativamente melhor que os demais.

A ferramenta desenvolvida por Flauzino (2007) realiza a decomposição dos sinais elétricos em uma base resultante dos próprios sinais analisados e não em bases pré-determinadas. Essa técnica é chamada de Decomposição em Componentes Ortogonais (DCO). A DCO pode ser considerada uma nova ferramenta de processamento de sinais, pensada inicialmente para o uso em sistemas elétricos, de forma a auxiliar, principalmente, as tarefas de detecção de eventos, estimação de parâmetros e filtragem de sinais. Em Batista e al. (2016), Flauzino (2007), a DCO foi utilizada de forma promissora na detecção e localização de faltas de alta impedância em sistemas de distribuição de energia elétrica. Pode-se afirmar que a técnica é altamente sensível à operação do sistema em análise e resulta em sinais no domínio do tempo, permitindo a aplicação de outras ferramentas existentes de processamento de sinais para melhorar a extração de características.

Já Liboni (2017) se aprofundou no entendimento da técnica DCO e em conjunto utilizou RNA e SVM tanto para a classificação multi-classes quanto para a detecção de novidades em sinais elétricos provenientes de máquinas elétricas rotativas. Os parâmetros desses algoritmos foram configurados para receber índices advindos do processamento de sinais elétricos de motores, e a partir deles, identificar padrões normais e os padrões anômalos. Ainda, a severidade do defeito também foi diagnosticada, a qual é representada pelo número de barras quebradas no rotor. Os resultados alcançados demonstraram a eficácia das ferramentas utilizadas, sendo que os índices criados se mostraram altamente correlacionados com o fenômeno do defeito. Mais especificamente, foi possível criar índices

monotônicos com a severidade do defeito e com baixa variabilidade, demonstrando-se que as ferramentas são eficientes extratores de características.

As RNA, sistemas *fuzzy* ou *neuro-fuzzy* são amplamente utilizados para velocidade, estimativa de torque e controle de acionamento de estado sólido de máquinas de corrente alternada e contínua. São particularmente adequados para aplicações de máquinas onde as relações entre a corrente e a velocidade do motor não são lineares. Essas técnicas de inteligência artificial estão sendo estendidas como uma ferramenta de tomada de decisão para os resultados da MCSA para monitoramento de condições e detecção de falhas de máquinas (FILIPPETTI et al., 2000; VAS, 1999; FILIPPETTI et al., 1996; SCHOEN et al., 1995; PENMAN; YIN, 1994; FILIPPETTI; FRANCESCHINI; TASSONI, 1993).

Um sistema de diagnóstico de falhas baseado em rede neural utilizando os espectros de corrente do estator foi desenvolvido em Schoen et al. (1995). O pré-processador extrai os componentes de frequência dos dados amostrados. Usando filtros de frequência baseados em regras, essas componentes de frequência são classificados em quatro categorias com um nível decrescente de importância. Com base nessas regras, uma rede neural treinada para todas as condições operacionais possíveis da máquina é usada para classificar os dados recebidos. Uma assinatura espectral que fica fora dos *clusters* treinados é marcada como uma falha potencial do motor. Para evitar um diagnóstico falso, o pós-processador envia um alarme apenas quando as assinaturas de falha são observadas persistentemente. Essa função é realizada mantendo um histórico de tempo do motor monitorado. Esse esquema foi implementado com sucesso para diagnosticar defeitos nos mancais e nos rotores desequilibrados dos motores de indução. As referências Kim, Parlos e Bharadwaj (2003), Kim e Parlos (2002) descrevem um método de previsão de falhas baseado em rede neural que não requer nenhum parâmetro da máquina ou informações de velocidade. A velocidade é estimada a partir da tensão e corrente terminais medidas. Máquinas de indução de diferentes classificações de potência podem ser acomodadas usando o ajuste mínimo da rede neural. A eficácia de detecção de 93% ou maior foi alcançada.

Da mesma forma, sistemas baseados em lógica difusa foram usados por Vas (1999) para classificar falhas relacionadas a barras quebradas, categorizando os dois componentes de banda lateral em torno do fundamental da corrente da linha do motor de indução por um conjunto de nove regras. Denotando as bandas laterais como *e*, que são as duas entradas do sistema e o número de barras quebradas como a saída do sistema, um exemplo dessas regras é "Se *e* for pequeno e grande, *n* é igual a uma barra quebrada". A inferência nebulosa é obtida usando os métodos de mínimo e máximo de implicação nebulosa e a técnica de defuzzificação do centroide.

Dentre as medidas comumente empregadas para avaliar o desempenho e a qualidade da predição de um modelo de classificação, destacam-se o desempenho preditivo e a acurácia (HAN; KAMBER; PEI, 2011; LEE; GIRAUD-CARRIER, 2008).

O desempenho preditivo com que o classificador prediz a classe dos exemplos está

relacionado a sua representação interna utilizando os atributos. Se os atributos não são capazes de representar o conhecimento implícito nos exemplos, o desempenho preditivo do modelo de classificação é deteriorado, ou seja, haverá uma grande chance da classe de novos exemplos submetidos ao classificador ser predita incorretamente (PARMEZAN, 2012). Ainda, o número de atributos utilizados para representar os exemplos é um dos fatores que pode influenciar diretamente no desempenho preditivo do modelo induzido, haja vista que alguns algoritmos de aprendizado não trabalham bem na presença de grande quantidade de atributos (LIU; MOTODA, 2008; LEE; MONARD, 2003).

Problemas como os descritos podem dificultar a extração de conhecimento a partir de informações contidas em uma base de dados. Desse modo, é necessário que os dados sejam representados e processados de maneira apropriada e o modelo construído seja avaliado e validado (HAN; KAMBER; PEI, 2011; LEE, 2005; WITTEN; FRANK, 2005).

A realização de pesquisas quanto à avaliação de medidas de importância e métodos para a combinação de atributos é de grande importância para a determinação de abordagens que possam guiar a tarefa de seleção de atributos (SA). Consequentemente, é possível obter dados reduzidos para a construção de modelos com menor complexidade de tempo ou mais compreensíveis ao usuário final e com qualidade potencialmente superior ou semelhante a modelos obtidos considerando todos os atributos. Adicionalmente, em domínios com um elevado número de atributos, pode-se contribuir com a redução dos efeitos associados à maldição da dimensionalidade (PARMEZAN, 2012).

A busca realizada por um algoritmo de SA pode ser instanciada em relação a basicamente quatro questões: (1) sentido de busca, (2) estratégia de busca, (3) definição do critério de parada e de (4) critérios de avaliação de importância de atributos.

A primeira questão influencia na determinação do ponto de partida (ou pontos de partida) no espaço de busca, isto é, na direção em que a busca será realizada e os operadores que serão utilizados para acessar os estados sucessores. As abordagens para a direção da busca são divididas em (LIU; MOTODA, 2008) e caracterizam-se como: *Forward Selection*; *Backward Elimination*; *Bidirectional Search*; *Random Search*.

A respeito da segunda questão, existem três diferentes táticas para lidar com o problema da estratégia de busca sem recorrer à busca exaustiva: busca completa, busca heurística e busca não-determinística (LIU; YU, 2002; SCHLIMMER, 1993). Na primeira estratégia nenhum subconjunto ótimo de atributos é perdido, embora nem todo subconjunto de atributos tenha que ser avaliado. Já a segunda estratégia emprega algum tipo de heurística para conduzir a busca, ou seja, utiliza conhecimento específico do problema para encontrar uma solução de modo mais eficiente em relação à busca não informada¹. No entanto, há o risco de não encontrar os subconjuntos ótimos. A terceira estratégia, diferentemente das anteriores, procura pelos subconjuntos de atributos aleatoriamente.

¹A busca não informada, também denominada de busca cega, trabalha somente com as informações necessárias para distinguir o estado alvo de um estado não alvo.

Desse modo, o subconjunto corrente não aumenta ou diminui, diretamente, a partir de algum subconjunto anterior segundo uma regra determinística.

Em relação à terceira questão, a criação de subconjuntos de atributos que deterioram ou não melhoram o desempenho preditivo de classificação pode ser adotada como critério de parada da busca. Outra possibilidade, aplicável a determinados algoritmos de SA, inclui parar a busca quando uma quantidade de atributos definida *a priori* é atingida. Uma quantidade de iterações também pode ser estabelecida com o intuito de finalizar esse processo.

As possíveis combinações entre estratégias e direções de busca são resumidas na Tabela 1 (LIU; MOTODA, 1998).

Direções de Busca	Estratégias de Busca		
	Completa	Heurística	Não-determinística
<i>Forward Selection</i>	✓	✓	✗
<i>Backward Elimination</i>	✓	✓	✗
<i>Bidirectional Search</i>	✓	✓	✗
<i>Random Search</i>	✗	✓	✓

Tabela 1 – Combinações de estratégias e direções de busca (modificado de Liu e Motoda (1998)).

Uma das maneiras de realizar SA é por meio da minimização/maximização de múltiplas medidas de importância, como objetivos, para a obtenção de soluções aproximadamente ótimas (ABEEL et al., 2010; SPOLAÔR, 2010; WANG; HUANG, 2009; LIU; MOTODA, 2008; LEE; MONARD; WU, 2006). Nesse contexto, os algoritmos genéticos multiobjetivo tentam produzir, por aproximações sucessivas, as melhores soluções para um determinado problema. Com isso, é possível evitar a computação de todas as soluções admissíveis e prover auxílio para a otimização de medidas de importância. Embora esse tema não seja tratado neste trabalho, pesquisas demonstram o interesse crescente no uso de meta-heurísticas como os algoritmos genéticos multiobjetivo na SA monorrótulo (SPOLAÔR; LORENA; LEE, 2010).

Associados à quarta questão, os métodos de SA selecionam os atributos pela avaliação individual ou pela avaliação de subconjuntos de atributos.

No caso de avaliação individual, a qual também é denominada de ordenação ou *ranking*, os atributos são avaliados individualmente e ordenados de acordo com algum critério de importância. Desse modo, os atributos que estiverem melhor posicionados podem ser selecionados. Diferentemente, na avaliação de subconjunto ocorre a busca por subconjuntos mínimos de atributos segundo algum critério de importância.

Os métodos de SA por avaliação individual geralmente removem atributos pouco ou descorrelacionados com a classe (atributos irrelevantes), pois espera-se que atributos relevantes com valores correlacionados (atributos redundantes), tenham a mesma importância na discriminação de classes. Contudo, métodos que avaliam subconjuntos de atributos podem remover tanto atributos irrelevantes quanto redundantes (LIU; MOTODA, 2008).

Outra característica específica da SA baseada em avaliação individual é que um limiar é necessário para definir quais são os atributos a serem selecionados.

Os métodos que realizam a seleção de subconjuntos de atributos podem ser agrupados nas abordagens filtro, *wrapper* e embutida, dependendo de como o algoritmo de SA se relaciona com os algoritmos de aprendizado (extração de padrões) (COVÕES, 2010; PILA, 2001; KOHAVI; JOHN, 1997).

Com relação aos critérios de avaliação de importância de atributos, diversas medidas têm sido propostas na literatura para combinar e identificar os atributos mais significativos em um conjunto de dados (PARMEZAN, 2012; PARMEZAN; LEE; WU, 2011; PARMEZAN; WU; LEE, 2011; COVÕES, 2010; PARMEZAN et al., 2010; SPOLAÔR, 2010; LIU; MOTODA, 2008; LEE, 2005; LIU; MOTODA; YU, 2004; YU; LIU, 2004; DASH; LIU, 2003; GUYON; ELISSEEFF, 2003; LEE; MONARD, 2003; LIU et al., 2003; SOUSA et al., 2002; PILA, 2001; DASH; LIU, 2000; TRAINA et al., 2000; LIU; MOTODA, 1998).

Existem diferentes definições na literatura para determinar a importância de um atributo. Algumas definições consideram a importância do atributo em relação à classe, enquanto outras trabalham com a SA não supervisionada (SPOLAÔR; LORENA; LEE, 2011; NOGUEIRA, 2009; HE; CAI; NIYOGI, 2005; MITRA; MURTHY; PAL, 2002; DY; BRODLEY, 2000; TALAVERA, 1999). Essa necessidade de estimativa da importância de atributos é comum, tanto à avaliação individual quanto à avaliação de subconjuntos de atributos. Por meio das medidas de importância é possível, por exemplo, avaliar se os atributos selecionados auxiliam a melhorar o desempenho preditivo do classificador, ou auxiliam a simplificar o modelo construído de modo que ele seja mais compreensível.

As medidas de avaliação de importância de atributos podem ser organizadas em três principais categorias: clássica, a qual inclui medidas de informação, correlação e distância, consistência e precisão.

Quando um atributo é removido e a medida de importância aplicada nos dados representados pelos atributos restantes é deteriorada, pode-se considerar que o atributo é importante (LIU; MOTODA, 1998). Com base nessa definição, a seguir são descritos os principais tipos de medidas de importância que compõem as categorias apresentadas.

Medidas de Informação: permitem determinar o ganho de informação a partir de um atributo, isto é, determinar a diferença entre a incerteza *a priori* e a incerteza *a posteriori* associada ao atributo investigado (YU; LIU, 2004). Desse modo, um atributo pode ser definido como sendo mais importante que outro quando implica em um maior ganho de informação. Um exemplo usual de medida de informação é a entropia (MITRA; MURTHY; PAL, 2002);

Medidas de Correlação: também conhecidas como medidas de dependência ou associação. Permitem quantificar a habilidade de prever o valor de um atributo a partir do valor de outro atributo (HALL, 2000; HALL, 1999). Em termos de correlação

entre atributos, é possível estimar a redundância de um conjunto de dados com o auxílio de medidas como a dimensão fractal (LEE; MONARD; WU, 2006), a qual é aplicável a um conjunto de dados rotulados ou não rotulados. Uma medida clássica de dependência é o coeficiente de correlação, o qual pode, por exemplo, ser utilizado em um problema de classificação para encontrar a correlação entre um atributo e a classe;

Medidas de Distância: também denominadas de medidas de separabilidade, divergência ou discriminação. Em SA supervisionada, essas medidas consideram como atributos mais importantes aqueles que possibilitam a separação de exemplos de rótulos distintos e a não separação de exemplos da mesma classe (ROBNIK-SIKONJA; KONONENKO, 2003; KONONENKO, 1994). Dentre essas medidas, destacam-se às de distância interclasse e intraclasse, as quais estimam respectivamente a separabilidade existente entre as classes na projeção correspondente ao subconjunto de atributos investigado e a densidade dos exemplos dentro de uma classe específica;

Medidas de Consistência: possuem características diferentes em relação às outras medidas, uma vez que são fortemente dependentes do conjunto de treinamento. A utilização dessas medidas permite encontrar um subconjunto mínimo de atributos que satisfaz a proporção de inconsistência aceita, geralmente definida pelo usuário. O conceito de consistência corresponde, em dados rotulados, à não ocorrência de um ou mais pares de exemplos com valores idênticos em cada atributo e rótulos distintos (ARAUZO-AZOFRA; BENITEZ; CASTRO, 2008; DASH; LIU, 2003). Desse modo, o intuito da análise por consistência é possibilitar a construção de hipóteses lógicas consistentes em um conjunto de treinamento. É importante observar que as medidas de consistência não detectam a ocorrência de atributos redundantes, pois não possibilitam a distinção entre dois atributos igualmente bons;

Medidas de Precisão: referem-se a tarefas de predição. Dados um determinado algoritmo de aprendizado e os diversos subconjuntos de atributos, o que maior precisão (desempenho preditivo) proporcionar ao modelo gerado será selecionado (KOHAVI; JOHN, 1997). Sendo assim, é usual a utilização do mesmo algoritmo tanto para realizar a tarefa de SA quanto para processar o conjunto de exemplos com os atributos selecionados.

É importante ressaltar que as medidas de precisão são, geralmente, dependentes do algoritmo de aprendizado considerado, pois os subconjuntos de atributos selecionados são importantes em relação ao desempenho preditivo do modelo induzido por um determinado algoritmo. Diferentemente, as medidas de separabilidade de classes são independentes do algoritmo de aprendizado escolhido para a posterior construção do modelo.

Capítulo 3

Aspectos da Bancada Experimental e Delineamento Experimental

Neste capítulo serão apresentados os principais componentes da estrutura responsável pela aquisição dos dados experimentais que foram utilizados para o estudo e o desenvolvimento do sistema de identificação e de monitoramento de defeitos em MITs. Ainda, serão apresentados os parâmetros e as variáveis isoladas para a aquisição de dados no delineamento experimental e as condições de operação utilizadas para a realização dos experimentos.

3.1 Disposição Geral da Bancada de Experimentos

A bancada de experimentos se encontra nas dependências da Escola de Engenharia de São Carlos (EESC) da Universidade de São Paulo (USP), mais especificamente no Laboratório de Automação Inteligente de Processos e Sistemas (LAIPS) e Laboratório de Controle Inteligente de Máquinas Elétricas (LACIME). A Figura 4 ilustra a composição da estrutura dos laboratórios utilizados para a aquisição dos sinais, com alguns componentes detalhados a seguir:

1. **Bancada experimental:** O laboratório é composto por uma bancada experimental, na qual estão instalados o MIT, a Máquina de Corrente Contínua (MCC), o *encoder* e o torquímetro. A MCC atua como gerador e é acoplada ao MIT mediante um torquímetro rotativo, sendo responsável pela simulação experimental de uma determinada carga acoplada ao eixo do MIT;
2. **Circuito de alimentação do gerador de corrente contínua (GCC):** Consiste em um circuito de alimentação da bobina de campo do GCC, cuja carga é conectada à bobina de armadura. O torque resistente imposto pelo gerador pode ser alterado, variando-se a tensão do enrolamento de campo, que é conectada a um filtro e um

retificador monofásico. Alternativamente, tal ajuste pode ser realizado mediante a variação da carga acoplada no enrolamento de armadura;

3. **Quadro de comando:** O laboratório possui um quadro de comando no qual se realiza a escolha da ligação, estrela ou triângulo, a escolha do tipo de alimentação, partida direta ou inversor, assim como o ajuste de tensão aplicada no MIT;
4. **Circuitos de acondicionamento de sinais:** São os circuitos utilizados no acondicionamento dos sinais dos transdutores e que enviam os sinais tratados para a placa de aquisição de dados localizada em um computador.

Figura 4 – Esquema geral do laboratório de ensaios de máquinas elétricas.

A Figura 5-(a) ilustra o Variac de 1800W, enquanto que a Figura 5-(b) apresenta o circuito da fonte de tensão de alimentação do enrolamento de campo.

O MIT utilizado no laboratório caracteriza-se por ser um motor linha *standard* do fabricante WEG, 1,0cv, 220V/380V, 3,02A/1,75A, 4 polos, 60Hz, 4,1Nm e velocidade nominal de 1715rpm. O motor possui 58 espiras por bobina no estator e 6 bobinas por fase, totalizando 348 espiras por fase. A sua gaiola de esquilo é formada por 34 barras.

A MCC possui 2kW de potência, 190V de tensão nominal de campo e 250V de tensão nominal de armadura. O *encoder* óptico incremental apresenta uma resolução de 2000 pulsos por rotação, cuja finalidade consiste na medição da velocidade angular do MIT para auxiliar no monitoramento do estado de funcionamento do motor.

As correntes do GCC são adquiridas por meio do sensor Hall de corrente LAH-25NP do fabricante LEM, como na Figura 6-(a). O acondicionamento de sinal é realizado

Figura 5 – Excitação do enrolamento de campo do gerador de corrente contínua (GCC).

utilizando-se o amplificador operacional *TL081* com ajuste de *offset*. As tensões do GCC são adquiridas por intermédio do sensor Hall *LV – 20P* do fabricante LEM, mostrado na Figura 6-(b). Uma única placa de acondicionamento é adaptada para receber os sinais dos sensores Hall de corrente e tensão. Já as variáveis elétricas do MIT são mensuradas por meio de sensores Hall inseridos internamente ao quadro de comando.

Figura 6 – Sensores Hall de corrente e tensão.

O laboratório de pesquisa é equipado com um sistema de aquisição de dados composto por uma placa *National Instruments* modelo *NIDAQmxPCIE – 6259*, a qual dispõe de 32 canais analógicos de entrada de 16 bits, com taxa de aquisição de 1,25 MHz, 4 canais analógicos de saída de 16 bits, com taxa de 2,8 MHz, e 48 canais de I/O digitais com capacidade de operar na frequência de *clock* de até 10MHz. A placa é conectada no barramento *PCI-Express* do computador por um bloco conector *SCB – 68* que, além de prover a interface de conexões, permite também a implementação de filtros de sinais em hardware. O laboratório também conta com o oscilógrafo *YOKOGAWA DL750*, que é capaz de realizar aquisições com taxas de amostragem variadas.

Alguns rotores e motores foram preparados para simular as falhas elétricas e mecânicas dos MIT. Neste trabalho serão analisados defeitos de barra com acionamento direto pela rede, porém a bancada conta com a possibilidade de simular falhas de curto-circuito e rolamentos defeituosos.

Os sinais adquiridos, segundo o procedimento experimental, foram processados nos ambientes *Python* e *Weka*.

3.2 Inserção de Defeitos nas Barras dos Rotores

A inserção de defeitos nos rotores dos MIT consistiu na perfuração, com caminho radial, do rotor, por meio de uma broca de 6mm de diâmetro, que é suficientemente maior que a largura da barra da gaiola de esquilo. Primeiramente, foi necessário encontrar o centro do rotor para que o orifício gerado apresente um caminho radial. A Figura 7-(a) mostra a metodologia empregada para se encontrar o centro, iniciando-se em marcar no raio externo de uma das extremidades do eixo com uma broca. Após, mensura-se o diâmetro deste eixo e o da broca. Desta maneira, a posição do centro é exatamente o raio do eixo somado ao raio da broca. Esse deslocamento é efetuado de forma precisa, uma vez que a furadeira apresenta sensores de posição vertical e horizontal.

A Figura 7-(b) ilustra a imagem do momento em que o rotor estava sendo perfurado, e consequentemente seccionando-se uma das barras do rotor. Nota-se também que as barras são de alumínio, devido à composição dos cavacos da perfuração. Já a Figura 7-(c) ilustra com nitidez o orifício de seccionamento de uma das barras da gaiola de esquilo e a Figura 7-(d) ilustra a imagem de um rotor com quatro barras seccionadas de forma adjacente.

Figura 7 – Procedimento para inserção de defeitos de barra em rotores.

Sucessivas inserções de defeitos nas barras dos rotores foram realizadas a fim de se constituir um banco de dados das variáveis monitoradas.

3.3 Delineamento Experimental

Para o desenvolvimento do sistema de diagnóstico e monitoramento de defeitos em MITs foram realizados experimentos utilizando a bancada supracitada para a construção do banco de dados. Foram realizados ensaios em motores saudáveis e motores com defeitos em partida direta com tensão de alimentação trifásica equilibrada e frequência de 60 Hz.

Para a construção do banco de dados de maneira sólida, possibilitando também a realização de trabalhos futuros, foram aplicados carregamentos de 0,5 Nm, 1,0 Nm, 1,5 Nm, 2,0 Nm, 2,5 Nm, 3,0 Nm, 3,5 Nm e 4,0 Nm ao eixo do MIT. Para cada condição de carregamento do eixo do motor foram realizados 10 repetições.

Desta maneira, utilizando o sistema de aquisição de dados, para cada experimento de cada carregamento, foram adquiridas as seguintes variáveis: tensões nas fases A (V_a), B (V_b) e C (V_c); correntes nas fases A (I_a), B (I_b) e C (I_c); velocidades de vibração mecânica tangencial na carcaça (Vibcarc), tangencial na base (Vibase), axial no lado acionado (Vibaxial), radial no lado acionado (Vibacpi) e radial no lado não acionado (Vibacpe). A Figura 8 ilustra os pontos de coleta das variáveis de velocidade de vibração mecânica na bancada experimental.

Figura 8 – Representação dos pontos de coletas das variáveis de vibração mecânica.

Ainda, para a detecção e o diagnóstico de falhas, esse processo experimental foi realizado para motores saudáveis e para motores com rotores contendo 1, 2, 3 e 4 barras quebradas de forma adjacente.

Todos os dados obtidos foram publicados e disponibilizados no artigo Trembl et al. (2020) na base de dados da IEEE DataPort.

Capítulo 4

Metodologia

Diante da justificativa e relevância do tema, apresentadas no Capítulo 1 e dos aspectos de tendência das pesquisas sobre o diagnóstico e monitoramento de defeitos em máquinas de indução identificados pela revisão sistemática da literatura no Capítulo 2, empregou-se uma metodologia específica para essa temática. A metodologia é formada por um conjunto de estratégias para o desenvolvimento de um sistema de detecção, diagnóstico, classificação e monitoramento de defeitos em MITs. Na Figura 9 é apresentado um fluxograma com as atividades metodológicas assim como suas interconexões.

A Etapa 1 da metodologia, estado da arte, consistiu na busca de trabalhos com a temática de diagnóstico e monitoramento de sistemas sujeitos a falhas, mais especificamente MITs. Essa etapa foi realizada por meio da aplicação da revisão sistemática da literatura apresentada no Capítulo 2, e serviu de apoio para a determinação das tendências quanto a essa linha de pesquisa. Não obstante, foi possível identificar métodos, ferramentas e métricas para o desenvolvimento desta pesquisa.

Desta maneira, a Etapa 2 da metodologia, aquisição de dados, apresentou os aspectos da bancada experimental e também o delineamento experimental, como mostrado no Capítulo 3. Essa etapa resultou na aquisição dos dados para a utilização nos processos metodológicos subsequentes, como serão descritos.

4.1 Pré-Processamento e Extração de Atributos

Bancos de dados reais são altamente suscetíveis a ruídos, ausência de dados e inconsistências devido a seu grande volume, e sua provável origem de múltiplas fontes heterogêneas. Consequentemente, é possível que a qualidade desses dados esteja comprometida, e dados de baixa qualidade irão levar a resultados de análise por sistemas de inferência de baixa qualidade. Portanto, necessita-se de uma fase anterior, que consiste na preparação dos dados para que possam ser aplicados os algoritmos de mineração de dados e de aprendizado de máquinas.

Figura 9 – Fluxograma com a sequência lógica da metodologia aplicada neste trabalho.

Desta maneira, neste trabalho, houve o cuidado com a identificação de dados faltantes, a normalização de dados em função do carregamento, e também na implementação de um método para a identificação da janela de dados do regime transitório de operação dos MITs que serve de subsídio para o alcance do objetivo central dessa pesquisa.

4.1.1 Dados Incoerentes

Essa etapa de tratamento dos dados visa eliminar os problemas de registros incompletos, valores errados e dados inconsistentes, de modo que eles não influenciem no resultado dos algoritmos usados. A técnica usada nesta etapa vai desde a remoção do registro com

problemas, passando pela atribuição de valores padrões para não existir influência sobre a população dos dados em questão.

O método de inserção de dados faltantes baseou-se, primeiramente, na identificação de dados inconsistentes de acordo com a distribuição de probabilidades para cada situação delineada experimentalmente no banco de dados. De acordo com os dados coletados, em concordância com a Seção 3.3, os dados foram agrupados em termos da severidade de falha, ou condição do rotor. Em seguida, os dados foram agrupados novamente em termos do carregamento imposto ao eixo do MIT. Esse agrupamento dos dados pode ser observado na Tabela 2.

Desse modo, todos os dados foram percorridos automaticamente para a identificação de dados incoerentes. Para isso, foram construídos gráficos em função do tempo para cada variável adquirida pelo delineamento experimental, em função de cada carregamento e para cada severidade de defeito. Com isso, por meio do cálculo da correlação entre os gráficos de uma mesma condição experimental pode-se ter base para inferir se houve erros na coleta de alguma amostra.

No banco de dados construído neste trabalho, houve apenas uma leitura inconsistente devido a uma possível falha no transdutor de corrente elétrica acoplado à bancada. Essa leitura foi a nona medição para a condição de rotor sem defeito e com carregamento de 0,5 Nm. Na Figura 10 é apresentado o gráfico da variável corrente na fase B em função do tempo para os dados considerados inconsistentes em comparação com os dados de uma medição considerada consistente para as mesmas condições experimentais. Ainda, é possível observar na Figura 10-(b) a presença do *trigger* no início da aquisição dos dados, o que não é verificado na Figura 10-(a).

Após esse processo de identificação de dados incoerentes, foi utilizado um método estatístico para o preenchimento dessa medição sem que houvesse alterações na distribuição de probabilidades das amostras referentes ao rotor sem defeitos e com carregamento de 0,5 Nm. Dessa maneira, foi aplicado o teste de normalidade de Kolmogorov Smirnov para todas as medições de todas as variáveis para a condição de rotor sem defeitos e carregamento de 0,5 Nm. Após atestado que todas as amostras seguiam uma distribuição normal, foi possível calcular as médias das medições para a então substituição da nona medição, que tinha se mostrado incoerente. Todos os processos automáticos dessa etapa de pré-processamento foram realizados por meio da linguagem de programação *Python*.

Condição do rotor	Carregamento (Nm)	Medições para cada Variável
Rotor saudável	0,5	10
	1,0	10
	1,5	10
	2,0	10
	2,5	10
	3,0	10
	3,5	10
	4,0	10
Rotor com 1 barra quebrada	0,5	10
	1,0	10
	1,5	10
	2,0	10
	2,5	10
	3,0	10
	3,5	10
	4,0	10
Rotor com 2 barras quebradas	0,5	10
	1,0	10
	1,5	10
	2,0	10
	2,5	10
	3,0	10
	3,5	10
	4,0	10
Rotor com 3 barras quebradas	0,5	10
	1,0	10
	1,5	10
	2,0	10
	2,5	10
	3,0	10
	3,5	10
	4,0	10
Rotor com 4 barras quebradas	0,5	10
	1,0	10
	1,5	10
	2,0	10
	2,5	10
	3,0	10
	3,5	10
	4,0	10

Tabela 2 – Maneira com a qual os dados foram agrupados e organizados para as análises posteriores.

Figura 10 – Exemplo de identificação da existência de dados incoerentes para o sinal de corrente em função do tempo (a) com relação a dados que representam o comportamento real esperado (b) do sinal de corrente na fase B.

4.1.2 Normalização dos Dados em Função da Variação dos Carregamentos

O banco de dados construído para o desenvolvimento desta pesquisa foi composto por dados provenientes de diversas condições de carregamento no eixo dos MITs. Desta maneira, há a necessidade de normalizar o dados em função do torque de carregamento para todos os dados serem passíveis de comparações em termos dos atributos que forem extraídos. Ainda, essa transformação dos dados, é uma prática para evitar que os algoritmos de aprendizado fiquem enviesados para as variáveis com maior ordem de grandeza.

Desta maneira, criou-se fatores de multiplicação para os dados em função do carregamento imposto aos eixos dos MITs. Para isso, utilizou-se uma interpolação linear para gerar os fatores de multiplicação por meio da linguagem de programação *Python*. Na Tabela 3 são apresentados os fatores de multiplicação empregados para os dados em função do carregamento nos eixos dos MITs.

Carregamento (Nm)	Equação de normalização	Fator de multiplicação
0,5	$\frac{0,5-0}{4,0-0}$	$\frac{1}{8}$
1,0	$\frac{1,0-0}{4,0-0}$	$\frac{2}{8}$
1,5	$\frac{1,5-0}{4,0-0}$	$\frac{3}{8}$
2,0	$\frac{2,0-0}{4,0-0}$	$\frac{4}{8}$
2,5	$\frac{2,5-0}{4,0-0}$	$\frac{5}{8}$
3,0	$\frac{3,0-0}{4,0-0}$	$\frac{6}{8}$
3,5	$\frac{3,5-0}{4,0-0}$	$\frac{7}{8}$
4,0	$\frac{4,0-0}{4,0-0}$	$\frac{8}{8}$

Tabela 3 – Fatores de multiplicação para a normalização dos dados em função do carregamento imposto ao eixo dos motores de indução trifásicos (MITs).

4.1.3 Método para Identificação do Regime Transitório

Neste trabalho, a principal questão de interesse é o diagnóstico e o monitoramento de defeitos em MITs por meio da análise de sinais no regime transitório de funcionamento dessas máquinas. Para isso, foi implementado um método para a identificação automática do regime transitório na operação de MITs utilizando-se de técnicas de diferenciação numérica de primeira e de segunda ordem.

Esse método é constituído pelo processo de análise das séries temporais representativas do comportamento dinâmico de MITs por meio da análise das variáveis adquiridas no delineamento experimental. A Figura 11 apresenta o fluxograma representativo do método para identificação da região transitória na operação de um MIT.

A primeira fase do método proposto, inicia-se com a aplicação da derivada numérica de primeira ordem para cada ponto das séries temporais adquiridas experimentalmente. Posteriormente, ocorre a identificação das regiões da série onde a derivada de primeira ordem é constante e ou nula por meio de parâmetros estatísticos. Então, a derivada numérica de segunda ordem é aplicada a cada ponto da curva referente ao sinal analisado para constatar as regiões da série temporal onde verifica-se a troca de sinal algébrico ou a troca de concavidade, sendo os pontos onde as derivadas de primeira e de segunda ordem apresentam valor nulo.

Com a caracterização dos pontos em locais de mudança de concavidade e em regiões cujas inclinações são nulas, a série temporal é então delimitada automaticamente em:

Trigger da aquisição de dados: tensão de alimentação de 5V, caracterizado pelo momento inicial do experimento até o ponto onde a derivada de primeira ordem deixa de ser constante e nula, por meio de parâmetros estatísticos ao nível de 5% de significância, e antecedida pelo ponto de derivada de primeira ordem máxima;

Região do regime transitório: definida e localizada entre os pontos de derivada de primeira ordem nula, sendo o primeiro antecedido pelo ponto de máximo e o segundo

Figura 11 – Diagrama representativo do método para identificação do regime transitório implementado neste trabalho.

precedido pelo ponto de mínimo da derivada de primeira ordem;

Região do regime permanente: determinada entre o ponto de derivada de primeira ordem nula precedido pelo ponto de mínimo da derivada de primeira ordem e o fim do experimento de cada série temporal.

Dessa maneira, as etapas subsequentes da metodologia irão ser implementadas apenas para a região do regime transitório, sendo o objetivo central dessa pesquisa. Ainda, a implementação desse método de identificação do regime transitório foi realizada por meio da linguagem de programação *Python*. Outra informação relevante é a de que seguindo a metodologia de identificação do regime transitório nas séries temporais, as janelas de tempo são de diferentes tamanhos para cada sinal.

4.1.4 Atributos Extraídos

Os atributos de entrada são descritores de dados que podem ser usados para descrever uniformemente diferentes conjuntos de dados, com tamanhos distintos e diferentes distribuições de dados. Tais medidas de caracterização devem ter duas propriedades essenciais. Precisam primeiro, serem simples matematicamente e em segundo lugar, devem capturar informações úteis para descrever e diferenciar o comportamento experimental imposto.

A eficiência do diagnóstico e monitoramento de defeitos em MITs, assim como qualquer outro problema de aprendizado de máquina ou mineração de dados, depende diretamente

da descrição dos conjuntos de dados. Uma escolha inadequada de atributos pode comprometer todo o processo de aprendizado, gerando um modelo de recomendação inadequado para a finalidade desejada.

A pesquisa sobre caracterização de conjuntos de dados no geral enfocam três categorias de atributos, como medidas estatísticas, baseadas em informação e baseadas em complexidade. Ainda, para atingir o objetivo geral desta pesquisa, foram incluídas as medidas baseadas na teoria de modelos e análise de sistemas que englobam as medidas específicas ao escopo de análise de séries temporais. A Tabela 4 mostra as medidas identificadas como as mais relevantes para descreverem séries temporais e as extraídas dos conjuntos de dados para este trabalho. Como citado no Capítulo 2, existem alguns atributos que se mostram eficazes na descrição de séries temporais e que não foram abordados nessa pesquisa. Essa circunstância é dada pela escolha de atributos mais simples matematicamente e que em conjunto podem ser equivalentes a atributos mais complexos, como o MCSA. Ainda, a descrição dos atributos extraídos pode ser observada no Apêndice A.

Como no delineamento experimental foram adquiridos dados para 11 sinais, sendo eles 3 de corrente elétrica de fase, 3 de tensão de alimentação e 5 de velocidade de vibração mecânica, e cada medida identificada na Tabela 4 foi extraída para cada um destes sinais, totalizando 319 atributos. Vale ressaltar que todos os atributos extraídos são referentes a todas as condições experimentais, sendo oito condições de carregamento para as quatro condições de defeito e uma condição de não defeito, totalizando 400 instâncias experimentais. As medidas foram extraídas com rotinas desenvolvidas em linguagem de programação *Python* com o auxílio das bibliotecas *NumPy*, *SciPy*, *Pandas*, *StatsModels*, *Matplotlib*, *Seaborn*, *Plotly*, *Bokeh* e *Pydot*.

4.2 Seleção dos Atributos Significativos

Após a extração dos atributos considerados relevantes para essa pesquisa de acordo a literatura, é necessária a implementação da tarefa de seleção de atributos. Essa tarefa é de suma importância para a redução da dimensionalidade do problema e também para evitar o sobreajuste nas etapas de mineração de dados e de aprendizado de máquinas.

Desse modo, neste trabalho considerou-se quatro algoritmos de seleção de atributos, sendo estes identificados na revisão sistemática da literatura como os que apresentaram maior capacidade de generalização:

CBF - Filtro Baseado em Consistência: Este método gera subconjuntos de atributos e os avalia sobre o tamanho e a inconsistência relacionados à classe. O subconjunto selecionado é aquele que, dentro do número máximo de tentativas no máximo de tentativas, tenha o menor tamanho e a menor inconsistência em relação à classe. Em outras palavras, é considerado um algoritmo probabilístico caracterizado por avaliar subconjuntos de recursos de acordo com sua consistência em relação a um

Categoria	Medida
Medidas estatísticas	Média
	Desvio padrão
	Variância
	Mediana
	Valor mínimo
	Valor máximo
	1º quartil
	2º quartil
	3º quartil
	Curtose
	Assimetria
	Covariância
	Soma dos valores
	Moda
Medidas baseadas em informação	Entropia
	Energia do sinal
	Contraste do sinal
	Homogeneidade do sinal
	Momento de 3ª ordem do sinal
	Potência do sinal
	Razão do ruído
	Distorção harmônica total do sinal
	Ruído e taxa de distorção do sinal
	Ponto de intersecção de 3ª ordem do sinal
	Intervalo dinâmico livre do sinal
Medidas baseadas em complexidade	Dimensão fractal
Medidas baseadas na análise de sistemas	Valor RMS
	Traço da matriz simétrica do sinal
	Autovalor de 1ª ordem do sinal

Tabela 4 – Categorias e medidas extraídas como descritores dos conjuntos de dados.

determinado rótulo, ou classe. Geralmente, a pesquisa favorece pequenos subconjuntos de recursos que apresentam alta consistência com a classe. O CBF possui uma complexidade temporal de $O(N.M^2)$, sendo N o número de instâncias e M o número de atributos (YU; LIU, 2004; HALL, 1999);

CFS - Seleção de Atributos Baseada em Correlação: Este algoritmo considera a capacidade preditiva individual de cada atributo e o grau de correlação entre esses atributos, incluindo a classe. Assim, é possível selecionar subconjuntos de atributos que são formados por características relevantes e não redundantes. Em termos práticos, pesos de relevância são atribuídos a subconjuntos de atributos de acordo com as medidas de avaliação da separabilidade, como a medida de incerteza simétrica. O CFS tem uma complexidade temporal de $O(N.M^2)$ (MITRA; MURTHY; PAL, 2002; HALL, 2000);

InfoGain: Este algoritmo seleciona atributos através de avaliação individual. A ideia central é calcular o ganho de informação, com base na medida de entropia, para avaliar a relevância de um atributo relacionado à classe. Finalmente, uma ordem (classificação) desses recursos é gerada em ordem decrescente em relação à sua relevância para a classe. Os primeiros t atributos são escolhidos para compor o subconjunto selecionado de atributos importantes, em que t é definido de acordo com alguns critérios, por exemplo, uma porcentagem do número original de atributos. O InfoGain possui uma complexidade de tempo de $O(M.T)$, em que T é o tempo para calcular o ganho de informação (PARMEZAN; LEE; WU, 2017; ARAUZO-AZOFRA; BENITEZ; CASTRO, 2008; DASH; LIU, 2003);

ReliefF: Este método funciona por amostragem aleatória de exemplos do conjunto de dados e da localização do vizinho mais próximo da mesma classe e do vizinho mais próximo da classe oposta. Então, os valores dos vizinhos mais próximos são comparados com a classe de amostras e usados para atualizar os pesos de importância de cada elemento em relação à classe. Este processo é repetido m vezes, onde m é o número de vezes que o algoritmo procura exemplos no conjunto de dados. O algoritmo trabalha com atributos quantitativos e qualitativos, mas está limitado a problemas de duas classes. Por esse motivo, o algoritmo original foi estendido em seis variações, e a versão chamada ReliefF, que tem complexidade de tempo de $O(m.N.M)$, é capaz de lidar com problemas de conjuntos de dados incompletos, com ruídos e com várias classes (PARMEZAN; LEE; WU, 2017; ROBNIK-SIKONJA; KONONENKO, 2003; KONONENKO, 1994).

Desta maneira, os quatro algoritmos de seleção de atributos descritos foram aplicados aos atributos extraídos para dois casos. O primeiro caso é um problema binário de classificação, onde existem duas classes, a condição de defeito e de não defeito. Já a segunda, é um problema multiclasse, tendo a condição de não falha e as condições de severidade de falha, sendo de 1, 2, 3 ou 4 barras quebradas no rotor. Para tanto, a tarefa de seleção de atributos foi prototipada no *software Weka*¹, utilizando os pacotes próprios da ferramenta, e implementada em linguagem de programação *Python*.

Quanto à estratégia de busca dos subconjuntos de atributos, visando o melhor desempenho dos algoritmos de seleção de atributos, foi utilizada a estratégia *Best First* para os algoritmos CBF e CFS e *Ranker*, ou ordenação, para os algoritmos InfoGain e ReliefF. Ainda, a complexidade computacional desses algoritmos foi descrita para comparações entre os mesmos.

¹<https://www.cs.waikato.ac.nz/ml/weka/>

4.3 Modelos de Aprendizado de Máquina

Com os atributos extraídos do banco de dados e selecionados de maneira automática pelos algoritmos de seleção de atributos, foi então realizada a tarefa de modelagem computacional por meio de algoritmos de aprendizado de máquinas para a detecção e o diagnóstico de defeitos.

Neste trabalho, foram implementados modelos de cinco algoritmos de aprendizado de máquinas. Esses algoritmos foram identificados com grande capacidade de generalização e também com a possibilidade de realizar trabalhos embarcados, em trabalhos futuros. Seguem os algoritmos de aprendizado de máquina implementados nessa pesquisa e uma descrição:

NB - Naive Bayes: Consiste em um algoritmo de aprendizado do paradigma estatístico. Assume que os dados seguem uma distribuição normal e emprega um modelo probabilístico para encontrar uma aproximação do conceito induzido. O NB pode ser entendido como uma rede bayesiana especializada em duas suposições importantes: os recursos são igualmente importantes e também estatisticamente independentes, de modo que, quando a classe é conhecida, conhecer o valor de um recurso não adiciona nenhuma informações sobre o valor de outro. Em termos práticos, o algoritmo classifica um novo exemplo de acordo com a classe mais provável, dado seu conjunto de dados. Esse algoritmo possui complexidade de tempo $O(2^N)$ (PARMEZAN; LEE; WU, 2017; ROBNIK-SIKONJA; KONONENKO, 2003; KONONENKO, 1994);

C4.5: Algoritmo para indução de árvores de decisão a partir do paradigma simbólico. O método usa a estratégia de dividir para conquistar, ou seja, um problema é decomposto em subproblemas mais simples e essa estratégia é usada recursivamente em cada subproblema decomposto. A árvore é uma disjunção de conjunções, na qual cada ramificação da árvore, da raiz à folha, é uma conjunção de condições sobre os recursos e o conjunto de galhos da árvore é separado. Portanto, cada nó da árvore especifica um teste em algum recurso do exemplo e em cada ramificação do nó um dos possíveis valores do recurso. C4.5 não assume nenhuma distribuição específica para o dados e permite que uma decisão complexa (predizer o valor da classe) seja decomposta em uma série de decisões elementares. Esse algoritmo possui complexidade de tempo $O(N \cdot \log(N)^2)$ (MITRA; MURTHY; PAL, 2002; HALL, 2000);

RD - Random Forest: Consiste em um conjunto de árvores de classificação ou regressão não podadas, criado usando amostras de autoinicialização dos dados de treinamento e seleção aleatória de recursos na árvore de indução. A previsão é feita agregando (voto majoritário ou média) as previsões do conjunto. Esse algoritmo possui complexidade de tempo $O(N \cdot t)$ (COVÕES, 2010; PILA, 2001; KOHAVI; JOHN, 1997);

MLP - *Multilayer Perceptron*: Constitui um tipo de rede neural artificial pertencente ao paradigma conexionista, inspirado no estudo das conexões neurais do sistema nervoso e baseia-se na combinação de unidades simples altamente conectadas. Nesse modelo, há uma ou mais camadas de neurônios entre as camadas de entrada de dados e saída de resultados. Estas camadas intermediárias são unidades que não interagem diretamente com o ambiente e funcionam como combinadores de características. Se houver conexões apropriadas entre as unidades de entrada e um conjunto considerável de unidades intermediárias, sempre é possível encontrar a representação que produzirá o mapeamento correto entre entrada de dados e saída de resultados (classificação). Esse algoritmo possui complexidade de tempo $O(N.W)$, sendo W o número de camadas da rede (PARMEZAN; LEE; WU, 2017; PAWLAK; KOWALSKI, 2010);

SVM - *Support Vector Machine*: Embora os modelos SVM apresentem uma estrutura semelhante a redes neurais, eles diferem na maneira como o aprendizado é conduzido. Enquanto as redes neurais trabalham por minimizar o risco empírico, ou seja, a minimização de erros do modelo induzido no treinamento, o SVM se baseia no princípio de minimizar o risco estrutural, que busca o menor erro de treinamento enquanto minimiza um limite superior no erro de generalização do modelo (erro de modelo quando aplicado aos dados de teste). O conceito de generalização é melhor compreendido no caso de classificação binária. Assim, dadas duas classes e um conjunto de pontos pertencentes a elas, o SVM determina o hiperplano que os separa, para colocar o maior número possível de pontos da mesma classe do mesmo lado, enquanto a distância de cada classe para essa superfície de decisão é maximizada. Nas situações em que as amostras não são linearmente separáveis, a solução concentra-se em mapear os dados de entrada para um espaço de maior dimensão (espaço de recurso). Esse mapeamento é alcançado pelo uso de uma função kernel. Esse algoritmo possui complexidade de tempo $O(N^3)$ (PARMEZAN; LEE; WU, 2017).

Os cinco algoritmos de aprendizado foram aplicados aos atributos selecionados por meio dos algoritmos de seleção de atributos do banco de dados no regime transitório para todas as condições experimentais. O treinamento dos modelos foi realizado de modo supervisionado para a criação de modelos preditivos para as situações binárias e multiclasse de classificação. O resultado do teste dos modelos baseou-se na classificação de novas instâncias comparando com o modelo para a determinação de sua classe. A distribuição de instâncias para o treino do modelo foi de 50%, 25% para a validação e 25% para o teste. Não obstante, a tarefa de treino do modelo foi realizada utilizando validação cruzada das instâncias com 10 partições para cada subconjunto.

Após a criação dos modelos de aprendizado, todos foram comparados entre si por meio de métricas de desempenho, gerando uma medida geral multicritério, como será descrita

na seguinte seção.

4.4 Avaliação de Desempenho

A avaliação do desempenho dos modelos de aprendizado construídos baseou-se em uma medida multicritério (PARMEZAN; LEE; WU, 2017). Como o primeiro passo para aplicar essa medida, é necessário escolher os índices de desempenho a serem considerados simultaneamente. De acordo com trabalhos identificados por meio da revisão sistemática da literatura, foram determinadas o uso de seis medidas:

1. Taxa de falsos positivos (TFP), ou alarmes falsos, de um classificador de conjunto induzido a partir dos subconjuntos de atributos selecionados pelos algoritmos de seleção de atributos, sendo conhecido também como erros do tipo I;
2. Taxa de falsos negativos (TFN) de um classificador de conjunto induzido a partir dos subconjuntos de atributos selecionados pelos algoritmos de seleção de atributos, sendo conhecido também como erros do tipo II;
3. Taxa de erro (TE) de um classificador de conjunto induzido a partir dos subconjuntos de atributos selecionados pelos algoritmos de seleção de atributos. É caracterizada como a medida contrária à acurácia do modelo;
4. Tempo de aprendizado (TA), em segundos normalizados, necessário para a construção do modelo de conjunto usando os subconjuntos de atributos, baseado na complexidade computacional dos algoritmos de seleção de atributos e de aprendizado de máquinas;
5. Porcentagem de atributos selecionados (PAS) por algoritmos de seleção de atributos;
6. Área complementar (ACROC) à área sob a curva ROC (*Receiver operating characteristics*).

Após configurar a medida multicritério com essas variáveis, é possível obter uma única medida para estimar o desempenho preditivo de cada configuração de modelo de aprendizado avaliado. Assim, a medida considera a área total de um hexágono irregular, que por sua vez é definido a partir dos valores alcançados pelas configurações do modelo de aprendizado de máquina nas seis variáveis. Quanto menor a área total, melhor o método empregado para a construção do modelo de aprendizado correspondente.

A Figura 12 mostra os hexágonos irregulares em relação a um modelo de decisão construído a partir de subconjuntos de atributos, cada um resultante de um algoritmo de aprendizado de máquina.

Figura 12 – Medida de desempenho multicritério.

Vale ressaltar que as medidas de desempenho referentes aos vértices do hexágono multicritério são normalizadas para o intervalo $[0, 1]$ para que a área interna dos hexágonos sejam passíveis de comparações.

4.5 Análise e Verificação da Robustez dos Sistemas Propostos

A análise e a verificação da robustez dos sistemas inteligentes propostos possui o objetivo de identificar o quão generalista é a configuração dos algoritmos de seleção de atributos em conjunto com os modelos de aprendizado de máquinas de acordo com os atributos de entrada (GUO et al., 2020; CHOQUEUSE et al., 2019; YAN et al., 2019; MAITRE et al., 2017; ZBEDE; GADOUE; ATKINSON, 2016; GUNER; USTUNDAG, 2015). Desse modo, essa análise mostra a maneira com a qual a configuração de métodos, de seleção de atributos e de aprendizado de máquinas, se comporta quando são inseridos ruídos nos sinais de entrada.

O protocolo de avaliação da robustez dos sistemas propostos é mostrado pelos fluxogramas das Figuras 13 e 14. Esse processo foi realizado em duas frentes, sendo a primeira de verificação e aferição da robustez dos modelos induzidos, Figura 13, para a identificação do comportamento dos modelos ao serem construídos utilizando-se sinais corrompidos por ruídos e testados também por sinais adicionados de ruídos. E, a segunda de análise da robustez, Figura 14, para a determinação do comportamento dos modelos induzidos a partir de sinais experimentais e testados com sinais corrompidos com ruídos.

Figura 13 – Fluxograma referente à metodologia empregada para a verificação da robustez dos sistemas propostos.

O processo de análise da robustez dos sistemas induzidos é caracterizado pelo teste dos modelos induzidos por meio dos dados experimentais, com sinais corrompidos com ruídos brancos. Já o processo de verificação da robustez foi definido com a indução de novos modelos por meio dos sinais corrompidos com ruídos brancos e o teste desses modelos com os sinais também corrompidos.

O processo de verificação da robustez dos modelos induzidos para a identificação da capacidade de generalização do conjunto de técnicas inteligentes é descrito a seguir em ordem cronológica de execução:

Inserção de Ruídos nos Sinais de Entrada: Foram inseridos ruídos brancos em todos os sinais coletados por meio do delineamento experimental descrito na Seção 3.3. Para a verificação da robustez dos sistemas induzidos, foram incorporados ruídos brancos com intensidade relativa à energia do sinal variando de zero a 100%, com passo de 5%, para todos os sinais coletados experimentalmente. Lembrando que os

Figura 14 – Fluxograma referente à metodologia empregada para a análise de robustez dos sistemas propostos.

sinais contendo 0% de ruído branco são os próprios sinais coletados pelo sistema de aquisição de dados;

Extração de Atributos: Os atributos descritos na Seção 4.1.4 foram novamente extraídos para todos os sinais compostos pelos sinais de entrada em adição aos ruídos brancos de intensidades relativas de zero a 100%, com passo de 5%, para a verificação da robustez. Cada intensidade de ruído branco adicionada ao sinal coletado experimentalmente gerou um conjunto de atributos com as respectivas classes. Dessa maneira, gerou-se 20 conjuntos de atributos, um conjunto para cada percentual de ruído branco adicionado ao sinal de entrada, variando de zero a 100%, com intervalos de 5%;

Seleção de Atributos: Para cada conjunto de atributos extraídos dos sinais compostos pelo sinal coletado experimentalmente com adição dos ruídos brancos de diferentes intensidades relativas, foram aplicadas as quatro técnicas de seleção de atributos

descritas na Seção 4.2. Assim, obteve-se os subconjuntos de atributos mais representativos dos conjuntos de dados, de acordo com algum critério inerente ao algoritmo de seleção de atributos utilizado, para cada cenário em função do percentual de ruído branco adicionado ao sinal coletado experimentalmente. Vale ressaltar que o processo de seleção de atributos foi realizado para dois contextos, de classificação binária e multiclasse;

Modelagem dos Algoritmos de Aprendizado de Máquinas: Por meio dos subconjuntos de atributos selecionados por cada algoritmo de seleção de atributos foram aplicados os cinco algoritmos de aprendizado de máquinas descritos na Seção 4.3, para a criação de modelos de classificação. Ressaltando que esse processamento foi realizado para os casos de classificação binária e multiclasse. A distribuição de instâncias para a o treino do modelo foi de 50%, 25% para a validação e 25% para o teste, e a tarefa de treino do modelo foi realizada utilizando validação cruzada das instâncias com 10 partições para cada subconjunto. Assim, para cada modelo induzido e testado foi possível obter as medidas de desempenho constituintes da medida multicritério, descritas na Seção 4.4;

Verificação da Robustez dos Modelos Induzidos: Com as medidas de avaliação do desempenho dos modelos induzido para cada percentual relativo de ruído branco adicionado aos sinais coletados experimentalmente, foi possível construir gráficos bidimensionais relacionando a porcentagem de ruído branco com o parâmetro de avaliação da configuração de técnicas utilizadas para a tarefa de classificação. Frente a isso, esses gráficos para a verificação da robustez dos modelos induzido apresentam no eixo das ordenadas o parâmetro de avaliação e o eixo das abcissas é referente ao percentual de ruído branco adicionado no sinal coletado experimentalmente. Os parâmetros de avaliação dos modelos induzidos utilizados foram os descritos para a medida multicritério e o próprio hexágono multicritério, citados na Seção 4.4. Com os gráficos de verificação de robustez concebidos, é possível identificar as regiões nas quais os parâmetros de avaliação são constantes, deixam de ser constante, apresentam comportamento crescente ou decrescente, ou seja, qualquer comportamento que diz respeito à estabilidade das técnicas utilizadas quando há a existência de ruídos nos sinais de entrada.

De mesmo modo, o processo de análise da robustez para os modelos criados especificamente para a detecção e o diagnóstico de defeitos em MITs é descrito a seguir de maneira ordenada:

Inserção de Ruídos nos Sinais de Entrada: Foram inseridos ruídos brancos em todos os sinais coletados por meio do delineamento experimental descrito na Seção 3.3.

Para a análise da robustez dos sistemas induzidos, foram incorporados ruídos brancos com intensidade relativa à energia do sinal variando de zero a 100%, com passo de 5%, para todos os sinais coletados experimentalmente. Lembrando que os sinais contendo 0% de ruído branco são os próprios sinais coletados pelo sistema de aquisição de dados;

Extração de Atributos: Os atributos descritos na Seção 4.1.4 foram novamente extraídos para todos os sinais compostos pelos sinais de entrada em adição aos ruídos brancos de intensidades relativas de zero a 100%, com passo de 5%, para a análise da robustez. Cada intensidade de ruído branco adicionada ao sinal coletado experimentalmente gerou um conjunto de atributos com as respectivas classes. Dessa maneira, gerou-se 20 conjuntos de atributos, um conjunto para cada percentual de ruído branco adicionado ao sinal de entrada, variando de zero a 100%, com intervalos de 5%;

Seleção de Atributos: Para o conjunto de atributos extraídos dos sinais coletados experimentalmente, com 0% de ruído, foram aplicadas as quatro técnicas de seleção de atributos descritas na Seção 4.2. Assim, obteve-se os subconjuntos de atributos mais representativos dos conjuntos de dados. Vale ressaltar que o processo de seleção de atributos foi realizado para dois contextos, de classificação binária e multiclasse.

Modelagem dos Algoritmos de Aprendizado de Máquinas: Por meio dos subconjuntos de atributos selecionados por cada algoritmo de seleção de atributos foram aplicados os cinco algoritmos de aprendizado de máquinas descritos na Seção 4.3, para a criação de modelos de classificação. Ressaltando que esse processamento foi realizado para os casos de classificação binária e multiclasse. A distribuição de instâncias para a o treino do modelo foi de 50%, 25% para a validação e 25% para o teste, e a tarefa de treino do modelo foi realizada utilizando validação cruzada das instâncias com 10 partições para cada subconjunto. Assim, para cada modelo induzido e testado foi possível obter as medidas de desempenho constituintes da medida multicritério, descritas na Seção 4.4;

Análise da Robustez dos Modelos Induzidos: Os modelos induzidos por meio dos sinais experimentais foram testados com os atributos corrompidos para cada nível de ruído branco adicionado aos próprios sinais experimentais. Com as medidas de avaliação do desempenho dos modelos induzido para cada intensidade relativa de ruído branco adicionado aos sinais coletados experimentalmente, foi possível construir gráficos bidimensionais relacionando a porcentagem de ruído branco com o parâmetro de avaliação da configuração de técnicas utilizadas para a tarefa de classificação. Frente a isso, esses gráficos para a análise da robustez dos modelos induzido apresentam no eixo das ordenadas o parâmetro de avaliação e o eixo das abcissas é referente

ao percentual de ruído branco adicionado no sinal coletado experimentalmente. Os parâmetros de avaliação dos modelos induzidos utilizados foram os descritos para a medida multicritério e o próprio hexágono multicritério, citados na Seção 4.4. Com os gráficos de análise de robustez concebidos, é possível identificar as regiões nas quais os parâmetros de avaliação são constantes, deixam de ser constante, apresentam comportamento crescente ou decrescente, ou seja, qualquer comportamento que diz respeito à estabilidade das técnicas utilizadas quando há a existência de ruídos nos sinais de entrada.

Para apoiar a decisão sobre o desempenho dos modelos induzidos quanto à robustez ao serem inseridos ruídos nos sinais de entrada, pode ser utilizado o parâmetro área sob as curvas analisadas. Nesse contexto, o parâmetro área sob a curva é empregado para fazer o posicionamento hierárquico dos modelos induzidos quanto à robustez. Desse modo, calculou-se a área sob as curvas de área do hexágono multicritério em função das porcentagens de ruído branco adicionados ao sinal de entrada. Assim, quanto menor a área sob a curva, melhor o desempenho do modelo induzido ao se corromper os dados de entrada por meio de ruídos. Para o cálculo da área sob as curvas, ou integral numérica, foi utilizada a regra $\frac{1}{3}$ de Simpson (JIANG; CHEN; MAO, 2007).

A constatação da robustez dos modelos induzidos foi realizada de duas maneiras. A análise da robustez possui o objetivo de determinar o quão generalista é o modelo ao ser testado com sinais corrompidos. Já a verificação da robustez apresenta o intuito de constatar o quanto um sistema, composto por um seletor de atributos e um classificador, se comporta na presença de ruídos desde a etapa do pré-processamento. Assim, a verificação da robustez tem como hipótese identificar as características intrínsecas dos algoritmos utilizados, enquanto que a análise de robustez confere as características do modelo construído para a aplicação específica de detecção e diagnóstico de defeitos em MITs. O confronto entre esses processos pode gerar base para uma recomendação de sistemas inteligentes mais assertivos e generalistas.

4.6 Sistema de Diagnóstico e Monitoramento de Defeitos em MITs

O sistema de detecção e de diagnóstico definitivo é apresentado de modo esquemático na Figura 15, sendo implementado com os dados advindos da bancada experimental.

Para a validação do sistema, foram ensaiados motores com rotores saudáveis e com falhas de barras quebradas, a fim de constituir um estudo de caso para este tipo de falha. Os motores foram acionados diretamente pela rede em diversas situações de carga no eixo constante.

Figura 15 – Sistema de detecção e de diagnóstico proposto.

O sistema monitora as tensões, as correntes e as velocidades de vibrações mecânicas e então processa esses sinais por meio dos atributos e ferramentas computacionais apresentadas anteriormente. A tarefa de seleção de atributos é realizada para a redução da dimensionalidade dos atributos, para a obtenção de subconjuntos de atributos relevantes de índices específicos quanto às falhas e para evitar o sobreajuste. A detecção de defeitos

é realizada pela implementação de classificadores. Ainda, as abordagens e estratégias de seleção de atributos são comparadas em função da acurácia na classificação das falhas. A classificação ocorre em duas frentes, sendo primeiramente para motores saudáveis e falhos e em seguida para a distinção da severidade da falha, representada pela quantidade de barras quebradas.

Na Tabela 5 são apresentados os códigos identificadores para os sistemas inteligentes construídos para facilitar a organização dos resultados.

Tipo de análise	Alg. de SA	Alg. de AM	ID. do sistema inteligente
(I) Binária	(a) CBF	(1) NB	Ia1
		(2) C4.5	Ia2
		(3) RD	Ia3
		(4) MLP	Ia4
		(5) SVM	Ia5
	(b) CFS	(1) NB	Ib1
		(2) C4.5	Ib2
		(3) RD	Ib3
		(4) MLP	Ib4
		(5) SVM	Ib5
	(c) InfoGain	(1) NB	Ic1
		(2) C4.5	Ic2
		(3) RD	Ic3
		(4) MLP	Ic4
		(5) SVM	Ic5
	(d) ReliefF	(1) NB	Id1
		(2) C4.5	Id2
		(3) RD	Id3
		(4) MLP	Id4
		(5) SVM	Id5
(II) Multiclasse	(a) CBF	(1) NB	IIa1
		(2) C4.5	IIa2
		(3) RD	IIa3
		(4) MLP	IIa4
		(5) SVM	IIa5
	(b) CFS	(1) NB	IIb1
		(2) C4.5	IIb2
		(3) RD	IIb3
		(4) MLP	IIb4
		(5) SVM	IIb5
	(c) InfoGain	(1) NB	IIc1
		(2) C4.5	IIc2
		(3) RD	IIc3
		(4) MLP	IIc4
		(5) SVM	IIc5
	(d) ReliefF	(1) NB	IId1
		(2) C4.5	IId2
		(3) RD	IId3
		(4) MLP	IId4
		(5) SVM	IId5

Tabela 5 – Organização das configurações de algoritmos de seleção de atributos (SA) e de aprendizado de máquinas (AM) dos sistemas inteligentes para diagnóstico por meio de códigos identificadores (IDs).

Capítulo 5

Resultados e Discussões

Neste capítulo são apresentados os resultados obtidos pela aplicação da metodologia descrita no Capítulo 4 e as discussões com comparações metodológicas e resultados já concebidos.

Conforme mencionado na metodologia, para o banco de dados contendo as variáveis de tensão e corrente no estator e velocidades de vibração mecânica para a condição saudável e de defeitos, inclusive de severidade de defeitos, foi realizada a seleção de atributos utilizando os algoritmos CBF, CFS, InfoGain e ReliefF, para cada estratégia de análise, binária ou multiclasse. Modelos de aprendizado foram gerados considerando os atributos selecionados por cada algoritmo de seleção de atributos para cada um dos algoritmos Naive Bayes, C4.5, Random Forest, Multilayer Perceptron e Support Vector Machine. Assim, foram totalizadas 40 configurações, sendo para um conjunto de dados, quatro algoritmos de seleção de atributos, cinco algoritmos de aprendizado de máquinas e duas abordagens, binária e multiclasse.

O desempenho preditivo dos classificadores induzidos foram estimados por meio de validação cruzada com 10 partições e comparados usando o teste estatístico não paramétrico *Kruskal-Wallis*, com nível de significância de 5%, e pós-teste de *Dunn*. Ainda, para atestar o desempenho dos sistemas de diagnóstico, foi proposta também uma métrica multicritério, dependendo de seis variáveis da construção do classificador e do seletor de atributos.

Os resultados obtidos, apresentados a seguir, estão organizados do seguinte modo:

- ❑ Banco de dados experimental construído por meio da bancada experimental;
- ❑ Atributos selecionados pelos algoritmos utilizados neste trabalho e o impacto das categoria nos atributos selecionados;
- ❑ Desempenho preditivo dos sistemas de diagnóstico induzidos e análise da significância estatística dos resultados;
- ❑ Verificação e análise da robustez dos sistemas de inferência induzidos;

- Recomendações para o sistema de diagnóstico.

5.1 Banco de Dados Experimental

O banco de dados construído para o desenvolvimento desse trabalho foi composto por ensaios experimentais de MITs saudáveis e contendo defeitos com diferentes severidades. O acionamento foi realizado com partida direta pela rede e com motores de 1,0cv, 220V/380V, 3,02A/1,75A, 4 polos, 60Hz, 4,1Nm e velocidade nominal de 1715rpm provenientes da bancada experimental e do delineamento experimental descritos no Capítulo 3.

Desta maneira, o banco de dados possui 400 exemplos, sendo que foram feitos 10 ensaios experimentais para cada configuração de carregamento do eixo e para motores saudáveis e também com severidades das falhas. Assim, o banco de dados se mostrou robusto para as condições avaliadas, se mostrando balanceado para as análises de diagnóstico de defeito e adaptável para a situação de detecção de defeito.

Ainda, para a análise do sistema de detecção de defeitos, o sistema de diagnóstico binário, o banco de dados foi separado em duas classes, motor saudável e motor com rotor contendo o defeito. Para o caso de diagnóstico da severidade de falhas, ou análise multiclasse, o banco de dados foi discretizado em cinco classes de acordo com as condições experimentais, motor saudável, motor com rotor com 1, 2, 3 e 4 barras quebradas.

5.2 Análise dos Atributos Seleccionados

Além das diversas características próprias do conjunto de dados, a maneira de avaliação dos atributos e a medida utilizada para determinar a importância dos atributos em relação à classe podem influenciar no subconjunto de atributos selecionado pelos algoritmos.

Nas Tabelas 6 e 7 são apresentados os atributos selecionados por cada um dos algoritmos de SA com a observação da estratégia de busca utilizada para os cenários binário e multiclasse, respectivamente, e os atributos selecionados são apresentados ordenados quanto à relevância com relação à classe. Ainda, lembrando que foram extraídos um total de 319 atributos do conjunto de dados.

Algoritmos de SA	Atributos seleccionados	Observações
CBF	Quartildois_Vibaxial	Estratégia Best First de busca
	Media_Vibaxial	
CFS	Desviopadrao_Vibacpe	Estratégia Best First de busca
	Mediana_Ib	
	Quartilum_Vb	
	Quartildois_Vibaxial	
	Quartiltres_Vibaxial	
	Quartiltres_Vb	
	Curtose_Va	
	Curtose_Vb	
	Assimetria_Vb	
	Moda_Ia	
	Moda_Ic	
	Razaoruido_Ib	
	Distorcaoharmonica_Va	
	Distorcaoharmonica_Vb	
	Ruidotaxadistorcao_Ic	
	Intervalodinamico_Va	
	Intervalodinamico_Vb	
	Dimensaofractal_Vibase	
InfoGain	Curtose_Va	Estratégia Ranker de busca
	Quartiltres_Vibaxial	
	Quartildois_Vibaxial	
	Distorcaoharmonica_Va	
	Curtose_Vb	
	Distorcaoharmonica_Vb	
	Moda_Ic	
	Razaoruido_Ib	
	Ruidotaxadistorcao_Va	
	Razaoruido_Ic	
	Intervalodinamico_Va	
	Razaoruido_Ia	
	Razaoruido_Va	
	Quartildois_Ia	
	Mediana_Ia	
	Quartiltres_Vb	
	Quartilum_Vb	
	Dimensaofractal_Vibase	
ReliefF	Mediana_Vb	Estratégia Ranker de busca
	Quartiltres_Vibaxial	
	Distorcaoharmonica_Va	
	Quartildois_Vibaxial	
	Distorcaoharmonica_Vb	
	Ruidotaxadistorcao_Va	
	Intervalodinamico_Va	
	Razaoruido_Ib	

Tabela 6 – Atributos seleccionados por cada um dos algoritmos de seleção de atributos (SA) para o caso de classificação binária.

Algoritmos de SA	Atributos selecionados	Observações
CBF	Desviopadrazo_Va	Estratégia Best First de busca
	Desviopadrazo_Ib	
	Mediana_Vibaxial	
	Quartiltres_Vibacpe	
	Quartiltres_Vb	
	Razaoruido_Va	
	Dimensaofractal_Vibacpe	
CFS	Desviopadrazo_Vibacpe	Estratégia Best First de busca
	Desviopadrazo_Ia	
	Variancia_Vibacpe	
	Mediana_Vibaxial	
	Mediana_Vibacpi	
	Mediana_Ib	
	Minimo_Ib	
	Maximo_Ib	
	Quartilum_Vibaxial	
	Quartilum_Vibcarc	
	Quartilum_Vibacpe	
	Quartilum_Va	
	Quartilum_Vb	
	Quartildois_Vibaxial	
	Quartiltres_Vibaxial	
	Quartiltres_Vibacpe	
	Quartiltres_Vb	
	Assimetria_Vibacpe	
	Moda_Vb	
	Moda_Ia	
	Moda_Ib	
	Moda_Ic	
	Entropia_Vibaxial	
	Entropia_Vibcarc	
	Razaoruido_Va	
	Distorcaoharmonica_Vibcarc	
InfoGain	Ruidotaxadistorcao_Va	Estratégia Ranker de busca
	Razaoruido_Va	
	Quartildois_Vibaxial	
	Quartiltres_Vibaxial	
	Intervalodinamico_Va	
	Distorcaoharmonica_Va	
	Distorcaoharmonica_Vb	
	Ruidotaxadistorcao_Vb	
	Moda_Ic	
	Ruidotaxadistorcao_Vc	
	Quartiltres_Vb	
	Intervalodinamico_Vc	
	Curtose_Va	
	Razaoruido_Vb	
	Curtose_Vb	
ReliefF	Quartiltres_Vibaxial	Estratégia Ranker de busca
	Quartildois_Vibaxial	
	Ruidotaxadistorcao_Va	
	Distorcaoharmonica_Vb	
	Distorcaoharmonica_Va	
	Intervalodinamico_Va	
	Pontointerseccao_Vibase	

Tabela 7 – Atributos selecionados por cada um dos algoritmos de seleção de atributos (SA) para o caso de classificação multiclasse.

Considerando os percentuais de atributos selecionados por cada método de SA, no caso binário, o algoritmo *InfoGain* foi o que selecionou os subconjuntos com mais atributos, variando de 7,52% até o máximo de 100,00%, ou todos os atributos.

Para o cenário de seleção de atributos contendo apenas duas classes, a categoria de medidas estatísticas foi a mais selecionada. Para o algoritmo *CBF* foram 100,00%, para o *CFS* foram 63,16%, para o *InfoGain* foram 61,11% e para o *ReliefF* foram 28,57% dos atributos selecionados pertencente a essa categoria. Não obstante, a maioria dos atributos extraídos pertencentes à categoria estatística são baseados na distribuição normal dos dados. Com isso, esses atributos apareceram com maior frequência, no caso binário, por serem melhores descritores da população das condições impostas experimentalmente.

Ainda na perspectiva binária de análise e seleção de atributos, a medida baseada em complexidade, dimensão fractal, apareceu selecionada em três algoritmos de seleção de atributos, *CFS*, *InfoGain* e *ReliefF*. Os fractais são formas complexas que não podem ser medidas apenas por dimensão topológica, sendo uma medida dimensional de maior complexidade computacional. A dimensão fractal surge então como uma alternativa de medição para obter o grau de complexidade de uma forma. Assim, essa medida é um valor que representa o quão densamente um conjunto de dados ocupa o espaço métrico em que ele existe. No caso de séries temporais, o significado matemático pode ser representado pela ausência ou presença de anomalias. Em contra-partida, esse atributo, dimensão fractal das séries temporais, apresenta uma complexidade computacional de tempo alta, dependendo do método empregado. Então, esse atributo apresenta uma desvantagem quando usado para aplicações em tempo real o que não ocorre com as medidas da categoria estatística.

A categoria de medidas baseadas em análise de sistemas, que são amplamente utilizadas como descritores de séries temporais na literatura, não foi selecionada por nenhum algoritmo de SA para o caso binário. Essa situação pode ter ocorrido pelo fato das medidas estatísticas, como média e desvio padrão, terem se sobrepostos às medidas em questão, tornando-as irrelevantes para essa análise. Quanto à categoria de medidas baseada em informação, que são representativas de algumas características de sinais, como ruído, intervalo dinâmico e taxas de distorção, foram selecionadas frequentemente para o caso binário.

Para o caso multiclasse, considerando os percentuais de atributos selecionados por cada método de SA, o algoritmo *CFS* foi o que selecionou os subconjuntos com mais atributos, com 8,15% dos atributos totais. Nesse cenário, as medidas estatísticas foram as mais selecionadas pelos algoritmos *CBF* e *CFS*, com 85,71% e 84,62% dos atributos selecionados, respectivamente. Ainda, esses algoritmos utilizaram a estratégia *Best First* de busca, encontrando os melhores descritores com base na consistência e correlação com as classes. Já os algoritmos com estratégia *Ranker* de busca, *InfoGain* e *ReliefF*, selecionaram majoritariamente os atributos da categoria baseada em informação. As medidas

da categoria estatística são mais simples e possuem uma relação com a distribuição dos dados, enquanto que as baseadas em informação são mais complexas e possuem relação com as características morfológicas dos sinais. Com isso, o uso das estratégias de busca referentes a cada algoritmo de SA prioriza a escolha desses atributos.

A categoria de medidas baseada em complexidade, representada apenas pela dimensão fractal, foi selecionada apenas pelo algoritmo *CBF*. Dessa maneira, esse atributo não foi capaz de descrever os dados experimentais para os outros algoritmos de SA utilizados, sendo observado o contrário para o caso binário de SA. Já a categoria de medidas baseada em análise de sistemas não foi selecionada por nenhum algoritmo, o mesmo foi constatado para a análise binária.

Ainda, todos os atributos extraídos do conjunto de dados foram avaliados quanto à sua distribuição de probabilidades para todas as 400 instâncias. Foi aplicado o teste de normalidade Kolmogorov Smirnov, com nível de significância de 5%, para cada atributo de todas as instâncias. Essa análise tem como objetivo identificar o comportamento dos atributos de acordo com suas tendências e distribuição. Lembrando que cada um dos atributos foram extraídos de cada um dos 11 sinais coletados pela bancada de experimentos. Dessa maneira, na Tabela 8 é apresentado a quantidade de atributos que seguem uma distribuição normal de acordo com todas as instâncias analisadas.

Categoria dos atributos	Nº total de atributos	Nº que apontam normalidade
Estatísticos	154	31
Baseados em informação	121	25
Baseados em complexidade	11	0
Baseados em análise de sistemas	33	0

Tabela 8 – Análise da distribuição normal pelos atributos extraídos em função de cada categoria.

Analisando essa tabela, atesta-se a frequência de atributos selecionados ser na maioria das categorias estatística e baseadas em informação. Essa ocorrência é explicada pelo fato que os atributos que passam no teste de normalidade e seguem uma distribuição normal, possam apresentar maior relevância com relação à classe por serem melhor descritos com medidas simples, como média, desvio padrão e curtose. Assim, recomenda-se, baseado nos resultados apresentados, que os atributos que maior se diferenciem com relação a classe e possuem maior impacto no sistema de diagnóstico sigam o comportamento de alguma distribuição de probabilidades, mais especificamente a distribuição normal, ou de Gauss.

De outra maneira, quanto ao tipo de transdutor utilizado para a coleta dos sinais experimentais para a então extração e seleção de atributos, pode-se observar algumas características sobre a relevância dos sinais, sendo de corrente elétrica, de tensão elétrica e de velocidade de vibração mecânica. A Tabela 9 apresenta os percentuais de atributos selecionado por cada algoritmo de seleção de atributos em função do sinal de referência para os atributos extraídos. Nessa tabela, é possível observar uma organização dos dados

das Tabelas 6 e 7 em função do tipo de sinal de entrada dos atributos extraídos, para os cenários binário e multiclasse.

	Contexto binário de SA (%)			
Tipo do sinal	CBF	CFS	InfoGain	ReliefF
Corrente elétrica	0,00	26,32	31,58	14,29
Tensão elétrica	0,00	47,37	52,63	57,14
Vibração mecânica	100,00	26,32	15,79	28,57
	Contexto multiclasse de SA (%)			
Tipo do sinal	CBF	CFS	InfoGain	ReliefF
Corrente elétrica	14,29	26,92	6,67	0,00
Tensão elétrica	42,86	19,23	80,00	57,14
Vibração mecânica	42,86	53,84	13,33	42,86

Tabela 9 – Porcentagem de atributos seleccionados em função do tipo de sinal experimental usado para a extração de atributos para cada algoritmo de seleção de atributos (SA) para os dois cenários experimentais.

Vale ressaltar que os atributos da Tabela 4 foram extraídos para cada sinal experimental, sendo três correntes de fase, nas fases A, B e C, três tensões de fase, nas fases A, B e C, e cinco velocidades de vibração mecânica, tangencial na carcaça, tangencial na base, axial no lado acionado, radial no lado acionado e radial no lado não acionado. Dessa maneira, os atributos oriundos dos sinais de velocidade de vibração mecânica apresentam uma maior quantidade em relação aos de tensão e corrente elétrica.

Analisando quanto ao tipo de sinal utilizado para extrair os atributos em função do algoritmo de SA, observa-se na Tabela 9 que para o CBF os sinais de vibração mecânica são hegemônicos para o contexto binário e multiclasse. No cenário multiclasse, os sinais de vibração estão empatados com os de tensão quanto ao quesito de atributos seleccionados. No algoritmo CFS, os sinais de vibração mecânica são mais representativos apenas para o caso multiclasse, e no caso binário os correspondentes são os de tensão elétrica.

Nos algoritmos ReliefF e InfoGain, os sinais predominantes são os de tensão elétrica. Outro aspecto, é a diferença na estratégia de busca desses algoritmos para a seleção dos subconjuntos de atributos representativos. Esse fato pode embasar a seleção de atributos majoritariamente provenientes dos sinais de tensão elétrica.

Em contrapartida, apenas com o processo de seleção de atributos não há evidências suficientes para elaborar uma ordem de importância entre os tipos de sinais avaliados. Dessa maneira, na seguintes seções serão apresentados os resultados dos sistemas de inferência como um todo para as tarefas de detecção e de diagnóstico da severidade de defeitos de barras quebradas em motores de indução. Após essas constatações, será possível avaliar a contribuição de cada tipo de sinal para o desempenho dos sistemas de inferência, possibilitando a elaboração de uma metodologia de recomendação de características e sinais para as tarefas de detecção e de diagnóstico de defeitos.

5.3 Comparação dos Modelos Induzidos

A Figura 16 apresenta a medida multicritério na forma de hexágonos para cada configuração de classificadores em função do algoritmo de SA para o caso de classificação binária. Ainda, a Figura 17 exibe um gráfico de barras, onde o eixo das ordenadas é representado pela área interna aos hexágonos da medida multicritério e o eixo das abcissas representa cada configuração de classificador e método de SA, também para o caso de classificação binária. Vale ressaltar que quanto menor a área interna ao hexágono multicritério, melhor habilitada é a configuração de métodos para o tema dessa pesquisa.

Figura 16 – Ilustração da medida multicritério para avaliação de desempenho de classificadores, para cada algoritmo de seleção de atributos (SA), para o caso de classificação binária.

É possível observar, na Figura 16, que o classificador *SVM* se sobressai sobre os outros, em termos de área do hexágono multicritério, em função de todos os métodos de SA utilizados. Ao se utilizar os métodos *CBF* e *CFS* para SA, a medida de Taxa de Erro (TE) são as maiores para esse método. Essa característica para o *CBF* é explicada pelo baixo número de atributos selecionados, sendo apenas um, de acordo com a Tabela 6. Desta maneira, o método *SVM* apresenta baixa generalização pela baixa dimensionalidade intrínseca a esse método, prejudicando a medida TE. Ao se aplicar o algoritmo *CFS*, baseado em correlação, a TE para o *SVM* foi prejudicada pelo fato de que os atributos selecionados serem pouco diferenciáveis entre si, prejudicando a correlação com a classe

Figura 17 – Áreas do hexágono multicritério para cada configuração de algoritmos de seleção de atributos (SA) e de aprendizado de máquina (AM) para o caso de classificação binária.

e a formação do hiperplano de separação das classes para esse método de aprendizado de máquinas.

Ao se aplicar os algoritmos *InfoGain* e *ReliefF* para o caso da classificação binária, o método *SVM* também é prejudicado, mas em componentes diferentes da medida multicritério. Nesses casos, a Área Complementar à curva *ROC* (ACROC) e a Taxa de Falsos Positivos (TFP) são as maiores encontradas. Pelas características de ranqueamento na tarefa de SA desses métodos, o SVM foi prejudicado pela alta taxa de alarmes falsos, ou TFP, que também faz parte da determinação da medida ACROC.

Verificando todas as configurações empregadas do sistema de detecção, método de SA em conjunto com o classificador, e de acordo com a Figura 17, observa-se que a área do hexágono multicritério é muito menor para todos os classificadores ao se aplicar os métodos *CBF* e *CFS*. No cenário de classificação binária era esperada essa característica, pois os algoritmos *CBF* e *CFS* são baseados em consistência e correlação, respectivamente, entre os atributos e as classes. Como nesse caso existem apenas duas classes, motor saudável e motor com rotor defeituoso, os classificadores puderam diferenciar com maior desempenho as classes para atributos com maior intercorrelação com as classes.

Dessa maneira, a Tabela 10 apresenta o *ranking* de sistemas de detecção de acordo com a medida multicritério para o caso da classificação binária. Assim, esse *ranking* pode ser utilizado para recomendar métodos de SA e de aprendizado de máquinas para se trabalhar

com dados e atributos semelhantes aos utilizados nesta pesquisa.

Ranking	Sistema de detecção	Área do hexágono multicritério
1 ^o	Ia5 (SVM)	0,016659
2 ^o	Ia3 (RD)	0,017495
3 ^o	Ia1 (NB)	0,017495
4 ^o	Ia2 (C4.5)	0,01786
5 ^o	Ib5 (SVM)	0,01824
6 ^o	Ia4 (MLP)	0,018681
7 ^o	Ib3 (RD)	0,019346
8 ^o	Ib4 (MLP)	0,019362
9 ^o	Ib2 (C4.5)	0,019452
10 ^o	Ib1 (NB)	0,019456
11 ^o	Ic1 (NB)	0,050578
12 ^o	Ic2 (C4.5)	0,050695
13 ^o	Id1 (NB)	0,050723
14 ^o	Id2 (C4.5)	0,050801
15 ^o	Ic4 (MLP)	0,051259
16 ^o	Ic3 (RD)	0,051338
17 ^o	Id4 (MLP)	0,051373
18 ^o	Id3 (RD)	0,051443
19 ^o	Id5 (SVM)	0,126549
20 ^o	Ic5 (SVM)	0,126551

Tabela 10 – *Ranking* para seleção dos sistemas de detecção para o cenário de classificação binária.

Com os resultados apresentados nessa tabela, verifica-se que os três primeiros lugares entre os sistemas de detecção foram construídos utilizando o algoritmo *CBF* para a tarefa de SA. Esse algoritmo trabalha baseado na consistência dos atributos com relação a classe, e para esse caso, foi selecionado apenas um atributo, que é o segundo quartil do sinal de vibração axial. Não obstante, o algoritmo SVM de aprendizado de máquinas se mostrou como o mais adequado para essa aplicação, pois, como discutido anteriormente, esse método gera melhores resultados quando a dimensionalidade é pequena e com isso o próprio SVM consegue gerar um hiperplano ótimo de separação das classes.

Em segundo e terceiro lugar estão métodos que se mostraram muito próximos em termos da medida multicritério. A RD baseada em árvores de decisão e o NB baseado em distribuição de probabilidades. A RD teve um bom desempenho também pelo fato de ser apenas um atributo selecionado pelo *CBF*, gerando uma árvore de baixa complexidade e com poucos galhos. Já o NB foi satisfatório devido ao atributo selecionado possuir uma distribuição normal dos valores para todas as instâncias, facilitando a classificação em termos estatísticos.

Para complemento dos resultados acerca dos sistemas de detecção de defeitos, na Tabela 11 são salientadas as médias e desvios padrão para algumas métricas importantes na comparação entre os modelos induzidos para a implementação de todos os algoritmos

de SA. Com essa tabela pode-se verificar aspectos gerais dos classificadores utilizados que podem auxiliar na escolha de uma configuração de métodos para aplicações, ressaltando que os testes foram realizados por meio de validação cruzada com 10 partições para cada configuração.

Modelo	Classe correta (%)	Ganho de entropia	Precisão
NB	99,4000(1,4510)	-97,5721(370,5309)	1,0000(0,0000)
C4.5	99,5667(1,3938)	-117,9615(449,0919)	0,9980(0,0200)
RD	100,0000(0,0000)	20,4317(0,2501)	1,0000(0,0000)
MLP	100,0000(0,0000)	21,6059(0,0149)	1,0000(0,0000)
SVM	80,0000(0,0000)	-6422,3415(0,0000)	0,7580(0,0270)
Modelo	Sensibilidade	Correlação de Mathews	AROC
NB	0,9700(0,0726)	0,9809(0,0463)	0,9874(0,0389)
C4.5	0,9800(0,0594)	0,9862(0,0450)	0,9898(0,0308)
RD	1,0000(0,0000)	1,0000(0,0000)	1,0000(0,0000)
MLP	1,0000(0,0000)	1,0000(0,0000)	1,0000(0,0000)
SVM	0,0000(0,0000)	0,0000(0,0000)	0,5000(0,0000)

Tabela 11 – Comparação dos modelos induzidos em termos de média e desvio padrão, entre parênteses, para medidas de importância para a classificação binária.

Já na Figura 18 são expostas as medidas multicritério ilustrada para cada configuração de sistema de diagnóstico em função do algoritmo de SA e do classificador para o caso de classificação multiclasse. Não obstante, na Figura 19 é apresentado um gráfico de barras, onde o eixo das ordenadas é representado pela área interna aos hexágonos da medida multicritério e o eixo das abcissas representa cada configuração de classificador e método de SA, também para o caso de classificação multiclasse.

Na abordagem de diagnóstico da severidade do defeito, é observado na Figura 18 que os classificadores se comportam de maneira totalmente diferente que para o caso de detecção binária de falhas. O primeiro aspecto desse comportamento, é que os atributos selecionados são diferentes, assim como a relação com as classes, e em segundo, o problema de análise é inteiramente diferente para alguns métodos, quando existem mais de duas classes.

Em todos os algoritmos de SA empregados, observa-se que as medidas de desempenho que mais prejudicaram a configuração de métodos foi a TFN e a TE. A TFN era esperada ser maior para o caso de análise multiclasse, pois os classificadores apresentam maior dificuldade na tarefa de diagnóstico. Já a TE, relacionada com o complemento da acurácia, também é prejudicada pelo fato de existirem múltiplas classes. Assim, ambas as medidas são prejudicadas pelo número de instâncias classificadas incorretamente, de forma direta ou indireta.

Nesse contexto, observando o gráfico de barras da Figura 19, os métodos de SA *CBF* e *CFS* apresentam os melhores resultados independente dos classificadores utilizados.

Figura 18 – Ilustração da medida multicritério para avaliação de desempenho de classificadores, para cada algoritmo de seleção de atributos (SA), para o caso de classificação multiclasse.

Como discutido anteriormente, esses métodos buscam, de acordo com alguma métrica, a intercorrelação dos atributos com as classes. Em contra-partida, os algoritmos *InfoGain* e *ReliefF* procuram aleatoriamente subconjuntos de atributos que melhor descrevam e diferenciem as classes. Assim, podem ser inseridos atributos pouco relevantes para os classificadores, denegrindo a tarefa de diagnóstico de defeitos.

De mesmo modo que para o caso binário de detecção de falhas, a Tabela 12 apresenta o *ranking* de sistemas de diagnóstico de acordo com a medida multicritério para o caso da classificação multiclasse.

Novamente, os dois primeiros lugares de sistema de diagnóstico apresentaram o *CBF* como seletor de atributos. O algoritmo RD e MLP se mostraram superiores a todos os outros em termos da área do hexágono multicritério. A MLP é um dos algoritmos mais utilizados para detecção e diagnóstico de defeitos na literatura, no entanto, apresenta uma desvantagem quanto ao tempo de treinamento do modelo. Já a RD, apresenta boas vantagem para o treinamento e para o teste, e ainda possibilita boa interpretabilidade do modelo, sendo um dos métodos mais adequados para aplicações embarcadas. Em terceiro lugar, a MLP em conjunto com o algoritmo *CFS* apresentou bons resultados, principalmente quanto a TFP, TFN e TE, tendo apenas maus resultados quanto à PAS, que é referente ao seletor de atributos *CFS*.

Figura 19 – Áreas do hexágono multicritério para cada configuração de algoritmos de seleção de atributos (SA) e de aprendizado de máquina para o caso de classificação multiclasse.

Ranking	Sistema de diagnóstico	Área do hexágono multicritério
1 ^o	Ila3 (RD)	0,002071
2 ^o	Ila4 (MLP)	0,002974
3 ^o	Ilb4 (MLP)	0,010156
4 ^o	Ilb2 (C4.5)	0,010926
5 ^o	Ila2 (C4.5)	0,011125
6 ^o	Ilb3 (RD)	0,011731
7 ^o	Ilb1 (NB)	0,012571
8 ^o	Ila5 (SVM)	0,017676
9 ^o	Ila1 (NB)	0,026036
10 ^o	Ilc2 (C4.5)	0,030459
11 ^o	IId2 (C4.5)	0,030959
12 ^o	Ilc4 (MLP)	0,031455
13 ^o	IId4 (MLP)	0,031998
14 ^o	Ilc3 (RD)	0,034474
15 ^o	IId3 (RD)	0,034851
16 ^o	Ilc1 (NB)	0,041256
17 ^o	IId1 (NB)	0,041451
18 ^o	Ilb5 (SVM)	0,157624
19 ^o	IId5 (SVM)	0,464311
20 ^o	Ilc5 (SVM)	0,464443

Tabela 12 – *Ranking* para seleção dos sistemas de diagnóstico para o cenário de classificação multiclasse.

Nesse contexto, avaliando todas as métricas de desempenho utilizadas no hexágono multicritério, existem algumas medidas mais importantes relacionadas a aplicação. Assim, pode-se identificar essas medidas como critério de desempate para os sistemas de detecção e diagnóstico de defeitos. Dentre as medidas de desempenho utilizadas, foi determinada que a TFP seja a mais adequada para ser o critério de desempate entre configurações de métodos. Quanto menor a taxa de alarmes falsos, ou TFP, mais conveniente é a configuração metodológica, pois para aplicações pode aumentar os custos de manutenção e de logística no tratamento do problema. Assim, essa medida deve ser usada de forma discriminante quando há o empate entre métodos, ou semelhança matemática, após a análise da área do hexágono multicritério.

Ainda, essa medida multicritério desenvolvida nesse trabalho, possui característica singular quanto às variáveis analisadas, que avalia tanto o processo de SA quanto a tarefa de classificação. Desse modo, essa medida pode ser aplicada para a comparação entre métodos e também para sistemas de recomendação de métodos para detecção e diagnóstico de defeitos.

Em acréscimo dos resultados a respeito dos sistemas de diagnóstico de defeitos, na Tabela 13 são exibidas as médias e desvios padrão para algumas métricas importantes na comparação entre os modelos induzidos para a implementação de todos os algoritmos de SA. Com essa tabela verifica-se tópicos gerais dos classificadores utilizados que podem auxiliar na escolha de uma configuração de métodos para aplicações, sendo que os testes foram realizados por meio de validação cruzada com 10 partições para cada configuração. Assim, cruzando os resultados do ranqueamento dos métodos por meio da medida multicritério com os resultados dessa tabela, a escolha do método de diagnóstico apresentará maior respaldo.

Modelo	Classe correta (%)	Ganho de entropia	Precisão
NB	83,7000(6,0469)	-1172,7077(854,2044)	1,0000(0,0000)
C4.5	89,2000(5,1709)	-2480,5797(1639,7885)	1,0000(0,0000)
RD	98,3667(2,5291)	58,0212(1,6978)	1,0000(0,0000)
MLP	98,6667(1,9534)	67,7513(1,8072)	1,0000(0,0000)
SVM	20,0000(0,0000)	-25706,3422(0,0000)	0,2000(0,0000)
Modelo	Sensibilidade	Correlação de Mathews	AROC
NB	0,9833(0,0503)	0,9894(0,0318)	0,9894(0,0321)
C4.5	0,9817(0,0575)	0,9883(0,0367)	0,9908(0,0288)
RD	1,0000(0,0000)	1,0000(0,0000)	1,0000(0,0000)
MLP	1,0000(0,0000)	1,0000(0,0000)	1,0000(0,0000)
SVM	0,2000(0,4020)	0,0000(0,0000)	0,5000(0,0000)

Tabela 13 – Comparação dos modelos induzidos em termos de média e desvio padrão, entre parênteses, para medidas de importância para a classificação multiclasse.

5.4 Verificação da Robustez dos Modelos Induzidos

Com a verificação da robustez¹ dos sistemas induzidos foi possível verificar o quanto de ruído branco um sistema de inferência pode suportar nos sinais de entrada até que apresente um comportamento instável quanto ao seu desempenho. Vale ressaltar que a metodologia para a verificação da robustez se baseia em gerar modelos de classificação por meio de dados corrompidos por ruídos, e após, o processo de teste, é realizado também por dados corrompidos com ruídos de mesma intensidade. Essa verificação tem por objetivo avaliar majoritariamente os algoritmos utilizados para a indução dos sistemas de inferência de maneira genérica, excluindo em partes a influência do sistema físico avaliado. Ainda, para a análise dos resultados é importante salientar que a métrica tempo de aprendizado, utilizada na medida multicritério de desempenho, não foi alterada para a verificação de robustez. Fato devido a essa medida ser relativa à complexidade computacional dos métodos, e não referente aos sinais de entrada. Essa medida foi apenas utilizada para as análises das áreas dos hexágonos multicritério.

5.4.1 Verificação da Robustez dos Modelos Induzidos para a Classificação Binária

Na Figura 20 é apresentada a PAS para cada algoritmo de SA em função da porcentagem de ruído branco adicionada aos sinais experimentais para o cenário da classificação binária. Observa-se que a PAS para os algoritmos Infogain e ReliefF aumenta até chegar em 10% de ruído e depois começa a diminuir. Ainda, o algoritmo CFS mantém-se constante quanto à PAS até 80% de ruído, e após tende a zero. Já o algoritmo CBF se mantém próximo de zero para todas as porcentagens de ruídos inseridos. Esses fatos são decorrentes da natureza de cada algoritmo, pois com o aumento da porcentagem de ruídos nos sinais de entrada, tem-se como hipótese que os atributos se tornam menos representativos do sinal original, evidenciando que a PAS decai a medida que a porcentagem de ruídos aumenta. Ainda, para os algoritmos baseados em ranqueamento, como o Infogain e o ReliefF, há uma instabilidade até os 10% de ruído que pode ser derivada do fato de que os subconjuntos de atributos selecionados sejam pouco representativos e redundantes aos sinais originais, aumentando a PAS.

Para o caso da classificação binária, nas Figuras 21, 22, 23, 24 e 25 são apresentados os resultados da verificação de robustez.

¹O termo **verificação da robustez** utilizado nesse trabalho diz respeito ao ato de conferir robustez ao sistema de inferência, devendo ser considerado distinto do termo **análise da robustez** que se deve ao ato de se determinar o quão robusto um sistema de inferência se mostrou para as condições experimentais impostas.

Figura 20 – Porcentagem de atributos selecionados por cada algoritmo de seleção de atributos (SA), para o caso da classificação binária, em função da porcentagem de ruído branco agregado ao sinal de entrada para a extração dos atributos.

Figura 21 – Taxa de falsos positivos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária.

Figura 22 – Taxa de falsos negativos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária.

Figura 23 – Taxa de erros em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária.

Figura 24 – Área complementar à curva ROC em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação binária.

Figura 25 – Área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação binária.

Quanto ao parâmetro de avaliação TFP, não há variações expressivas em função da porcentagem de ruído branco adicionado ao sinal de entrada para todos os algoritmos de SA, com exceção dos casos em que o classificador é o algoritmo SVM. Pode ser observado na Figura 21 que o algoritmo SVM passa por uma instabilidade quando são adicionados ruídos nos sinais de entrada para o caso do algoritmo CFS como seletor de atributos. Isso é atestado pela própria natureza desse método de classificação, que apresenta uma hipótese de separar as duas classes por meio de hiperplanos, e com a inserção de ruídos, os atributos são alterados de forma estatística tornando esse método instável. Ainda, para o caso dos algoritmos Infogain e ReliefF para SA, o SVM também apresenta um comportamento distinto dos demais algoritmos, mas, nesses casos, a TFP é diminuída até 10% de ruído e permanece constante até o final dos testes, para o ReliefF, e até 90%, para o Infogain. De maneira geral, após realizar comparações estatísticas, por meio do teste de normalidade de Shapiro-Wilk, para cada configuração de técnicas de SA e de AM em função do percentual de ruído adicionado nos sinais de entrada não foi possível encontrar diferenças estatisticamente significativas para um intervalo de confiança de 95% para todas as configurações, com exceção da configuração do algoritmo CFS com o SVM. Ainda, vale ressaltar que os testes foram feitos para verificar se em uma mesma configuração de técnicas haveria diferenças estatísticas em função do percentual de ruído, e, assim, o teste de normalidade funcionou como um teste de hipótese de comparação entre os ensaios.

A TFN, no caso da classificação binária, é apresentada em função do percentual de ruído na Figura 22. O comportamento do parâmetro de avaliação TFN se mostrou semelhante ao da TFP, mostrando a instabilidade do algoritmo SVM e também o aumento desse parâmetro para os seletores de atributos baseados no ranqueamento. Ainda, fica nítida o aumento da TFN para a configuração do ReliefF com a RD após os 80% de ruído adicionado. Esse fato é relacionado à PAS, discutido de acordo com a Figura 20, que se mostra instável após 50% de ruído. Essa circunstância pode gerar árvores de decisão com muitas condições para uma classificação binária, provocando o sobreajuste. A comparação entre os percentuais de ruído para cada uma das configurações, mostrou diferença estatisticamente significativa para os casos de CFS em conjunto com o SVM, do Infogain em conjunto com o SVM, do ReliefF em conjunto com o SVM e com a RD, para o intervalo de confiança de 95%. Ainda, vale ressaltar, que a configuração do ReliefF com a RD não apresentou diferenças estatísticas até 80% de ruído para o nível de significância de 5%.

Para o parâmetro de avaliação TE para a classificação binária, observa-se os resultados da verificação de robustez na Figura 23. Nesse caso, nota-se novamente o comportamento instável do algoritmo SVM com o CFS como seletor de atributos. Ainda, o comportamento do SVM é semelhante ao discutido para as medidas de avaliação TFP e TFN. Um ponto importante observado é para o algoritmo RD associado ao ReliefF, que nesse caso se mostrou instável quanto a TE, sendo possivelmente decorrente do sobreajuste e da PAS. Quanto às comparações estatísticas, para um intervalo de confiança de 95%, houve

diferença estatisticamente significativa para a configuração do CFS com o SVM e para o ReliefF com a RD.

Os resultados para a verificação da robustez levando em consideração a medida ACROC, Figura 24, foram semelhantes aos apresentados pelas outras medidas. Essa constatação era esperada pois essa medida é derivada dos parâmetros da classificação como TFP, TFN e TE. Quanto às comparações estatísticas, a configuração do CFS com o SVM foi a única que apresentou diferença estatisticamente significativa para um nível de significância de 5%.

Como desfecho da verificação de robustez para a classificação binária, foi analisada também a área do hexágono multicritério em função dos percentuais de ruído adicionado no sinal de entrada, apresentado na Figura 25. Vale lembrar, que quanto menor o valor da área do hexágono multicritério, melhor é o desempenho da configuração de técnicas de acordo com as medidas de avaliação que compõem os vértices do hexágono.

Na análise dos classificadores que utilizaram o CBF como seletor de atributos, não houve diferença estatisticamente significativa para a área do hexágono multicritério, para um intervalo de confiança de 95%. Já para os algoritmos CFS e Infogain, houve diferença estatisticamente significativa apenas para o classificador SVM. E, para o algoritmo ReliefF, nenhum dos classificadores passaram no teste de hipótese. No entanto, para esse caso, do ReliefF como seletor de atributos, os classificadores apresentaram comportamento com morfologia semelhante. Esse fato demonstra a importância do seletor de atributos nos modelos induzidos, pois cada classificador possui uma diferente natureza, estatística, probabilística ou lógica.

Quanto à robustez, o algoritmo CBF para SA, se mostrou o mais estável entre os quatro analisados para essa tarefa. Quanto aos classificadores, o NB, o C4.5 e a MLP em todas as medidas de avaliação se mostraram estáveis para todas as porcentagens de ruído adicionados ao sinal de entrada. Assim, recomenda-se que essa verificação de robustez seja utilizada em conjunto com a avaliação de desempenho, apresentada na Seção 5.3, e a aplicabilidade para a escolha dos métodos mais adequados para compor o sistema de inferência para a classificação binária.

Como suporte para a tomada de decisão na recomendação de sistemas inteligentes, foram determinadas as áreas sob as curvas de área do hexágono multicritério em função do percentual de ruído, Figura 25. Dessa maneira, foi possível obter um ranqueamento dos modelos induzidos que apresentaram menor área para o parâmetro de avaliação multicriterial, sendo apresentado na Tabela 14.

Para o problema de detecção de falhas sem a especificação da severidade do defeito, classificação binária, o seletor de atributos CBF ocupou os cinco primeiros lugares no *ranking* para a verificação de robustez. Essa característica é semelhante à apresentada na comparação dos modelos induzidos na Seção 5.3. O algoritmo CBF, baseado na consistência dos subconjuntos de atributos, foi o que elegeu o menor número de atributos

Ranking	Sistema de detecção	Área sob a curva de verificação de robustez
1 ^o	Ia3 (RD)	0,002787
2 ^o	Ia1 (NB)	0,002803
3 ^o	Ia4 (MLP)	0,003040
4 ^o	Ia2 (C4.5)	0,003047
5 ^o	Ia5 (SVM)	0,004110
6 ^o	Ib2 (C4.5)	0,006080
7 ^o	Ib1 (NB)	0,006092
8 ^o	Ib4 (MLP)	0,006117
9 ^o	Ib3 (RD)	0,006161
10 ^o	Ic1 (NB)	0,286187
11 ^o	Ic4 (MLP)	0,286919
12 ^o	Ic2 (C4.5)	0,288534
13 ^o	Ic3 (RD)	0,290505
14 ^o	Ib5(SVM)	0,417827
15 ^o	Id1 (NB)	1,044991
16 ^o	Id4 (MLP)	1,047317
17 ^o	Id2 (C4.5)	1,048087
18 ^o	Id3 (RD)	1,087128
19 ^o	Ic5(SVM)	1,110960
20 ^o	Id5(SVM)	1,577159

Tabela 14 – *Ranking* para identificação dos sistemas de detecção para o cenário de classificação binária em função da verificação da robustez.

para todos os níveis de ruído nos sinais. Dessa maneira, os subconjuntos de atributos se tornam mais simples de serem diferenciados pelos classificadores, e, conseqüentemente melhorando a magnitude das medidas de avaliação, como TFP, TFN, TE, e ACROC.

Os primeiros lugares do *ranking* apresentaram área sob as curvas com ordem de grandeza semelhante. O primeiro tendo como classificador a RD baseada em um conjunto de árvores de decisão e em segundo o NB fundamentado em aproximações de modelos probabilísticos. Como o ruído branco adicionado segue uma distribuição de probabilidades normal, era de se esperar que o classificar NB obtivesse um desempenho satisfatório, pois a partir de um certo nível de ruído, todos os atributos seguiriam uma distribuição normal. No entanto, a contar desse nível de ruído que todos os atributos seguem uma distribuição normal, o classificador apresenta dificuldade na diferenciação entre as classes, pela ocorrência de redundâncias. Problema esse que é pouco verificado no caso da classificação binária, como apresentado pelas curvas de verificação de robustez.

Dessa maneira, constatou-se que quanto maior o nível de ruído adicionado ao sinal de entrada mais os atributos se tornam irrelevantes para com a classe. A verificação de atributos irrelevantes às instâncias da base de dados, geralmente, confunde o algoritmo de aprendizado. Esse fato ocorre pelo fato dos atributos irrelevantes serem escolhidos pelo seletor de atributos e, com isso, ocorrem erros aleatórios durante o teste (HAND; ADAMS,

2014).

Outro aspecto a ser considerado é o motivo da seleção dos atributos irrelevantes. Na medida em que o modelo é construído, como uma árvore de decisão ou uma rede neural, cada vez menos dados estão disponíveis para auxiliar a escolha do atributo. Assim, alcança-se um estado em que os atributos aleatórios aparentam serem bons apenas por acaso. No caso das árvores de decisão, como os algoritmos C4.5 e RD, a chance desse processo ocorrer aumenta com a profundidade da árvore. De maneira geral, como os algoritmos baseados em árvores de decisão reduzem continuamente a quantidade de dados em que se baseiam suas escolhas, é necessário, para manter um desempenho adequado, que o número de instâncias cresça exponencialmente com o número de atributos irrelevantes. Caso esse apresentado pela inserção de ruídos aos sinais de entrada (HAND; ADAMS, 2014; BHARGAVA et al., 2013; CHUNG; GRAY, 1999).

Já os algoritmos de aprendizado de máquinas baseados em distribuições de probabilidades, como o NB, assumem que todos os atributos são independentes entre si. Essa suposição foi constatada pelos resultados, Tabela 14, como correta para os atributos irrelevantes. No entanto, essa suposição pode afetar o modelo de aprendizado quando existem atributos redundantes (SARITAS; YASAR, 2019; JIANG; ZHANG; CAI, 2008).

Nesse contexto, para findar a verificação de robustez no caso da detecção de defeitos, observou-se que os modelos que criam ajustes matemáticos, ou estatísticos, se comportaram de maneira mais comportada quando há a presença de dados corrompidos por ruídos. Outra constatação obtida por meio dos resultados da PAS, foi a de que quanto menor o número de atributos selecionados, melhor foi o desempenho e mais generalista foram os modelos.

5.4.2 Verificação da Robustez dos Modelos Induzidos para a Classificação Multiclasse

Para o caso da seleção de atributos no cenário multiclasse, na Figura 26 é apresentado os resultados para essa tarefa em função da porcentagem de ruídos adicionada ao sinal de entrada para os quatro algoritmos implementados. Observa-se que o comportamento dos algoritmos de SA no caso multiclasse é semelhante ao apresentado para a classificação binária na Seção 5.4.1.

Nas Figuras 27, 28, 29, 30 e 31 são apresentados os resultados para verificação de robustez dos modelos induzidos para o cenário de classificação multiclasse. Vale ressaltar, que para a análise de desempenho e verificação da robustez dos modelos, quanto menor for o valor dos parâmetros de avaliação, melhor será o desempenho do modelo.

Figura 26 – Porcentagem de atributos selecionados por cada algoritmo de seleção de atributos (SA), para o caso da classificação multiclasse, em função da percentagem de ruído branco agregado ao sinal de entrada para a extração dos atributos.

Figura 27 – Taxa de falsos positivos em função da percentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse.

Figura 28 – Taxa de falsos negativos em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse.

Figura 29 – Taxa de erros em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse.

Figura 30 – Área complementar à curva ROC em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos no caso da classificação multiclasse.

Figura 31 – Área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação multiclasse.

O parâmetro de avaliação TFP, Figura 27, demonstrou um comportamento, no geral, crescente à medida que a porcentagem de ruído aumenta. A exceção foi constatada pela implementação do algoritmo SVM para a tarefa de classificação em conjunto com os algoritmos CFS, Infogain e ReliefF de SA. Nesses casos, houve instabilidade e crescimentos expressivos da TFP, demonstrando a sensibilidade do SVM aos ruídos e também quando há um problema multiclasse. Já os outros modelos apresentaram uma reação comportada e esperada na presença de ruídos para a geração e testes dos modelos.

Ainda, pode-se evidenciar que os modelos induzidos deixam de ter um comportamento estatisticamente constante, ao nível de significância de 5%, após 70% de ruído branco ser adicionado ao sinal de entrada. Com as exceções do SVM como classificador em conjunto com os algoritmos CFS, Infogain e ReliefF para SA, sendo que nesses casos não houve semelhanças estatísticas para o parâmetro de avaliação TFP. Essas características são semelhantes às constadas pela medida de avaliação TFN, Figura 28, apenas com magnitudes reduzidas. Ou seja, para o caso de classificação multiclasse os erros do tipo I e do tipo II são correlacionados.

Quanto à medida de avaliação TE, Figura 29, que relaciona-se como o inverso da acurácia, os resultados apresentados foram os previstos. Ou seja, com o aumento de ruído adicionado ao sinal entrada, houve um aumento na taxa de erro. Com a exceção do algoritmo SVM, que apresentou instabilidade quando associado ao seletor de atributos CFS e valores relativamente altos quando associado ao Infogain e ReliefF. Essas últimas constatações foram discutidas na Seção 5.3 com base na natureza desse método de classificação para a formação de hiperplanos que diferenciam as classes.

Ao se comparar estatisticamente os valores de TE para cada modelo induzido em função da porcentagem de ruído branco por meio do teste de Shapiro-Wilk, não foram evidenciadas diferenças estatisticamente significativas, para um intervalo de confiança de 95%, até o percentual de 70% de ruído branco. Com a exceção dos modelos baseados no classificador SVM associado aos seletores de atributos CFS, Infogain e ReliefF. No entanto, como se almeja recomendar sistemas de inferência, deve-se atentar ao valor da magnitude dessas medidas de avaliação, como foi feito na comparação dos modelos induzidos na Seção 5.3.

Outra medida de avaliação que apresenta acentuada importância é a ACROC, Figura 30. A análise dos resultados para essa medida de avaliação é semelhante às realizadas anteriormente, com uma diferença de que o nível de ruído admitido pelos modelos sem que hajam diferenças estatisticamente significativas foi de 80% para todos os modelos, com a exceção do SVM em conjunto com o CFS. Ainda, o modelo resultante da associação do SVM com o ReliefF não apresentou diferença estatisticamente significativa para a ACROC até a adição de 100% de ruído branco. No entanto, a magnitude dessa medida torna esse modelo insatisfatório para os dados utilizados.

Para se ter um panorama geral sobre a verificação de robustez, foram construídos os

gráficos da medida multicritério utilizada em função da porcentagem de ruído branco, Figura 31, associando todas as medidas de avaliação discutidas anteriormente. Por meio da comparação estatística, ao nível de 5% de significância, o limite de robustez encontrado foi de 80% de ruído branco adicionado ao sinal de entrada. Esse limite foi constatado para todos os modelos induzidos para os seletores de atributos CBF e CFS, esse com a exceção do SVM como classificador. Após esse limite, a área do hexágono multicritério aumenta de maneira considerável, constatando-se a degradação dos modelos induzidos.

Nesse sentido, para os modelos induzidos por meio dos algoritmos Infogain e ReliefF não houve relações estatísticas que representaram uma distribuição normal para o intervalo de confiança de 95%. No entanto, para esses casos as curvas de área do hexágono multicritério em função da porcentagem de ruído branco apresentam uma morfologia semelhante. Ou seja, classificadores de diferentes naturezas matemáticas, lógicas e ou probabilísticas apresentam a mesma conduta quando os mesmos atributos são utilizados para diferentes níveis de ruído branco. Lembrando que esse fato foi constatado para os seletores de atributos baseados na estratégia de ranqueamento dos atributos, e isso pode esclarecer que os métodos de classificação foram mais genéricos nesses casos, mesmo com a medida de avaliação tendo maior magnitude para os seletores de atributos baseados na correlação ou consistência.

De mesmo modo que para o caso de classificação binária, foi determinado o ranqueamento dos sistemas de diagnóstico de defeitos em função da área sob a curva de área do hexágono multicritério em função do percentual de ruído branco para o problema multiclasse. O *ranking* dos modelos induzidos para a verificação da robustez pode ser observado na Tabela 15.

Quanto ao método de seleção de atributos, o algoritmo CBF ocupa a primeira posição nesse *ranking*. Os fatores para isso foram discutidos para o caso da classificação binária levando em consideração a quantidade de atributos selecionados, tornando o sistema de inferência mais generalista. Após, da segunda a quinta posição estão os modelos induzidos utilizando o algoritmo CFS de seleção de atributos. A natureza desse algoritmo é baseada na correlação dos subconjuntos de atributos com as classes. Dessa maneira, para o caso multiclasse, houve elevada correlação entre os atributos e a classe, sendo perceptível pelos valores de PAS apresentados na Figura 26.

No que diz respeito aos métodos utilizados para o aprendizado de máquinas, e, consequentemente aos modelos induzidos, novamente o algoritmo RD apresentou o melhor resultado para a verificação de robustez. Esse resultado é complementado com os apresentados pela comparação dos modelos induzidos apresentado na Seção 5.3.

Em segundo lugar está a MLP e em terceiro a RD, agora associada ao algoritmo CFS, com a mesma ordem de grandeza para o valor da área sob a curva da verificação de robustez. A MLP se mostrou nesse trabalho como um método de classificação generalista e capaz de suportar ruídos nos sinais de entrada. Esse fato é constatado pelos resultados

Ranking	Sistema de diagnóstico	Área sob a curva de verificação de robustez
1 ^o	Ila3 (RD)	0,075019
2 ^o	Iib4 (MLP)	0,094578
3 ^o	Iib3 (RD)	0,095223
4 ^o	Iib2 (C4.5)	0,105121
5 ^o	Iib1 (NB)	0,106626
6 ^o	Iia4 (MLP)	0,106863
7 ^o	Iia2 (C4.5)	0,107369
8 ^o	Iia1 (NB)	0,176168
9 ^o	Iia5 (SVM)	0,217231
10 ^o	Iib5 (SVM)	0,434357
11 ^o	Iic2 (C4.5)	0,538261
12 ^o	Iic4 (MLP)	0,562390
13 ^o	Iic3 (RD)	0,567203
14 ^o	Iic1 (NB)	0,567840
15 ^o	Iid2 (C4.5)	0,915931
16 ^o	Iid1 (NB)	0,956078
17 ^o	Iid4 (MLP)	0,959026
18 ^o	Iid3 (RD)	0,973601
19 ^o	Iic5 (SVM)	1,225448
20 ^o	Iid5 (SVM)	1,466952

Tabela 15 – *Ranking* para identificação dos sistemas de diagnóstico para o cenário de classificação multiclasse em função da verificação da robustez.

já apresentados nas análises anteriores.

Nesse sentido, para finalizar a verificação de robustez no caso do diagnóstico da severidade de defeitos, observou-se que os modelos que criam ajustes matemáticos, ou estatísticos, se apresentaram de maneira mais comportada quando há a presença de dados corrompidos por ruídos. Outra constatação obtida por meio dos resultados da PAS, foi a de que quanto menor o número de atributos selecionados, melhor foi o desempenho e mais generalista foram os modelos.

5.5 Análise da Robustez dos Modelos Induzidos

Para a avaliação dos sistemas de inferência induzidos para o caso específico de detecção e diagnóstico de defeitos em MITs, foi proposto o protocolo para a análise de robustez. A análise de robustez permitiu determinar o quanto os sistemas de inferência é robusto para os sinais avaliados e as condições experimentais impostas. Para isso, os modelos induzidos por meio dos atributos extraídos dos sinais obtidos experimentalmente, Seção 5.3, foram testados com atributos extraídos de sinais corrompidos com porcentuais de ruído branco. Vale ressaltar que a métrica tempo de aprendizado, utilizada na medida multicritério de desempenho, não foi alterada para a análise de robustez. Fato devido a essa medida ser

relativa a complexidade computacional dos métodos, e não referente aos sinais e entrada. Ainda, a métrica porcentagem dos atributos selecionados possui o mesmo comportamento que o apresentado na verificação da robustez, Seção 5.4, pois os atributos selecionados com os sinais corrompidos foram os mesmos. Assim, essas medidas foram apenas utilizadas para as análises das áreas dos hexágonos multicritério.

Na Figura 32 podem ser observadas as curvas de análise de robustez para o caso de detecção de defeitos, classificação binária, composta pela área do hexágono multicritério em função da porcentagem de ruído branco adicionado ao sinal de entrada para os conjuntos de atributos de teste. Já na Tabela 16 é apresentado o ranqueamento dos sistemas de inferência para a tarefa de detecção de defeitos em função da área sob as curvas da Figura 32 no cenário de classificação binária.

Figura 32 – Análise da robustez por meio da área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação binária.

O limite de robustez de um dado sistema de inferência é a quantidade de ruído branco que o sistema suporta até que exista diferença estatisticamente significativa em suas medidas de avaliação, como a medida multicritério. Por meio do teste de normalidade de Shapiro-Wilk, com nível de significância de 5%, observou-se que o limite de robustez para todos os sistemas induzidos no caso da classificação binária foi de 60%, com exceção das configurações do CFS com o SVM, Infogain com o SVM, Relief com o SVM e RD que não apresentaram semelhanças estatisticamente significativas no valor da área do hexágono multicritério.

Dessa maneira, observando o *ranking* da Tabela 16, pode-se verificar de maneira quantitativa os sistemas de inferência que melhor se desempenharam quando há a existência

Ranking	Sistema de detecção	Área sob a curva de análise de robustez
1º	Ia3 (RD)	0,344820
2º	Ia1 (NB)	0,346817
3º	Ia5 (SVM)	0,351734
4º	Ia2 (C4.5)	0,354746
5º	Ib1 (NB)	0,379815
6º	Ia4 (MLP)	0,380073
7º	Ib4 (MLP)	0,449591
8º	Ib3 (RD)	0,464997
9º	Ib2 (C4.5)	0,473017
10º	Ic4 (MLP)	0,535574
11º	Id2 (C4.5)	0,543721
12º	Id4 (MLP)	0,545962
13º	Id1 (NB)	0,549598
14º	Ic1 (NB)	0,562571
15º	Ic2 (C4.5)	0,571113
16º	Ic3 (RD)	0,605781
17º	Id3 (RD)	0,765521
18º	Ib5(SVM)	1,189687
19º	Id5(SVM)	1,861969
20º	Ic5(SVM)	1,891519

Tabela 16 – *Ranking* para identificação dos sistemas de detecção para o cenário de classificação binária em função da análise de robustez.

de ruídos nos atributos de teste. Assim, tem-se do primeiro ao sexto lugar configurações de algoritmos que apresentaram a área sob a curva com mesma ordem de grandeza. Uma característica dessa constatação é a de que cinco destas configurações são baseadas no seletor de atributos CBF, que apresentou bom desempenho nas análises anteriores. Quanto ao algoritmo de aprendizado de máquinas utilizado, a RD e o NB foram os melhores colocados, mesmo tendo naturezas distintas. Assim, constata-se novamente a importância da escolha do seletor de atributos adequado para o desempenho de um sistema de inferência como um todo.

O algoritmo SVM que ficou na terceira posição dessa classificação é o que apresenta a melhor definição matemática para a classificação binária, quando não há grandes quantidades de atributos e nem atributos redundantes. Nesse caso, a associação com o seletor CBF justifica a terceira posição no ranqueamento. Outras constatações sobre o SVM associado com diferentes seletores de atributos, podem ser observadas na Figura 32. Além disso, as outras três configurações que o SVM faz parte são os últimos lugares do *ranking*.

Nessa análise de robustez, os atributos começam a perder a qualidade de diferenciação a medida que a porcentagem de ruído aumenta. Assim, é importante que os atributos selecionados sejam representativos do sistema físico o bastante para suportarem ruídos e não se tornarem redundantes. Ainda, com a existência de atributos pouco representativos,

os algoritmos tendem a se confundir e considerar atributos ruins como bons apenas por acaso, gerando erros aleatórios que são observados no crescimento das curvas de análise de robustez apresentadas.

As curvas de análise de robustez de cada uma das métricas que compõem o vértice do hexágono multicritério podem ser observadas no Apêndice B. No entanto, as discussões aqui apresentadas foram realizadas considerando todas as métricas.

Para o cenário de diagnóstico da severidade de defeitos em MITs, classificação multi-classe, os resultados da análise de robustez são apresentados na Figura 33 e na Tabela 17.

Figura 33 – Análise da robustez por meio da área do hexágono multicritério em função da porcentagem de ruído branco inserido ao sinal de entrada para todos os modelos induzidos, separado por algoritmos de seleção de atributos, no caso da classificação multiclasse.

Analisando estatisticamente as curvas de análise de robustez no caso multiclasse, Figura 33, ao nível de confiança de 5% com o teste Shapiro-Wilk, constatou-se que o limite de robustez para os sistemas de inferência foi de 80% para todos os casos, com exceção das configurações do CBF com o SVM e NB, CFS com o SVM, Infogain e ReliefF com o SVM. Estas exceções apresentaram diferenças estatisticamente significativa antes do limite de 80% de ruído branco.

Vale lembrar que as métricas que compõem os vértices do hexágono multicritério apresentado em termos de área na Figura 33 são apresentadas no Apêndice B.

Uma constatação importante na comparação entre a análise de robustez nos cenários binários e multiclasse é a diferença entre os limites de robustez identificados, sendo 60% e 80% respectivamente. Uma hipótese para essa ocorrência é o fato de os sistemas de inferência para o caso multiclasse apresentem maior quantidade de regras e relações matemáticas entre os atributos para descreverem uma determinada classe. Sendo assim,

Ranking	Sistema de diagnóstico	Área sob a curva de análise de robustez
1 ^o	I Ib4 (MLP)	0,200440
2 ^o	I Ia3 (RD)	0,205015
3 ^o	I Ib2 (C4.5)	0,214167
4 ^o	I Ia2 (C4.5)	0,219927
5 ^o	I Id4 (MLP)	0,223871
6 ^o	I Ib3 (RD)	0,226437
7 ^o	I Ic4 (MLP)	0,229875
8 ^o	I Ic2 (C4.5)	0,231314
9 ^o	I Ib1 (NB)	0,237047
10 ^o	I Ia4 (MLP)	0,249584
11 ^o	I Id2 (C4.5)	0,287198
12 ^o	I Ic3 (RD)	0,293722
13 ^o	I Ic1 (NB)	0,305044
14 ^o	I Id3 (RD)	0,327247
15 ^o	I Id1 (NB)	0,342471
16 ^o	I Ia5 (SVM)	0,406418
17 ^o	I Ia1 (NB)	0,474644
18 ^o	I Ic5 (SVM)	0,836828
19 ^o	I Ib5 (SVM)	0,893055
20 ^o	I Id5 (SVM)	1,031193

Tabela 17 – *Ranking* para identificação dos sistemas de diagnóstico para o cenário de classificação multiclasse em função da análise de robustez.

com a maior discretização dos sistemas de inferência em termos das relações de atributos e classe, constroem-se sistemas mais robustos e menos probabilidade de ocorrência de erros aleatórios durante a construção dos modelos de aprendizado de máquinas. Essa observação é exatamente o oposto ao apresentado pelo processo de verificação da robustez, Seção 5.4, na qual os modelos são construídos e testados por meio dos atributos já corrompidos com ruídos.

Quanto ao ranqueamento das áreas sob as curvas de análise de robustez, Tabela 17, nota-se que os dois primeiros lugares apresentam valores relativamente próximos. Ambos sistemas de inferência, um constituído pelo CFS com a MLP e outro pelo CBF com a RD, são generalistas por sua natureza de construção. A MLP quando testada com atributos corrompidos resolve uma série de sistemas lineares que correspondem as camadas da rede. Esse fato também pode ser observado de forma análoga na RD que gera um conjunto de árvores de decisão, se assemelhando a um sistema especialista. Esses algoritmos apresentam uma linearidade para a classificação, o que foi benéfico para os testes com atributos corrompidos, pela simplicidade na identificação das classes. Esse fato não ocorre com os algoritmos baseados em estatística, nesse caso, pois os atributos se tornam cada vez mais irrelevantes e redundantes para com a classe quando analisados isoladamente.

Assim, para a classificação multiclasse, observou-se a importância de se relacionar os

subconjuntos de atributos com a classe, e não os atributos isoladamente, como demonstrado pelos resultados.

5.6 Recomendação de Sistema de Detecção e Diagnóstico

Os resultados apresentados permitem aferir diversas considerações a respeito dos processos de construção de um sistema de detecção e de diagnóstico de severidade de defeitos em MITs. É mostrado também que o processo perde a generalidade quando a medida de importância usada for baseada no desempenho preditivo do classificador induzido. Contudo, estudos, como Han, Kamber e Pei (2011), Covões (2010), Brazdil et al. (2009), Lee e Giraud-Carrier (2008), evidenciam que em algumas situações atributos selecionados utilizando o desempenho preditivo de um determinado algoritmo de AM para classificação podem ser utilizados com sucesso por outros classificadores (KOHAVI; JOHN, 1997).

Neste trabalho as vantagens da redução do conjunto de atributos foram apresentadas, analisadas experimentalmente e discutidas. Entretanto, é importante notar que a redução do conjunto de atributos pode não ser interessante em algumas situações, como em casos onde há poucos atributos, quando há mais de duas classes ou, então, quando se sabe sobre a relevância dos dados.

Outra questão a ser considerada é que os algoritmos de SA existentes não são de uso geral. Comumente cada um desses métodos apresentam algumas restrições com relação aos tipos de dados aos quais podem ser aplicados. A seguir são descritas três das principais características associadas aos dados que devem ser levadas em consideração para a escolha de um método de SA.

- Tipo dos dados: é necessário observar os tipos de valores que os atributos podem assumir. Alguns métodos podem não ser capazes de tratar todos esses tipos de dados. Além disso, considerando as classes, a maioria dos métodos de aprendizado trabalham apenas com problemas de classificação em que cada exemplo possui uma única classe associada. Porém, estudos demonstram a proposição de diferentes algoritmos de aprendizado para lidar com múltiplas classes (TSOUMAKAS; KATAKIS, 2007);
- Dimensão dos dados: devem ser considerados tanto a capacidade de o algoritmo trabalhar com um conjunto de treinamento reduzido quanto com um grande volume de dados (LIU; MOTODA, 2008);
- Presença de ruído: alguns métodos podem não ser capazes de manipular dados com ruído, conflitantes ou, então, dados desconhecidos. Neste último caso, muitas téc-

nicas tem sido aplicadas, sendo algumas delas bastante simples como a substituição dos valores desconhecidos pela média ou mediana do atributo (BATISTA, 2003).

Ainda, é importante ressaltar que a SA já foi explorada com outras intenções além da redução do volume de dados e de evitar o sobreajuste como, por exemplo, a aplicação de técnicas não supervisionadas de SA em mineração de textos (NOGUEIRA, 2009) e a utilização de atributos selecionados como atributos de saída em aprendizado multitarefa (CARUANA; SA, 2003; SWEDEN; CARUANA; SA, 1998).

Neste trabalho, com o auxílio da medida de desempenho multicritério, foi possível identificar as configurações de técnicas mais relevantes para os processos de detecção e de diagnóstico de severidade de defeitos em MITs. Essa medida levou em consideração parâmetros referentes ao processo de SA e de AM. Desta maneira, as medidas de desempenho utilizadas para quantificar as configurações de métodos foram a TFP, TFN, TE, TA, PAS e ACROC. Para esse propósito, a medida multicritério se mostrou simples, ágil e fidedigna aos resultados observados. Ainda, foi possível verificar e analisar a robustez dos sistemas de inferência induzidos por meio da inserção de ruídos desde a construção até o teste dos modelos utilizados.

Desta maneira, para o caso da detecção de defeitos, ou classificação binária, o sistema mais recomendado para essa tarefa possui as seguintes características:

Tipos de atributos: Atributos da categoria estatística, mais especificamente o segundo quartil do sinal de velocidade de vibração axial mecânica;

Desempenho experimental: Configuração do algoritmo *CBF* de seleção de atributos com o algoritmo SVM de aprendizado de máquinas;

Verificação da robustez: A configuração das técnicas *CBF* com a RD foi a mais robusta em termos generalistas desde a construção até o teste;

Análise da robustez: A configuração das técnicas *CBF* com a RD foi a mais robusta para a aplicação específica de detecção de defeitos em MITs com um limite de robustez de 60% de ruídos brancos admitidos.

De acordo com os resultados, observou-se que os algoritmos RD e NB também são adequados para esse processo. Assim, a escolha deve ser feita baseada, primeiramente, na área de aplicação, como em tecnologias embarcadas ou pela internet, e também quanto à TFP, ou taxa de alarmes falsos.

Como discutido na Seção 5.2 é importante elencar a relevância de cada tipo de sinal para o sistema de inferência, sendo de natureza elétrica, sinais de corrente e tensão elétrica, ou mecânica, sinais de velocidade de vibração mecânica. Para isso, os atributos foram organizados de acordo com o tipo de sinal que originou os atributos. Em seguida, foram aplicadas as técnicas de seleção de atributos e de aprendizado de máquinas identificadas

com o melhor desempenho, nas discussões anteriores desse capítulo para os atributos de cada tipo de sinal. Por exemplo, para o caso de detecção de defeitos, classificação binária, foram construídos novos sistemas de inferência constituídos pelo CBF e pelo SVM por meio dos atributos de cada tipo de sinal. Dessa maneira, a Figura 34 apresenta os resultados da avaliação do sistema de inferência constituído pelo CBF e SVM para os atributos oriundos de apenas uma modalidade de sinal, para o cenário binário de classificação. Essa avaliação se deu por meio da medida multicritério e tem a finalidade de ordenar os sinais em função de sua importância e da capacidade preditiva por meio de seus atributos.

Figura 34 – Avaliação do desempenho preditivo dos atributos oriundos de cada tipo de sinal experimental avaliado utilizando o seletor de atributos CBF e o classificador SVM.

Desse mesmo modo, a Tabela 18 complementa a Figura 34 e apresenta os atributos selecionados e a área interna do hexágono multicritério para a avaliação dos atributos em termos da capacidade preditiva para o cenário de detecção de defeitos.

Tipo do sinal	Atributos selecionados	Área do hexágono
Corrente elétrica	Moda_Ic	0,000973
	Razaoruido_Ic	
Tensão elétrica	Curtose_Va	0,394289
Velocidade de vibração mecânica	Quartildois_Vibaxial	0,000822

Tabela 18 – Relação dos atributos selecionados e do desempenho do sistema de inferência composto pelo seletor de atributos CBF e pelo classificador SVM quando utilizado apenas um dos tipos dos sinais experimentais coletados para o sistema de detecção de defeitos, classificação binária.

Com os resultados, observa-se que os sinais de velocidade de vibração mecânica, mais especificamente a velocidade de vibração axial no lado acionado (Vibaxial) são superiores

em termos da medida multicritério. No entanto, o uso de sensores mecânicos apresentam custos elevados, complexidade de instrumentação e são mais vulneráveis à falhas com relação aos sensores elétricos. Esse fato justifica a grande quantidade de estudos na literatura que implementam sensores que coletam sinais elétricos para a detecção e diagnóstico de falhas em MITs.

No que se refere aos sensores elétricos, neste trabalho, os sinais de corrente elétrica apresentaram adequada capacidade preditiva, sendo a área do hexágono multicritério da mesma ordem de grandeza da encontrada utilizando os sinais de velocidade de vibração mecânica. Dessa maneira, para o caso da classificação binária, não restam objeções quanto à utilização dos sinais de origem mecânica, no entanto, os sinais de corrente elétrica podem substituí-los com um desempenho apropriado. Além disso, os sinais de origem elétrica, mais especificamente os sinais de corrente elétrica, apresentam baixa complexidade de instrumentação com relação aos mecânicos, mesmo que precisem de mais atributos para o sistema de inferência.

Para o caso de diagnóstico da severidade dos defeitos, ou classificação multiclasse, o sistema recomendado com base nos resultados deve possuir as seguintes características:

Tipos de atributos: Atributos majoritariamente da categoria estatística, tais como o desvio padrão dos sinais de tensão na fase A e de corrente na fase B, mediana do sinal de vibração axial, terceiro quartil dos sinais de vibração radial do lado acionado e da tensão na fase B. Atributo razão do ruído do sinal de tensão na fase A, pertencente a categoria baseada em informação, e a dimensão fractal do sinal de vibração radial do lado acionado, pertencente a categoria baseada em complexidade;

Desempenho experimental: Configuração do algoritmo *CBF* de seleção de atributos com o algoritmo RD de aprendizado de máquinas;

Verificação da robustez: A configuração dos algoritmos *CBF* com a RD foi a mais robusta em termos generalistas desde a construção até o teste com um limite de 80% de ruídos brancos admitidos;

Análise da robustez: A configuração dos algoritmos *CFS* com a MLP foi a mais robusta para a aplicação específica de diagnóstico de defeitos em MITs com um limite de robustez de 80% de ruídos brancos admitidos.

Outros métodos ficaram bem colocados quanto a área do hexágono multicritério, como a MLP utilizando o *CBF* e o *CFS* como seletores de atributos. Assim, a determinação do método a ser empregado depende da aplicação, sendo que os três métodos citados apresentaram bons resultados em termos de precisão, de sensibilidade e de acordo com as seis variáveis do hexágono multicritério.

Novamente, com o intuito de listar a pertinência entre os tipos de sinais utilizados para extrair os atributos, agora para o caso de classificação multiclasse, os atributos foram organizados de acordo com o tipo de sinal que os originou. Posteriormente, foram aplicadas as técnicas de seleção de atributos e de aprendizado de máquinas identificadas com o melhor desempenho, sendo as técnicas CBF e RD respectivamente, identificadas nas discussões anteriores desse capítulo. Foram construídos novos sistemas de inferência constituídos pelo CBF e pela RD por meio dos atributos de cada tipo de sinal. Dessa maneira, a Figura 35 apresenta os resultados da avaliação do sistema de inferência constituído pelo CBF e RD para os atributos oriundos de apenas uma modalidade de sinal, para o cenário multiclasse de classificação. Não obstante, a Tabela 19 integra os resultados e apresenta os atributos selecionados e a área interna do hexágono multicritério para a avaliação dos atributos em termos da capacidade preditiva para o cenário de diagnóstico da severidade de defeitos.

Figura 35 – Avaliação do desempenho preditivo dos atributos oriundos de cada tipo de sinal experimental avaliado utilizando o seletor de atributos CBF e o classificador RD.

De acordo com os resultados de área interna do hexágono multicritério para o cenário multiclasse do sistema de inferência, nota-se que o desempenho preditivo dos atributos provenientes dos sinais mecânicos são proeminentes com relação aos sinais elétricos, de tensão e corrente. Nesse cenário, de diagnóstico da severidade de defeitos, os sinais de origem mecânica apresentaram maior representatividade quanto ao tipo de defeito, sendo os atributos mais sensíveis e discriminantes com relação a classe. Já os sinais de origem elétrica, apresentaram sistemas de inferência que selecionaram uma quantidade maior de atributos e com desempenho inferior aos sistemas originados por meio de sinais mecânicos.

O baixo desempenho dos sistemas de inferência construídos por meio de sinais elétricos

Tipo do sinal	Atributos selecionados	Área do hexágono
Corrente elétrica	Desviopadrao_Ia	0,024170
	Desviopadrao_Ib	
	Desviopadrao_Ic	
	Mediana_Ib	
	Minimo_Ib	
	Maximo_Ib	
	Quartilum_Ib	
	Moda_Ia	
	Moda_Ic	
	Entropia_Ia	
	Entropia_Ic	
	Ruidotaxadistorcao_Ic	
	Intervalodinamico_Ic	
	Dimensaofractal_Ia	
	Dimensaofractal_Ib	
	Dimensaofractal_Ic	
Tensão elétrica	Media_Va	0,021777
	Mediana_Va	
	Maximo_Va	
	Maximo_Vc	
	Quartiltres_Vb	
	Quartiltres_Vc	
	Moda_Vb	
	Energia_Va	
	Razaoruido_Va	
	Distorcaoharmonica_Vb	
	Ruidotaxadistorcao_Vc	
	Intervalodinamico_Va	
Velocidade de vibração mecânica	Media_Vibaxial	0,007178
	Desviopadrao_Vibacpe	
	Mediana_Vibase	
	Minimo_Vibacpi	
	Quartilum_Vibcarc	
	Quartilum_Vibase	
	Quartildois_Vibaxial	
	Entropia_Vibcarc	
	Distorcaoharmonica_Vibcarc	
	Dimensaofractal_Vibaxial	

Tabela 19 – Relação dos atributos selecionados e do desempenho do sistema de inferência composto pelo seletor de atributos CBF e pelo classificador RD quando utilizado apenas um dos tipos dos sinais experimentais coletados para o sistema de diagnóstico da severidade de defeitos, classificação multiclasse.

foi constatado principalmente, segundo a Figura 35, pelas medidas taxa de falsos positivos, taxa de erro e porcentagem dos atributos selecionados. Em outras palavras, esses sistemas apresentaram taxas de alarmes falsos e de classes incorretamente classificadas superior

em referência ao sistema baseado nos mecânicos.

Entretanto, como explicitado anteriormente, o uso de sensores mecânicos requer características de instrumentação complexa, de manutenção e de maior custo com relação aos elétricos. No que diz respeito aos sensores elétricos para o sistema de inferência multi-classe, o desempenho encontrado não é considerado adequado levando-se em consideração a diferença significativa entre os valores da área interna do hexágono multicritério com relação ao sistema de inferência criado por meio dos sinais mecânicos.

Assim, para solucionar o problema de se usar apenas uma modalidade de sinal, sendo escolhido os sinais elétricos, recomenda-se construir estimadores de atributos mecânicos por meio dos atributos de origem elétrica. Essa hipótese se mostra plausível tendo como suporte a distribuição de probabilidades dos atributos, segundo a Seção 5.2. Com isso, seria possível estabelecer modelos probabilísticos que estimam os atributos mecânicos em função dos atributos elétricos, melhorando o desempenho geral do sistema de inferência.

No que se refere a robustez, tanto para classificação binária quanto para a multiclasse, devem ser observados primariamente a aplicação e o instrumental disponível. Esses fatores podem influenciar na escolha da configuração de técnicas para serem mais ou menos robustas, por meio de parâmetros do sistema físico, como frequência de amostragem e condições de operação.

Dessa maneira, essa pesquisa busca elevar o estado da arte na recomendação de técnicas de detecção e de diagnóstico da severidade de defeitos em MITs, provendo ferramentas validadas estatisticamente em diversas condições para essa finalidade.

Capítulo 6

Conclusão

Neste trabalho foi apresentado um estudo comparativo entre sistemas de detecção e de diagnóstico de severidade de defeitos de barras quebradas em rotores de MITs. No setor industrial, esses equipamentos são amplamente utilizados, portanto, o monitoramento da saúde estrutural traz benefícios econômicos, aumenta a disponibilidade de toda a planta industrial e permite um planejamento otimizado da manutenção.

Para tanto, foi utilizado um banco de dados próprio, adquiridos por meio de uma bancada experimental, e extraídos atributos de diferentes categorias. Ainda, foram investigados alguns dos principais algoritmos de SA que se baseiam em medidas representativas dessas categorias de atributos. Resultados experimentais demonstraram que a SA, por meio da redução da dimensionalidade, pode auxiliar na melhora da qualidade dos dados sob a perspectiva de desempenho preditivo, o que contribui para a construção de modelos de indução mais compreensíveis e com um menor custo computacional.

No decorrer deste trabalho, verificou-se que a pesquisa em SA tem permitido a proposição e o desenvolvimento de diversos algoritmos com o intuito de selecionar atributos importantes, com respeito a um critério de avaliação específico.

Do ponto de vista prático, a escolha pela utilização de um determinado algoritmo de SA, ou um conjunto de algoritmos, deve ser conduzida de acordo com o conhecimento do domínio, em geral detido pelos especialistas do domínio, e o conhecimento de detalhes técnicos sobre os algoritmos de SA, usualmente detido por especialistas da área computacional e das engenharias.

No entanto, ao mesmo tempo em que diversos algoritmos estão disponíveis para a realização da tarefa de SA, cresce a dificuldade em se determinar, *a priori*, qual ou quais desses algoritmos seriam mais apropriados, de acordo com as características do problema e as características dos algoritmos de SA, uma vez que nenhum algoritmo pode ser considerado o melhor independentemente do problema abordado.

Uma das maneiras para auxiliar no problema da seleção de algoritmos é aplicar Meta-Aprendizado. Essa técnica baseia-se no acúmulo de experiência por meio do desempenho de múltiplas aplicações de um determinado método, diferentemente das estratégias tradi-

onais, as quais empregam conhecimento especialista ou processos delongados de avaliação empírica.

Como diferentes algoritmos de aprendizado são baseados em distintos conjuntos de suposições sobre os dados, *bias* ou viés de indução, é importante que haja flexibilidade. Em outras palavras, o algoritmo de aprendizado deve gerar bons modelos se o *bias* dos dados coincidir com o *bias* do algoritmo, além da relevância dos dados. Desse modo, a utilização de meta-dados, ou atributos, como propriedades do problema, propriedades dos algoritmos, baseadas, por exemplo, em medidas de precisão e de robustez, ou ainda padrões previamente construídos, podem auxiliar a selecionar e/ou combinar diferentes métodos.

Para a definição de um sistema de inferência capaz de identificar a condição do defeito, foram utilizadas técnicas de pré-processamento de sinais, técnicas de seleção de atributos e sistemas de aprendizado de máquina, testados com sinais de corrente e tensão elétrica da fase do estator e velocidade de vibração mecânica.

Com a análise dos atributos selecionados, foi possível concluir que os que seguem uma distribuição de probabilidade normal têm maior relevância no sistema de inferência.

Desse modo, nesse trabalho, foi aplicada uma medida multicritério para o desbravamento do melhor conjunto de técnicas para a detecção e o diagnóstico de defeitos em MITs. Com essa medida, tendo como variáveis seis características dos classificadores e dos algoritmos de SA foi possível criar um método para recomendação ferramentas para a resolução desse problema. Ainda, verificou-se características necessárias aos atributos extraídos que conferem maior poder de classificação, robustez e agilidade em aplicações industriais. Tal análise foi realizada utilizando os sinais experimentais e sinais experimentais parcialmente corrompidos com ruído para verificar se o sistema inteligente é robusto.

Ressalta-se na análise de sinais corrompidos que os sinais de velocidade de vibração são menos sensíveis ao ruído, indicando que possuem mais informações sobre a saúde estrutural do MIT. Isso contrasta com o grande número de pesquisas documentadas na literatura relacionadas à detecção e diagnóstico de barras quebradas em motores de indução trifásicos que utilizam sinais de corrente elétrica. Este trabalho contribui para mostrar o potencial dos sinais mecânicos para essa finalidade.

Nesse contexto, durante o desenvolvimento desse trabalho, surgiram novos caminhos de investigação e são listados como trabalhos futuros:

- ❑ Avaliar a influência de cada categoria de atributos no sistema preditivo por meio da medida multicritério;
- ❑ Aplicar a metodologia concebida nessa pesquisa para outros bancos de dados, mais especificamente em motores acionados por meio de inversores de frequência;
- ❑ Empregar a metodologia de verificação e análise de robustez em outros conjuntos de dados para fins de comparação;

- ❑ Modelar um aproximador de funções baseado nos sinais elétricos para estimar os atributos mecânicos.

Referências

ABEEL, T. et al. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. **Bioinformatics**, 2010. v. 26, n. 3, p. 392–398, 2010.

ANTONINO-DAVIU, J. et al. Detection of broken outer-cage bars for double-cage induction motors under the startup transient. **IEEE Transactions on Industry Applications**, 2012. 2012. ISSN 00939994.

ARAUZO-AZOFRA, A.; BENITEZ, J. M.; CASTRO, J. L. Consistency measures for feature selection. **Journal of Intelligent Information Systems**, 2008. Kluwer Academic Publishers, v. 30, n. 3, p. 273–292, 2008.

BACCARINI, L. M. R. e. a. Svm practical industrial application for mechanical faults diagnostic. **Expert Systems with Applications**, 2011. v. 38, p. 6980–6984, 2011.

BATISTA, G. E. A. P. A. Pré-processamento de dados em aprendizado de máquina supervisionado. Tese de Doutorado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. 2003.

BATISTA, O.; AL. et. Methodology for information extraction from oscillograms and its application for high-impedance faults analysis. **International Journal of Electrical Power and Energy Systems**, 2016. v. 76, p. 23–34, 2016.

BENBOUZID, M. E. H.; HACHEMI, E. A review of induction motors signature analysis as a medium for faults detection. **IEEE Transactions on Industrial Electronics**, 2000. v. 47, p. 984–993, 2000.

BENBOUZID, M. E. H.; KILMAN, G. B. What stator current processingbased technique to use for induction motor rotor faults diagnosis. **IEEE Transactions on Energy Conversion**, 2003. v. 18, p. 238–244, 2003.

BHARGAVA, N. et al. Decision tree analysis on j48 algorithm for data mining. **Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering**, 2013. v. 3, n. 6, 2013.

BINDU, S.; THOMAS, V. Diagnoses of internal faults of three phase squirrel cage induction motor-a review. **International Conference on Advances in Energy Conversion Technologies (ICAECT)**, 2014. p. 48–54, 2014.

BOSSIO, G. e. a. Separating broken rotor bars and load oscillations on im fault diagnosis through the instantaneous active and reactive currents. **IEEE Transactions on Industrial Electronics**, 2009. v. 56, p. 4571–4580, 2009.

BRAZDIL, P. B. et al. **Metalearning: applications to data mining**. Berlin: Springer-Verlag, 2009.

CAMPOS, T. P. de. Uma ferramenta para apoiar a pesquisa exploratória na *Web*. Dissertação de Mestrado, Universidade Estadual de Maringá, Departamento de Informática. 2013.

CARUANA, R.; SA, V. R. D. Benefitting from the variables that variable selection discards. **Journal of Machine Learning Research**, 2003. JMLR.org, v. 3, p. 1245–1264, 2003.

CHISSOM, B. S. Interpretation of the kurtosis statistic. **The American Statistician**, 1970. v. 24, n. 4, p. 19–22, 1970.

CHOQUEUSE, V. et al. Monitoring of three phase signals based on singular value decomposition. **IEEE Transactions on Smart Grid**, 2019. v. 10, n. 6, p. 6156–6166, 2019.

CHUNG, H. M.; GRAY, P. Data mining. **Journal of management information systems**, 1999. Taylor & Francis, v. 16, n. 1, p. 11–16, 1999.

CLAVERIA, O.; TORRA, S. Forecasting tourism demand to catalonia: Neural networks vs. time series models. **Economic Modelling**, 2014. v. 36, p. 220–228, 2014.

COBLE, J.; HINES, J. W. Identifying optimal prognostic parameters from data: a genetic algorithms approach. **Annual Conference of the Prognostics and Health Management Society**, 2009. 2009.

COSTA, C.; KASHIWAGI, M.; MATHIAS, M. H. Rotor failure detection of induction motors by wavelet transform and fourier transform in non-stationary condition. **Mechanical Systems and Signal Processing**, 2015. v. 1, p. 15–26, 2015.

COVÕES, T. F. Seleção de atributos via agrupamento. Dissertação de Mestrado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. 2010.

DAS, S. Filters, wrappers and a boosting based hybrid for feature selection. In: **International Conference on Machine Learning**. [S.l.]: Williams College, 2001. p. 74–81.

DASH, M.; LIU, H. Feature selection for clustering. In: **Pacific-Asia Conference on Knowledge Discovery and Data Mining**. [S.l.: s.n.], 2000. p. 110–121.

DASH, M.; LIU, H. Consistency-based search in feature selection. **Artificial Intelligence**, 2003. Elsevier Science Inc., v. 151, n. 1-2, p. 155–176, 2003.

DIAO, L. et al. Optimized design of discrete traction induction motor model at low-switching frequency. **IEEE Transactions on Power Electronics**, 2013. v. 28, n. 10, p. 4803–4810, 2013.

- DOMINGOS, P. Why does bagging work? A bayesian account and its implications. In: **International Conference on Knowledge Discovery and Data Mining**. [S.l.: s.n.], 1997. p. 155–158.
- DRIF, M.; CARDOSO, A. J. M. Stator fault diagnostics in squirrel cage three-phase induction motor drives using instantaneous active and reactive power signature analyses. **IEEE Transactions on Industrial Informatics**, 2014. v. 10, p. 1348–1360, 2014.
- DY, J. G.; BRODLEY, C. E. Feature subset selection and order identification for unsupervised learning. In: **International Conference on Machine Learning**. [S.l.: s.n.], 2000. p. 247–254.
- EBRAHIMI, B. M.; FAIZ, J.; AKIN, B. Pattern recognition for broken bars fault diagnosis in induction motors under various supply conditions. **European Transactions on Electrical Power**, 2012. v. 22, p. 1176–1190, 2012.
- EREN, L.; DEVANEY, M. J. Bearing damage detection via wavelet packet decomposition of the stator current. **IEEE Transactions on Instrumentation and Measurement**, 2004. v. 53, n. 2, p. 431–436, 2004.
- FALOUTSOS, C.; KAMEL, I. Beyond uniformity and independence: analysis of R-trees using the concept of fractal dimension. In: **Proceedings of the Thirteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems**. Minneapolis, United States of America: ACM, 1994. p. 4–13.
- FILIPPETTI, F.; FRANCESCHINI, G.; TASSONI, C. Neural networks aided on-line diagnostics of induction motor faults. **IEEE Industry Applications Society**, 1993. p. 316–323, 1993.
- FILIPPETTI, F. et al. Ai techniques in induction machines diagnosis including the speed ripple effect. **IEEE Industry Applications Society**, 1996. p. 655–662, 1996.
- FILIPPETTI, F. et al. Recent developments of induction motor drvies fault diagnosis using ai techniques. **IEEE Transactions on Industrial Electronics**, 2000. v. 47, p. 994–1004, 2000.
- FLAUZINO, R. A. Identificação e localização de faltas de alta impedância em sistemas de distribuição baseadas em decomposição por componentes ortogonais e inferência fuzzy. Tese de Doutorado, Universidade de São Paulo. 2007.
- GARCIA-ESCUDEIRO, L. A. e. a. Robust condition monitoring for early detection of broken bars in induction motors. **Expert Systems with Applications**, 2011. v. 38, p. 2653–2660, 2011.
- GARCIA-PEREZ, A. e. a. The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors. **IEEE Transactions on Industrial Electronics**, 2011. v. 58, p. 2002–2010, 2011.
- GEROMEL, J. C.; KOROGUI, R. H. **Controle linear de sistemas dinâmicos 2ª edição: teoria, ensaios práticos e exercícios**. [S.l.]: Editora Blucher, 2019.
- GHANBARI, T. Autocorrelation function-based technique for stator turn-fault detection of induction motor. **IET Science, Measurement Technology**, 2016. v. 10, p. 100–110, 2016.

- GHORBANIAN, V.; FAIZ, J. A. A survey on time and frequency characteristics of induction motors with broken rotor bars in line-start and inverter-fed modes. **Mechanical Systems and Signal Processing**, 2015. p. 427–456, 2015.
- GODOY, W. F. e. a. Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter. **IET Electrical Power Applications**, 2016. v. 10, p. 430–439, 2016.
- GRITLI, Y. et al. Experimental validation of doubly fed induction machine electrical faults diagnosis under time-varying conditions. **Electric Power Systems Research**, 2011. 2011. ISSN 03787796.
- GU, F. e. a. A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals. **Mechanical Systems and Signal Processing**, 2015. v. 50, p. 400–413, 2015.
- GUNER, A.; USTUNDAG, M. Performance analysis of fourth order cumulant method in direction of arrival estimation. In: **2015 23rd Signal Processing and Communications Applications Conference (SIU)**. [S.l.: s.n.], 2015. p. 1693–1696.
- GUO, Z. et al. A new fault diagnosis classifier for rolling bearing united multi scale permutation entropy optimize vmd and cuckoo search svm. **IEEE Access**, 2020. v. 8, p. 153610–153629, 2020.
- GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. **Journal of Machine Learning**, 2003. v. 3, p. 1157–1182, 2003.
- HALL, M. **Correlation-based feature subset selection for machine learning**. 178 p. Tese (Doutorado) — Department of Computer Science, University of Waikato, 1999.
- HALL, M. A. Correlation-based feature selection for discrete and numeric class machine learning. In: **International Conference on Machine Learning**. [S.l.: s.n.], 2000. p. 359–366.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques**. 3. ed. California, United States of America: Morgan Kaufmann, 2011.
- HAND, D. J.; ADAMS, N. M. Data mining. **Wiley StatsRef: Statistics Reference Online**, 2014. Wiley Online Library, p. 1–7, 2014.
- HE, X.; CAI, D.; NIYOGI, P. Laplacian score for feature selection. In: **Advances in Neural Information Processing Systems**. [S.l.: s.n.], 2005. p. 507–514.
- HENAO, H. et al. Trends in fault diagnosis for electrical machines: A review of diagnostic techniques. **IEEE Industrial Electronics Magazine**, 2014. 2014. ISSN 19324529.
- HO, T.; BASU, M.; LAW, M. Measures of geometrical complexity in classification problems. In: **Data complexity in pattern recognition**. London, England: Springer, 2006, (Advanced Information and Knowledge Processing). p. 1–23.
- HO, T. K.; BASU, M. Complexity measures of supervised classification problems. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2002. IEEE Computer Society, Washington, United States of America, v. 24, n. 3, p. 289–300, 2002.

- JIANG, L.; CHEN, Z.; MAO, J. On the numerical stability of the precise integration time-domain (pitd) method. **IEEE Microwave and Wireless Components Letters**, 2007. v. 17, n. 7, p. 471–473, 2007.
- JIANG, L.; ZHANG, H.; CAI, Z. A novel bayes model: Hidden naive bayes. **IEEE Transactions on knowledge and data engineering**, 2008. IEEE, v. 21, n. 10, p. 1361–1371, 2008.
- KALOUSIS, A.; GAMA, J.; HILARIO, M. On data and algorithms: understanding inductive performance. **Machine Learning**, 2004. v. 54, p. 275–312, 2004.
- KANG, T. et al. Electrical monitoring of mechanical defects in induction motor driven v belt pulley speed reduction couplings. **IEEE Transactions on Industry Applications**, 2018. v. 54, n. 3, p. 2255–2264, 2018.
- KIM, K.; PARLOS, A. G. Induction motor fault diagnosis based on neuropredictors and wavelet signal processing. **IEEE/ASME Transactions on Mechatronics**, 2002. v. 7, p. 201–219, 2002.
- KIM, K.; PARLOS, A. G.; BHARADWAJ, R. M. Sensorless fault diagnosis of induction motors. **IEEE Transactions on Industrial Electronics**, 2003. v. 50, p. 1038–1051, 2003.
- KITCHENHAM, B. A.; CHARTERS, S. Technical report, **Guidelines for performing systematic literature reviews in software engineering, evidence-based software engineering**. 2007.
- KLIMAN, G. B.; STEIN, J. Induction motor fault detection via passive current monitoring. **International Conference on Electrical Machines**, 1990. p. 13–17, 1990.
- KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. **Artificial Intelligence**, 1997. Elsevier Science Inc., Essex, United Kingdom, v. 97, n. 1-2, p. 273–324, 1997.
- KONONENKO, I. Estimating attributes: analysis and extention of Relief. In: **European Conference on Machine Learning**. Amsterdam: Springer-Verlag, 1994. p. 171–182.
- LATHI, B. P. **Sinais e Sistemas Lineares**. [S.l.]: Bookman, 2006.
- LEE, H. D. Seleção de atributos importantes para a extração de conhecimento de bases de dados. Tese de Doutorado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. 2005.
- LEE, H. D.; MONARD, M. C. Seleção de atributos para algoritmos de aprendizado de máquina supervisionado utilizando como filtro a dimensão fractal. **Revista de La Sociedad Chilena de Ciencia de La Computación**, 2003. v. 4, n. 1, p. 1–8, 2003.
- LEE, H. D.; MONARD, M. C.; WU, F. C. A fractal dimension based filter algorithm to select features for supervised learning. **Lecture Notes in Computer Science**, 2006. v. 4140, p. 278–288, 2006.

- LEE, J. W.; GIRAUD-CARRIER, C. Predicting algorithm accuracy with a small set of effective meta-features. In: **International Conference on Machine Learning and Applications**. [S.l.: s.n.], 2008. p. 808–812.
- LIBONI, L. Diagnóstico de falhas em motores de indução trifásicos baseado em decomposição em componentes ortogonais e aprendizagem de máquinas. Tese de Doutorado, Universidade de São Paulo. 2017.
- LIMA, P. C.; LIMA, R. R. Estatística experimental. **Guia de Estudos, Universidade Federal de Lavras**, 2014. v. 20, n. 06, 2014.
- LIU, H.; MOTODA, H. **Feature selection for knowledge discovery and data mining**. Massachusetts, United States of America: Kluwer Academic Publishers, 1998.
- LIU, H.; MOTODA, H. **Computational methods of feature selection**. Minnesota, United States of America: Chapman & Hall/CRC data mining and knowledge discovery, 2008.
- LIU, H.; MOTODA, H.; YU, L. A selective sampling approach to active feature selection. **Artificial Intelligence**, 2004. v. 159, p. 49–74, 2004.
- LIU, H.; YU, L. Feature selection for data mining. Department of Computer Science and Engineering, Arizona State University. 2002.
- LIU, H.; YU, L. Toward integrating feature selection algorithms for classification and clustering. **IEEE Transactions on Knowledge and Data Engineering**, 2005. v. 17, p. 491–502, 2005.
- LIU, H. et al. Active feature selection using classes. In: **Pacific-Asia Conference on Knowledge Discovery and Data Mining**. [S.l.: s.n.], 2003. p. 474–485.
- LOPES, T. D. et al. Análise comparativa da detecção de defeitos em rolamentos de motores de indução trifásicos utilizando rbf e sfam. **Simpósio Brasileiro de Automação Inteligente**, 2015. v. 209, 2015.
- MACIEJEWSKI, N. A.; TREML, A. E.; FLAUZINO, R. A. A systematic review of fault detection and diagnosis methods for induction motors. In: **2020 FORTEI-International Conference on Electrical Engineering (FORTEI-ICEE)**. Bandung, Indonesia: IEEE, 2020.
- MAITRE, J. et al. Classification algorithms comparison for interturn short circuit recognition in induction machines using best-fit 3d ellipse method. **Canadian Journal of Electrical and Computer Engineering**, 2017. v. 40, n. 4, p. 255–265, 2017.
- MALEK, M. Abd-el; ABDELSALAM, A. K.; HASSAN, O. E. Induction motor broken rotor bar fault location detection through envelope analysis of start-up current using hilbert transform. **Mechanical Systems and Signal Processing**, 2017. 2017. ISSN 10961216.
- MATIC, D. e. a. Support vector machine classifier for diagnosis in electrical machines: Application to broken bar. **Expert Systems with Applications**, 2012. v. 39, p. 8681–8689, 2012.

- MEHRJOU, M. R. e. a. Rotor fault condition monitoring techniques for squirrel-cage induction machine-a review. **Mechanical Systems and Signal Processing**, 2011. v. 25, p. 2827–2848, 2011.
- MITRA, P.; MURTHY, C. A.; PAL, S. K. Unsupervised feature selection using feature similarity. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2002. IEEE Computer Society, v. 24, n. 3, p. 301–312, 2002. ISSN 0162-8828.
- MOLLINEDA, R. A.; SÁNCHEZ, J. S.; SOTOCA, J. M. Data characterization for effective prototype selection. In: **Proceedings of the Second Iberian Conference on Pattern Recognition and Image Analysis**. Berlin, Heidelberg: Springer, 2005. p. 27–34.
- NANDI, S.; TOLİYAT, H. A.; LI, X. Condition monitoring and fault diagnosis of electrical motors-a review. **IEEE Transactions on Energy Conversion**, 2005. v. 20, p. 719–729, 2005.
- NASCIMENTO, F. A.; BERGER, P. A.; ROCHA, A. F. da. Algoritmo para cálculo da distorção harmônica por meio da transformada rápida de hartley. **a a**, 2007. v. 60, p. 4, 2007.
- NOGUEIRA, B. M. **Avaliação de métodos não-supervisionados de seleção de atributos para mineração de textos**. 2009. Dissertação de Mestrado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo.
- NUNES, C. S. Gestão em educação a distância: um framework baseado em boas práticas. Tese de Doutorado, Universidade Federal de Santa Catarina, Departamento de Engenharia do Conhecimento. 2017.
- OSTOJIC, P. e. a. Advanced motor monitoring and diagnostics. **Industry Application Society 60th Annual Petroleum and Chemical Industry Conference**, 2013. p. 1–9, 2013.
- PALÁCIOS, R. H. C. et al. Reconhecedor neural de defeitos no estator em motores de indução trifásicos apoiado por análise de componentes principais. **XII Simpósio Brasileiro de Automação Inteligente (SBAI), SBA**, 2015. p. 1206–1211, 2015.
- PARMEZAN, A. R. S. Proposta de um módulo de monitoramento de qualidade de dados com assistência inteligente: um estudo de caso para o sistema médico de auxílio à cirurgia coloproctológica. Monografia de Graduação, Centro de Engenharias e Ciências Exatas, Universidade Estadual do Oeste do Paraná. 2012.
- PARMEZAN, A. R. S. et al. Estudo comparativo entre métodos de seleção de atributos baseados em medidas de precisão e correlação aplicados a bases de dados. **XVIII Simpósio Internacional de Iniciação Científica da Universidade de São Paulo**, 2010. p. 1–1, 2010.
- PARMEZAN, A. R. S.; LEE, H. D.; WU, F. C. Redução da dimensionalidade em bases de dados naturais através de métodos de filtro para seleção de atributos importantes. **XIX Simpósio Internacional de Iniciação Científica da Universidade de São Paulo**, 2011. p. 1–1, 2011.

- PARMEZAN, A. R. S.; LEE, H. D.; WU, F. C. Estudo preliminar da construção de um modelo de recomendação de algoritmos de seleção de atributos utilizando meta-aprendizado. **XX Simpósio Internacional de Iniciação Científica da Universidade de São Paulo**, 2012. p. 1–1, 2012.
- PARMEZAN, A. R. S.; LEE, H. D.; WU, F. C. Metalearning for choosing feature selection algorithms in data mining: Proposal of a new framework. **Expert Systems with Applications**, 2017. Elsevier, v. 75, p. 1–24, 2017.
- PARMEZAN, A. R. S.; WU, F. C.; LEE, H. D. Estudo de medidas de importância e algoritmos para seleção de atributos para mineração de dados. **XX Encontro Anual de Iniciação Científica**, 2011. p. 1–4, 2011.
- PARMEZAN, A. R. S.; WU, F. C.; LEE, H. D. Meta-aprendizado no auxílio à seleção de atributos: um estudo para medidas de correlação e consistência. **XXI Encontro Anual de Iniciação Científica**, 2012. p. 1–4, 2012.
- PAWLAK, M.; KOWALSKI, C. Low-cost embedded system for the im fault detection using neural networks. **XIX International Conference on Electrical Machines (ICEM)**, 2010. v. 19, p. 1–5, 2010.
- PENMAN, J.; YIN, C. M. Feasibility of using unsupervised learning, artificial neural networks for the condition monitoring of electrical machines. **Proc. Inst. Elect. Eng., Elect. Power Appl.**, 1994. v. 141, p. 317–322, 1994.
- PILA, A. D. **Seleção de atributos relevantes para aprendizado de máquina utilizando a abordagem de rough sets**. 2001. Dissertação de Mestrado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo.
- PONS-LLINARES, J. et al. Induction motor diagnosis based on a transient current analytic wavelet transform via frequency b-plines. **IEEE Transactions on Industrial Electronics**, 2011. 2011. ISSN 02780046.
- PONS-LLINARES, J. et al. Pursuing optimal electric machines transient diagnosis: The adaptive slope transform. **Mechanical Systems and Signal Processing**, 2016. 2016. ISSN 10961216.
- PUCCINI, L. R. et al. Comparativo entre as bases de dados pubmed, scielo e google acadêmico com o foco na temática educação médica. **Cadernos UniFOA, Volta Redonda**, 2015. v. 28, p. 75–82, 2015.
- RANGEL-MAGDALENO, J. J. e. a. Fpga-based broken bars detection on induction motors under different load using motor current signature analysis and mathematical morphology. **IEEE Transactions on Instrumentation and Measurement**, 2014. v. 63, p. 1032–1040, 2014.
- RAZIK, H.; CORREA, M. R.; SILVA, E. A novel monitoring of load level and broken bar fault severity applied to squirrel-cage induction motors using a genetic algorithm. **IEEE Transactions on Industrial Electronics**, 2009. v. 56, p. 4615–4626, 2009.
- RISTANOSKI, G.; LIU, W.; BAILEY, J. A time-dependent enhanced support vector machine for time series regression. **International Conference on Knowledge Discovery and Data Mining**, 2013. p. 946–954, 2013.

- ROBNIK-SIKONJA, M.; KONONENKO, I. Theoretical and empirical analysis of ReliefF and RReliefF. **Machine Learning**, 2003. Kluwer Academic Publishers, Massachusetts, v. 53, n. 1-2, p. 23–69, 2003.
- ROCHA, E. M. et al. A fuzzy type-2 fault detection methodology to minimize false alarm rate in induction motor monitoring applications. **Applied Soft Computing Journal**, 2020. 2020. ISSN 15684946.
- ROMERO-TRANCOSO, R. e. a. Fpga-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference. **IEEE Transactions on Industrial Electronics**, 2011. v. 58, p. 5263–5270, 2011.
- ROMERO-TRONCOSO, R. J. et al. Fpga-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference. **IEEE Transactions on Industrial Electronics**, 2011. v. 58, n. 11, p. 5263–5270, 2011.
- SADEGHIAN, A.; YE, Z.; WU, B. Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks. **IEEE Transactions on Instrumentation and Measurement**, 2009. v. 58, p. 2253–2263, 2009.
- SANTOS, F. M. C.; SILVA, I. N.; SUETAKE, M. Sobre a aplicação de sistemas inteligentes para diagnóstico de falhas em máquinas de indução - uma visão geral. **Revista Controle e Automação, Setembro e Outubro**, 2012. v. 23, 2012.
- SANTOS, F. M. d. C. Identificação de falhas em motores de indução trifásicos usando sistemas inteligentes. Tese de Doutorado, Universidade de São Paulo. 2013.
- SARITAS, M. M.; YASAR, A. Performance analysis of ann and naive bayes classification algorithm for data classification. **International Journal of Intelligent Systems and Applications in Engineering**, 2019. v. 7, n. 2, p. 88–91, 2019.
- SCHLIMMER, J. C. Efficiently inducing determinations: a complete and systematic search algorithm that uses optimal pruning. In: **International Conference on Machine Learning**. [S.l.: s.n.], 1993. p. 284–290.
- SCHOEN, R. R. et al. An unsupervised on-line system for induction motor fault detection using stator current monitoring. **IEEE Transactions on Industrial Application**, 1995. v. 31, p. 1280–1286, 1995.
- SEERA, M. e. a. Condition monitoring of induction motors: A review and application of an ensemble of hybrid intelligent models. **Expert Systems with Applications**, 2014. v. 41, p. 4891–4903, 2014.
- SERONATO, A. L.; NETO, P. M. **Detecção de defeitos em motores de indução utilizando a análise das correntes de linha**. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2015.
- SILVA, A.; POVINELLI, R.; DEMERDASH, N. Rotor bar fault monitoring method based on analysis of air-gap torques of inducton motors. **IEEE Transactions on Industrial Informatics**, 2013. v. 9, p. 2274–2283, 2013.

SINGH, G.; NAIKAN, V. N. Detection of half broken rotor bar fault in vfd driven induction motor drive using motor square current music analysis. **Mechanical Systems and Signal Processing**, 2018. 2018.

SOUALHI, A.; CLEC, G.; RAZIC, H. Detection and diagnosis of faults in induction motor using an improved artificial ant clustering technique. **IEEE Transactions on Industrial Electronics**, 2013. v. 60, p. 4053–4062, 2013.

SOUSA, E. P. M. et al. How to use fractal dimension to find correlations between attributes. In: **Workshop on Fractals and Self-similarity in Data Mining: Issues and Approaches**. [S.l.: s.n.], 2002. p. 26–30.

SOUZA, B. F. Meta-aprendizagem aplicada à classificação de dados de expressão gênica. Tese de Doutorado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. 2010.

SPOLAÔR, N. Aplicação de algoritmos genéticos multiobjetivo ao problema de seleção de atributos. Dissertação de Mestrado, Universidade Federal do ABC. 2010.

SPOLAÔR, N.; LORENA, A. C.; LEE, H. D. Uma revisão sistemática sobre aplicações de metaheurísticas multiobjetivo ao problema de seleção de atributos. 2010. 2010. Relatório Técnico, Universidade Federal do ABC.

SPOLAÔR, N.; LORENA, A. C.; LEE, H. D. Algoritmos genéticos multi-objetivo para a seleção de atributos. **VIII Encontro Nacional de Inteligência Artificial**, 2011. Natal, p. 938–949, 2011.

SWEDEN, S.; CARUANA, R.; SA, V. R. D. Using feature selection to find inputs that work better as extra outputs. In: **International Conference on Artificial Neural Networks**. [S.l.: s.n.], 1998. p. 1–6.

TALAVERA, L. Feature selection as a preprocessing step for hierarchical clustering. In: **International Conference on Machine Learning**. [S.l.: s.n.], 1999. p. 389–397.

TRAINA, C.; SOUSA, E. P. M.; TRAINA, A. J. M. **Using fractals in data mining**. New Jersey, United States of America: Wiley-IEEE Press, 2005. 599–630 p.

TRAINA, C. et al. Fast feature selection using fractal dimension. In: **Brazilian Data Base Symposium**. [S.l.: s.n.], 2000. p. 158–171.

TREML, A. E. et al. Experimental database for detecting and diagnosing rotor broken bar in a three-phase induction motor. **IEEE Dataport**, 2020. IEEE Dataport, 2020.

TSOUMAKAS, G.; KATAKIS, I. Multi-label classification: an overview. **International Journal of Data Warehousing and Mining**, 2007. v. 2007, p. 1–13, 2007.

VAS, P. **Parameter estimation, condition monitoring, and diagnosis of electrical machines**. Oxford, UK: Clarendon, 1993.

VAS, P. **Intelligence-Based Electrical Machines and Drives: Applications of Fuzzy, Neural, Fuzzy-Neural and Genetic Algorithm Based Techniques**. New York, USA: Oxford Univ. Press, 1999.

- VEDREÑO-SANTOS, F. et al. Diagnosis of rotor and stator asymmetries in wound-rotor induction machines under nonstationary operation through the instantaneous frequency. **IEEE Transactions on Industrial Electronics**, 2014. 2014. ISSN 02780046.
- VILALTA, R.; DRISSI, Y. A perspective view and survey of meta-learning. **Artificial Intelligence Review**, 2002. v. 18, p. 77–95, 2002.
- WANG, C.-M.; HUANG, Y.-F. Evolutionary-based feature selection approaches with new criteria for data mining: a case study of credit approval data. **Expert Systems with Applications**, 2009. v. 36, n. 3-2, p. 5900–5908, 2009.
- WITTEN, I. H.; FRANK, E. **Data mining: practical machine learning tools and techniques**. California, United States of America: Morgan Kaufmann, 2005.
- YAN, L. et al. Robustness improvement of fcs mptc for induction machine drives using disturbance feedforward compensation technique. **IEEE Transactions on Power Electronics**, 2019. v. 34, n. 3, p. 2874–2886, 2019.
- YANG, C. et al. Reliable detection of induction motor rotor faults under the rotor axial air duct influence. **IEEE Transactions on Industry Applications**, 2014. 2014. ISSN 00939994.
- YANG, T. e. a. Feature knowledge based fault detection of induction motors through the analysis of stator current data. **IEEE Transactions on Instrumentation and Measurement**, 2016. v. 65, p. 549–558, 2016.
- YU, L.; LIU, H. Efficient feature selection via analysis of relevance and redundancy. **Journal of Machine Learning Research**, 2004. v. 5, p. 1205–1224, 2004.
- ZAMUDIO-RAMIREZ, I. et al. Detection of winding asymmetries in wound-rotor induction motors via transient analysis of the external magnetic field. **IEEE Transactions on Industrial Electronics**, 2020. v. 67, n. 6, p. 5050–5059, 2020.
- ZBEDE, Y. B.; GADOUE, S. M.; ATKINSON, D. J. Model predictive mras estimator for sensorless induction motor drives. **IEEE Transactions on Industrial Electronics**, 2016. v. 63, n. 6, p. 3511–3521, 2016.
- ZHANG, X. et al. Applications and comparisons of four time series models in epidemiological surveillance data. **PLOS ONE**, 2014. v. 9, 2014.

Apêndices

APÊNDICE A

Descrição dos Atributos Extraídos

Para definir os atributos, consideremos que são estimadas a partir de um conjunto de dados D (ou parte dele) organizado de acordo com o formato de arquivo de atributo-valor, como retratado na Tabela 20.

Tabela 20 – Formato atributo-valor

Exemplos	Atributos				Classe ($\mathbb{C}$)
	A_1	A_2	$\dots$	A_M	
E_1	a_{11}	a_{12}	$\dots$	a_{1M}	c_1
E_2	a_{21}	a_{22}	$\dots$	a_{2M}	c_2
E_3	a_{31}	a_{32}	$\dots$	a_{3M}	c_3
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
E_N	a_{N1}	a_{N2}	$\dots$	a_{NM}	c_N

De acordo com o formato de arquivo de entrada mostrado na Tabela 20, um conjunto de dados contém N pares de exemplos (E_i, c_i) , onde $E_i = (a_{i1}, a_{i2}, \dots, a_{iM})$ e $c_i \in \{1, \dots, C\}$. Isso é, cada exemplo E_i é descrito por M atributos preditivos (características) e tem um rótulo c_i de C classes. O atributo classe, por sua vez, representa o conceito a ser aprendido e descrito pelos modelos construídos usando algoritmos de aprendizado de máquina supervisionado.

A.1 Medidas Estatísticas

Uma das formas mais comuns de caracterizar um conjunto de dados consiste em extrair medidas de uma área da estatística bem conhecida, a estatística descritiva. Muitas dessas medidas têm um baixo custo computacional e pode ser usado para verificar a ordem, descrição e distribuição de dados numéricos.

As medidas estatísticas descritivas pressupõem que um processo estatístico gera os dados. Como diferentes parâmetros, frequentemente, descrevem o processo, as medidas podem ser visualizadas como estimativas de parâmetros estatísticos da distribuição que originou os dados. As medidas média, desvio padrão, variância, covariância, mediana,

moda, valores extremos, soma dos valores e quartis são as descritas pela estatística clássica (LIMA; LIMA, 2014).

Curtose:

A medida de curtose expressa o grau de achatamento de uma distribuição, que na maioria dos casos é considerada em relação a uma distribuição normal (ou gaussiana). A forma da curva de distribuição de acordo com a curtose pode ser categorizada como mesocúrtica, platicúrtica ou leptocúrtica (Figura 36).

Figura 36 – Categorias genéricas de curtose considerando a distribuição gaussiana.

A curva gaussiana, que é um modelo teórico referencial de distribuição, é chamada de mesocúrtica. Quando a distribuição tem uma curva de frequência mais aberta quando comparada à gaussiana, é denominada platicúrtica. Em contraste, quando a curva de frequência de distribuição tem uma curva de frequência mais fechada do que a Gaussiana, ela é chamada de leptocúrtica (CHISSOM, 1970).

Em termos práticos, a caracterização do nivelamento é significativa apenas quando a distribuição é aproximadamente simétrica. Assim, para um atributo numérico A com valores N , o coeficiente de curtose é dado pelo quociente do momento de quarta ordem (m_4) pelo quadrado da variância (s^2) como mostrado na Equação 1.

$$Curtose(A) = \frac{m_4}{(s^2)^2} \quad (1)$$

O momento de quarta ordem centrado na média é definido de acordo com a Equação 2.

$$m_4 = \frac{\sum_{i=1}^N (a_i - \bar{A})^4}{N} \quad (2)$$

Assimetria:

O coeficiente de assimetria de Pearson, também chamado de primeiro coeficiente de assimetria de Pearson, é definido de forma não dimensionada pela Equação 3.

Esta formulação matemática visa resumir quantitativamente a assimetria de uma distribuição.

$$Assimetria(A) = \frac{3 \times (\bar{A} - Md)}{s} \quad (3)$$

Na Equação 3, de um atributo numérico A , $\bar{A}$ representa a média, Md indica a mediana e s é o desvio padrão da amostra. O coeficiente busca caracterizar como e quanto a distribuição dos valores de um atributo se afasta da condição de simetria. Dependendo do valor assumido por esta medida estatística, é possível categorizar a distribuição como simétrica, assimétrica moderada ou assimétrica forte.

A.2 Medidas Baseadas em Informação

Medidas baseadas na teoria da informação buscam caracterizar os atributos nominais e sua relação com o atributo de classe. As medidas de informação mais populares são descritas abaixo.

Entropia:

A entropia da distribuição das classes em um conjunto de dados D , definida pela Equação 4, origina-se da área de teoria da comunicação. Indica a quantidade aproximada de informações necessárias para identificar o rótulo da classe de um exemplo do conjunto de dados original (HAN; KAMBER; PEI, 2011; MITRA; MURTHY; PAL, 2002).

$$Entropy(D) = - \sum_{i=1}^n [p_i \times \log_2(p_i)] \quad (4)$$

Na Equação 4 p_i , $1 \leq i \leq n$, representa a probabilidade de ocorrência de cada c_i da classe C , dada pela razão entre o número de casos em que c_i ocorre e o número total de exemplos. Além disso, a função log de base 2 é usada porque a informação é codificada em bits.

Energia e Potência do Sinal:

A energia e a potência traz informações quanto a capacidade e a morfologia de um sinal. A Equação 5 é referente a energia, e a Equação 6 a potência para um sinal discreto $x(n)$, com n sendo o intervalo discreto entre os dados do sinal e N o período que se deseja considerar (LATHI, 2006).

$$Energia(x) = \sum_{n=-\infty}^{\infty} [x^2[n]] \quad (5)$$

$$Potencia(x) = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{n=-N}^N [x^2[n]] \quad (6)$$

Contraste, Homogeneidade e Momento de 3ª Ordem do Sinal:

O contraste de um sinal é a informação que depende da posição dos dados na série temporal. Essa medida é utilizada na análise imagens, no entanto, nesse trabalho utilizamos para séries temporais. Para o cálculo dessa medida é realizado o somatório da multiplicação de cada intervalo de tempo discreto n pelo valor da variável x correspondente até o fim da série temporal T , segundo a Equação 7 (LATHI, 2006).

$$Contraste(x) = \sum_{n=0}^T [n^2 x[n]] \quad (7)$$

A homogeneidade de um sinal descreve como o sinal está distribuído no seu eixo independente, ou nesse caso, no decorrer do tempo de aquisição. A Equação 8 descreve os aspectos matemáticos de cálculo (LATHI, 2006).

$$Homogeneidade(x) = \sum_{n=0}^T \left[\frac{x[n]}{(1 + n^2)} \right] \quad (8)$$

A métrica momento de 3ª do sinal evidencia ainda mais a diferença entre os valores da variável independente com relação a variável mensurada. A Equação 9 descreve o processo de cálculo (LATHI, 2006).

$$Momento3(x) = \sum_{n=0}^T [n^3 x[n]] \quad (9)$$

Razão do Ruído:

A relação sinal/ruído é definida como a subtração da informação mútua média, entre classes e atributos, da entropia média dos atributos e então o resultado é dividido novamente pela informação mútua média entre classes e atributos. A equação 10 formaliza esta relação matemática considerando um conjunto de dados genérico D (KALOUSIS; GAMA; HILARIO, 2004).

$$NS.ratio(D) = \frac{Entropy(D) - AMI(D)}{AMI(D)} \quad (10)$$

Em termos teóricos, a Equação 10 estima a quantidade de informações não úteis do conjunto de dados D .

Distorção Harmônica Total do Sinal:

A distorção harmônica total é determinada a partir da frequência fundamental e dos

primeiros cinco harmônicos usando um periodograma modificado do mesmo comprimento do sinal de entrada. É caracterizada como o somatório das razões entre o sinal de entrada e o sinal harmônico considerado equivalente (NASCIMENTO; BERGER; ROCHA, 2007).

A.3 Medida Baseada em Complexidade

As medidas de complexidade buscam representar particularidades locais e globais de conjuntos de dados (HO; BASU; LAW, 2006). Algumas dessas medidas também analisam o alcance e propagação de valores no conjunto de dados relativos a cada recurso e verificação para sobreposições entre diferentes classes (MOLLINEDA; SÁNCHEZ; SOTOCA, 2005; HO; BASU, 2002).

Dimensão Fractal do Conjunto de Dados:

O uso do conceito de dimensão fractal está associado à existência de redundância nos conjuntos de dados e a possibilidade desses conjuntos serem aproximados por dimensões menores (TRAINA; SOUSA; TRAINA, 2005).

Para fractais estatisticamente auto-semelhantes, como conjuntos de dados reais, uma maneira de definir a dimensão fractal é dada pelo Fractal de Correlação de Dimensão D_2 , que pode ser calculado pelo método Box-Count Plot (FALOUTSOS; KAMEL, 1994). Nesse método, a ideia consiste, inicialmente, na construção de uma rede no conjunto de dados à parte r células. Então, o número de pontos dentro do i^{th} célula de tamanho r , chamada de $C_{r,i}$, são contadas. O Fractal de Correlação da Dimensão D_2 é definido pela Equação 11.

$$D_2 = \frac{\partial \log(S_2(r))}{\partial \log(r)}, r \in [r_{\min}, r_{\max}], \quad (11)$$

na qual,

$$S_2(r) = \sum_i C_{r,i}^2. \quad (12)$$

Em teoria, os fractais exatamente auto-semelhantes são infinitos. Na prática, conjuntos de dados reais, que têm um número finito de pontos, são considerados fractais estatisticamente auto-semelhante a um intervalo de escala $r \in [r_{\min}, r_{\max}]$ se obedecem a uma regra de construção bem definida nesta faixa. Assim, a dimensão intrínseca de um determinado conjunto de dados pode ser medida como a inclinação da linha reta que melhor se ajusta às porções lineares do gráfico em uma escala logarítmica $S_2(r)$ por r (TRAINA; SOUSA; TRAINA, 2005).

A.4 Medidas Baseadas na Análise de Sistemas

As medidas baseadas na teoria na análise de sistemas dinâmicos buscam caracterizar os dados em termos de características do sistema físico. Essas medidas são usadas como base para a análise de sistemas dinâmicos.

Valor RMS:

O valor RMS (*root mean square*) de um sinal depende de sua forma de onda. Para o valor RMS leva-se em conta a equivalência com um sinal contínuo, é uma medida estatística da magnitude de uma quantidade variável (GEROMEL; KOROGUI, 2019).

Traço e Autovalores da Matriz Simétrica do Sinal

Formalmente, uma matriz é um conjunto ordenado de elementos dispostos em linhas e colunas. Usando K para designar o conjunto dos números reais ($\mathbb{R}$) em união com o conjunto dos números complexos ($\mathbb{C}$) e m e n como números inteiros positivos.

Definição

Chama-se matriz do tipo $m \times n$ sobre K a todo o quadro que se obtém dispondo $m \times n$ números segundo m linhas e n colunas, como na Equação 13 a matriz A . Nesse caso, usa-se como notação os termos a_{ij} , com $i = 1...m$ e $j = 1...n$, como os elementos da matriz A , em que i indica a linha em que se situa o elemento e j indica a coluna onde o elemento está situado.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (13)$$

Enquadramento de matrizes

Para conjuntos bidimensionais de dados, ou matrizes, que não são quadradas é restringido a aplicação de algumas funções matemáticas, como a extração de determinantes e do polinômio característico. Como os conjuntos de dados, que são representados por matrizes, podem ter dimensões distintas, é necessário aplicar sempre uma técnica para enquadrar a matriz, ou seja, torná-la quadrada.

Para isso, com uma matriz retangular A_{mn} , com $m \neq n$, representativa de um conjunto de dados, faz-se a multiplicação matricial pela sua forma transposta, como na Equação 14. Assim, o resultado dessa multiplicação é uma matriz quadrada (SM_{nn}) e simétrica, possibilitando o cálculo de diversas funções matemáticas que

não são possíveis com matrizes retangulares, sendo ainda representativas do conjunto de dados.

$$SM_{nn} = A_{nm}^T \times A_{mn} \quad (14)$$

Determinante da matriz de dados

O determinante de uma matriz é um número que depende apenas dos elementos da matriz. Desse modo, a toda matriz quadrada de ordem n (M_n), é possível associar um número do conjunto $\mathbb{R}$ segundo uma determinada lei. Os determinantes são utilizados, para caracterizar certas operações algébricas, ou mesmo para sintetizar certas expressões matemáticas complexas.

Menor complementar

Considerando uma matriz A , de ordem n , como na Equação 15.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (15)$$

O menor complementar D_{ij} , relativo ao elemento a_{ij} , é o determinante da matriz quadrada de ordem $n - 1$, que se obtém de A retirando-se a linha i e a coluna j , como exemplo na Equação 16.

$$D_{12} = \begin{bmatrix} -- & -- & \dots & -- \\ a_{21} & -- & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & -- & \dots & a_{nn} \end{bmatrix} \quad (16)$$

Cofator

Dada uma matriz A , de ordem n , o cofator do elemento a_{ij} é o número de A_{ij} que se obtém multiplicando-se o parâmetro $(-1)^{i+j}$ pelo menor complementar de a_{ij} , segundo a Equação 17.

$$A_{ij} = (-1)^{i+j} \times D_{ij} \quad (17)$$

Determinante de matrizes quadradas de ordem n

Para uma matriz quadrada de ordem n como a mostrada na Equação 15, o determinante de A é definido pela Equação 18.

$$\det A = a_{11} \times A_{11} + a_{12} \times A_{12} + \dots + a_{1n} \times A_{1n} \quad (18)$$

Essa equação é a soma dos números que se obtém da multiplicação de cada termo da primeira linha pelo seu respectivo cofator. Ainda, observa-se que cada cofator envolve o determinante de matrizes de ordem $n - 1$. Para se calcular o determinante de A deve-se aplicar a Equação 18 para cada submatriz encontrada até que a ordem da matriz seja 1. Desta forma, sabe-se que o determinante de matrizes de ordem 1 é o próprio e único elemento dessa matriz.

Assim, há duas formas de se calcular o determinante de matrizes de ordem n . A primeira é como mostrado na Equação 18, onde se fixou a linha i . A segunda forma é alcançada fixando-se a coluna j , como na Equação 19

$$\det A = a_{1j} \times A_{1j} + a_{2j} \times A_{2j} + \dots + a_{nj} \times A_{nj} \quad (19)$$

Traço de matrizes

O traço, em conjunto com o determinante, são números representativos de matrizes. Esse valor é determinado pela soma dos elementos da diagonal principal de matrizes quadradas de ordem n . O traço de uma matriz A de ordem n é denotado pela Equação 20 (GEROMEL; KOROGUI, 2019).

$$Tr(A) = \sum_{i,j=1}^n a_{ij} \quad (20)$$

Autovalor de uma matriz

Considerando uma matriz quadrada A que pertence ao conjunto dos $\mathbb{R}$ de ordem n . Se λ for um número complexo e v um vetor complexo não nulo de dimensão n que satisfaça a Equação 21, então λ é um autovalor de A e v é um autovetor de A associado ao autovalor λ (GEROMEL; KOROGUI, 2019).

$$A \times v = \lambda \times v \quad (21)$$

Da definição da Equação 21, tem-se a representação da Equação 22, e para que a solução seja não nula, o determinante deve ser igual a zero, sendo I a matriz identidade.

$$(\lambda \times I - A) \times v = 0 \quad (22)$$

Desta maneira, aplicando o determinante para a parte nula da Equação 22 é determinado o polinômio característico da matriz, como na Equação 23.

$$\det(\lambda \times I - A) = 0 \Rightarrow \alpha_0 + \alpha_1 \times \lambda + \dots + \alpha_n \times \lambda^n \quad (23)$$

Nessa equação, os termos α de $0 \dots n$ são parâmetros alcançados com a aplicação do determinante e λ os autovalores. Cada raiz do polinômio característico é caracterizada por um autovalor, e para uma matriz de ordem n são determinados n autovalores.

APÊNDICE B

Resultados Complementares da Análise de Robustez

Nesse apêndice são apresentados os resultados complementares da análise de robustez, discutidos na Seção 5.5.

Figura 37 – Resultados para a análise de robustez segundo o parâmetro taxa de falsos positivos para o caso da classificação binária.

Figura 38 – Resultados para a análise de robustez segundo o parâmetro taxa de falsos negativos para o caso da classificação binária.

Figura 39 – Resultados para a análise de robustez segundo o parâmetro taxa de erros para o caso da classificação binária.

Figura 40 – Resultados para a análise de robustez segundo o parâmetro área complementar a curva ROC para o caso da classificação binária.

Figura 41 – Resultados para a análise de robustez segundo o parâmetro taxa de falsos positivos para o caso da classificação multiclasse.

Figura 42 – Resultados para a análise de robustez segundo o parâmetro taxa de falsos negativos para o caso da classificação multiclasse.

Figura 43 – Resultados para a análise de robustez segundo o parâmetro taxa de erros para o caso da classificação multiclasse.

Figura 44 – Resultados para a análise de robustez segundo o parâmetro área complementar a curva ROC para o caso da classificação multiclasse.