

**UNIVERSIDADE NOVE DE JULHO - UNINOVE
PROGRAMA DE PÓS GRADUAÇÃO EM INFORMÁTICA E GESTÃO
DO CONHECIMENTO - PPGI**

RODRIGO CARDOZO DE JESUS

**UM ESTUDO COMPARATIVO ENTRE TÉCNICAS DE
APRENDIZAGEM DE MÁQUINA E PROCESSAMENTO DE SINAIS
PARA DETECTAR BARRAS QUEBRADAS EM UM MOTOR DE
INDUÇÃO TRIFÁSICO COM ROTOR DO TIPO GAIOLA DE ESQUILO**

**São Paulo
2019**

RODRIGO CARDOZO DE JESUS

**UM ESTUDO COMPARATIVO ENTRE TÉCNICAS DE
APRENDIZAGEM DE MÁQUINA E PROCESSAMENTO DE SINAIS
PARA DETECTAR BARRAS QUEBRADAS EM UM MOTOR DE
INDUÇÃO TRIFÁSICO COM ROTOR DO TIPO GAIOLA DE ESQUILO**

Exame de Defesa apresentado a Universidade
Nove de Julho - UNINOVE, como parte dos
requisitos para a obtenção do título de Mes-
tre em Informática e Gestão do Conheci-
mento.

Prof. Orientador: Dr. Cleber Gustavo Dias

**São Paulo
2019**

Jesus, Rodrigo Cardozo de.

Um estudo comparativo entre técnicas de aprendizagem de máquina e processamento de sinais para detectar barras quebradas em um motor de indução trifásico com rotor do tipo gaiola de esquilo. / Rodrigo Cardozo de Jesus. 2019.

156 f.

Dissertação (Mestrado) - Universidade Nove de Julho - UNINOVE, São Paulo, 2019.

Orientador (a): Prof. Dr. Cleber Gustavo Dias.

1. Motor de indução. 2. Barras quebradas. 3. Transformada rápida de Fourier. 4. Método dos vizinhos mais próximos.

I. Dias, Cleber Gustavo.

II. Título.

CDU 004

São Paulo, 31 de janeiro de 2019.

Aos meus pais Wilson Carvalho de Jesus e Marli da Penha Cardozo por sempre me apoiarem e me darem forças para continuar seguindo em frente.

Ao meu irmão Thiago Cardozo de Jesus e a minha cunhada Priscila Araújo pelo apoio constante e pelos incentivos.

*Em memória a minha Avó Alzira, a meu Avô
João e a minha Tia Leoseni, pessoas mara-
vilhosas*

AGRADECIMENTOS

Agradeço, primeiramente, ao meu orientador Prof. Dr. Cleber Gustavo Dias, pela confiança depositada em mim e pelas orientações sempre precisas que me ajudaram muito ao longo da realização dessa dissertação. Meu muito obrigado!

Ao Prof. Dr. Luís Carlos Silva, colega de pesquisa, pela ajuda e dicas importantíssimas, pois sem elas, esse trabalho não seria possível.

Agradeço também ao Prof. Dr. Sidnei Alves de Araújo pela ajuda e pelos importantes apontamentos que foram de grande ajuda na elaboração desse texto.

Obrigado ao Prof. MSc. Jan Novaes Recicar pela concessão do espaço para a elaboração e montagem do laboratório e dos equipamentos de pesquisa.

Agradeço ao meu colega de trabalho Sílvio de Souza Jardim, pela assistência e excelentes ideias na montagem e idealização do laboratório de pesquisa, onde está o motor estudado.

Agradeço minha amiga Patricia de Paula Thomaz sempre torcendo por mim, me ajudando com os textos da dissertação e com as nossas conversas sempre divertidas e inteligentes que ajudaram e muito a realização dessa dissertação.

Agradeço ao meu colega de pesquisa e amigo André Clemente Lira pela parceria e aos meus amigos de programa Aurélio Briani Matias Ramos, Celia Hatsumi Aihara, Dimitria Theophanis Boukouvalas, Fernanda Moran Menezes Pereira, Jonathan Farias Bordignon, Késsia Lucas Rodrigues, Luis Carlos dos Santos Junior pela força e pela amizade nesses últimos anos que foram importantes para a realização desse trabalho.

À Universidade Nove de Julho ao programa de pós-graduação em informática (*PPGI*) ao Prof. Dr. André Felipe Henriques Librantz e a todos os professores do programa pelo apoio e confiança, a *FAPESP* pelo apoio financeiro e a todos que colaboraram direta ou indiretamente para a execução desse trabalho.

"Você ganha força, coragem e confiança através de cada experiência em que você realmente para e encara o medo de frente."

Eleanor Roosevelt

- *R.C.Jesus. UM ESTUDO COMPARATIVO ENTRE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E PROCESSAMENTO DE SINAIS PARA DETECTAR BARRAS QUEBRADAS EM UM MOTOR DE INDUÇÃO TRIFÁSICO COM ROTOR DO TIPO GAIOLA DE ESQUILO. 2019. Dissertação (Mestrado) - Uninove, Universidade Nove de Julho - Programa de pós-graduação em informática (PPGI), São Paulo, SP, Brasil.*

A detecção de falhas em motores de indução trifásico do tipo gaiola de esquilo revelou um interesse crescente nos últimos anos, especialmente pelo estudo e a implementação de novas técnicas de diagnóstico, dada a sua utilização no segmento industrial em todo o mundo. Dentre as falhas existentes, destaca-se o rompimento de uma ou mais barras que fazem parte da gaiola rotórica da máquina. Uma das técnicas mais estudadas na literatura, e empregadas ainda hoje, para a detecção de barras quebradas, é a análise da assinatura da corrente elétrica de uma das fases do motor, a fim de encontrar determinadas harmônicas que indicam a presença do defeito. Esta técnica utiliza a Transformada Rápida de Fourier para a análise das componentes harmônicas, e outros estudos mais recentes empregaram, por exemplo, a Transformada de Hilbert, de modo a melhorar a resolução em frequência do sinal, bem como o uso de medidas estatísticas para encontrar alguns parâmetros do sinal de corrente do motor, na condição de rotor defeituoso. Adicionalmente, outras pesquisas têm investigado o uso das técnicas de aprendizagem de máquina para auxiliar na avaliação das condições do motor, a partir das características extraídas nos domínios do tempo e da frequência. Deste modo, o presente trabalho desenvolveu um estudo comparativo entre algumas técnicas de processamento de sinais, usadas no diagnóstico de barras rotóricas quebradas, e os algoritmos de aprendizagem de máquina mais empregados no monitoramento e diagnóstico de falhas em máquinas elétricas. Tais algoritmos foram parametrizados em diversas condições sendo possível comparar não apenas a acurácia de cada modelo, mas também as taxas de falso positivo e falso negativo em cada caso. Os resultados obtidos mostraram que o uso das características estatísticas em conjunto com aquelas extraídas em frequência apresentou a melhor performance. Os experimentos realizados com um motor de 7.5 kW, para diversas condições de carga do motor, e especialmente em baixa carga, permitiriam avaliar assim a melhor combinação das aludidas técnicas para a detecção e a classificação do defeito na gaiola rotórica.

Palavras-chave: Motor de Indução, Barras Quebradas, Transformada Rápida de Fourier, Transformada de Hilbert, Máquina de Suporte Vetorial, Método dos Vizinhos mais Próximos, Regressão Logística, Aprendizagem de Máquina, Processamento de Sinais.

- R.C.Jesus. **A COMPARATIVE STUDY BETWEEN MACHINE LEARNING AND SIGNAL PROCESSING TECHNIQUES TO DETECT BROKEN BARS IN A THREE-PHASE SQUIRREL CAGE INDUCTION MOTOR.** 2019. Thesis (Master's degree) - *Uninove, Universidade Nove de Julho - Programa de pós-graduação em informática (PPGI), São Paulo, SP, Brasil.*

Detection of failures in three-phase squirrel-cage induction motors has shown increasing interest in recent years, especially in the study and implementation of new diagnostic techniques, given their use in the industrial segment worldwide. Among the existing faults, stands out the rupture of one or more bars that are part of the machine's rotor cage. One of the most studied techniques in the literature, and still used today for the detection of broken rotor bars, is the electric current signature analysis of one of the motor phases, in order to find certain harmonics that indicate the presence of the defect. This technique uses the Fast Fourier Transform for the analysis of harmonic components, and other more recent studies have employed, for example, the Hilbert Transform, in order to improve the frequency signal resolution, as well as the use of statistical measures to find some parameters of the motor current signal, in the condition of a defective rotor. In addition, other researches have investigated the use of machine learning techniques to aid in the evaluation of machine conditions, based on the characteristics extracted in the time and frequency domains. Thus, the present work developed a comparative study between some signal processing techniques, used in the diagnosis of broken rotor bars, and the machine learning algorithms most used in the monitoring and diagnosis of failures in induction motors. These algorithms were parameterized in several conditions and it was possible to compare not only the accuracy of each model, but also the false positive and false negative rates in each case. The results showed that the use of the statistical characteristics with those extracted in frequency domains presented the best performance. Experiments with a 7.5 kW motor, for several load conditions, and especially at low load, would allow to evaluate the best combination of the aforementioned techniques for the detection and classification of the defect in the rotor cage .

Keywords: Induction Motor, Broken Bars, Fast Fourier Transform, Hilbert Transform, Support Vector Machine, K-Nearest Neighbor, Logistic Regression, Machine Learning, Signal processing.

Lista de Figuras	15
Lista de Abreviaturas	19
Lista de Símbolos	20
Lista de Símbolos	21
1 Introdução	22
1.1 A Relevância dos Motores de Indução Trifásicos no Cenário Industrial . . .	22
1.2 Problema e Questão de Pesquisa	24
1.3 Objetivo	24
1.3.1 Objetivo Geral	24
1.3.2 Objetivos Específicos	25
1.4 Justificativa e Motivação	25
1.4.1 Importância em resolver o problema de barras rompidas no <i>MIT</i> . .	25
1.4.2 Justificativa para o uso das técnicas empregadas na detecção de barras rompidas em <i>MIT</i>	28
1.5 Limitação da Pesquisa	29
1.6 Organização do Trabalho	29
2 Fundamentação Teórica	31
2.1 Os Motores de Indução Trifásicos com Rotor Gaiola de Esquilo	31
2.2 A Ocorrência de falhas nos Motores de Indução Trifásicos	32
2.2.1 Princípios de Funcionamento do Motor de Indução Trifásico	35
2.3 O Escorregamento do Motor	38
2.4 A Análise de <i>Fourier</i>	39
2.4.1 A Transformada Rápida de Fourier (<i>FFT</i>)	41
2.5 A Transformada de Hilbert (HT)	42
2.6 Medidas Estatísticas Extraídas da Corrente do Motor	44
2.7 Aprendizagem de Máquina	45
2.7.0.1 Aprendizagem Supervisionada	46
2.7.0.2 A Regressão Logística	47
2.7.0.3 As Máquinas de Vetores de Suporte	49
2.7.0.4 Método do Vizinho Mais Próximo	52
2.8 Medidas para a avaliação de desempenho dos classificadores	54
2.8.1 A Curva <i>ROC</i>	55

3	Revisão da Literatura	57
3.1	As Técnicas de Processamento de Sinais Usadas no Diagnóstico de Defeitos em MIT	58
3.1.1	O Uso da Técnica <i>MCSA</i> na Detecção de Barras Quebradas	59
3.2	Técnicas de Aprendizagem de Máquina para a detecção de barras rompidas no <i>MIT</i>	61
4	Metodologia Proposta	64
4.1	Caracterização Metodológica	64
4.2	Procedimentos para a Condução dos Experimentos	64
4.3	Montagem do Aparato Experimental	67
4.4	Definição dos Cenários Operacionais Para o Motor	68
5	Resultados	71
5.1	Limitações das abordagens utilizadas no pré-processamento do sinal	71
5.1.1	Exemplo do espectro de frequências extraído com a <i>MCSA</i> para diferentes janelas amostrais	71
5.1.2	A Transformada de Hilbert no Apoio a Detecção de Barras Quebradas	74
5.1.3	Exemplo do espectro de frequência extraído com as técnicas <i>HT</i> e <i>FFT</i> para diferentes janelas amostrais	74
5.2	Avaliação dos melhores parâmetros para os classificadores	76
5.3	Avaliação das melhores abordagens para cada janela amostral	77
5.3.1	Melhores abordagens obtidas para uma janela de 10s	78
5.3.2	Melhores abordagens obtidas para uma janela de 5s	79
5.3.3	Melhores abordagens obtidas para uma janela de 3s	81
5.3.4	Melhores abordagens obtidas para uma janela de 1s	83
5.3.5	Melhores abordagens obtidas para uma janela de 0,5s	85
5.3.6	Melhores abordagens obtidas para uma janela de 0,25s	87
6	Conclusão	91
	Referências Bibliográficas	93
A	Apêndice	101
A.1	Avaliação do classificador <i>KNN</i> usando a <i>FFT</i> em uma janela de 10s . . .	101
A.1.1	Avaliação do classificador <i>SVM</i> usando a <i>FFT</i> em uma janela de 10s	102
A.1.2	Avaliação do classificador <i>RL</i> usando a <i>FFT</i> em uma janela de 10s	104
A.1.3	Avaliação do classificador <i>KNN</i> usando <i>FFT</i> e características <i>Estadísticas</i> em uma janela de 10s	105
A.1.4	Avaliação do classificador <i>SVM</i> com características extraídas usando a <i>FFT</i> e Estatísticas em uma janela de 10s	106

A.1.5	Avaliação do classificador <i>RL</i> com características extraídas usando a <i>FFT</i> e <i>Estatísticas</i> em uma janela de 10s	108
A.1.6	Avaliação do classificador <i>KNN</i> com características extraídas usando a <i>HT</i> e a <i>FFT</i> em uma janela de 10s	109
A.1.7	Avaliação do classificador <i>SVM</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 10s	110
A.1.8	Avaliação do classificador <i>RL</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 10s	112
A.1.9	Avaliação do classificador <i>KNN</i> usando as grandezas <i>Estatísticas</i> e combinadas com aquelas extraídas por meio da técnica <i>HT</i> e <i>FFT</i> em uma janela de 10s	113
A.1.10	Avaliação do classificador <i>SVM</i> usando as grandezas <i>Estatísticas</i> e aquelas extraídas com as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 10s	114
A.1.11	Avaliação do classificador <i>RL</i> usando as grandezas <i>Estatísticas</i> e aquelas extraídas com as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 10s	116
A.1.12	Avaliação do classificador <i>KNN</i> usando apenas as grandezas <i>Estatísticas</i> como entradas em uma janela de 10s	117
A.1.13	Avaliação do classificador <i>SVM</i> usando apenas as grandezas <i>Estatísticas</i> como entradas em uma janela de 10s	119
A.1.14	Avaliação do classificador <i>RL</i> usando apenas as grandezas <i>Estatísticas</i> como entradas em uma janela de 10s	120
A.1.15	Avaliação do classificador <i>KNN</i> usando a <i>FFT</i> em uma janela de 5s	122
A.1.16	Avaliação do classificador <i>SVM</i> usando a <i>FFT</i> em uma janela de 5s	123
A.1.17	Avaliação do classificador <i>RL</i> usando a <i>FFT</i> em uma janela de 5s	124
A.1.18	Avaliação do classificador <i>KNN</i> com características extraídas usando a <i>FFT</i> e <i>Estatísticas</i> em uma janela de 5s	126
A.1.19	Avaliação do classificador <i>SVM</i> com características extraídas usando a <i>FFT</i> e <i>Estatísticas</i> em uma janela de 5s	127
A.1.20	Avaliação do classificador <i>RL</i> com características extraídas usando a <i>FFT</i> e <i>Estatísticas</i> em uma janela de 5s	128
A.1.21	Avaliação do classificador <i>KNN</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 5s	130
A.1.22	Avaliação do classificador <i>SVM</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 5s	131
A.1.23	Avaliação do classificador <i>RL</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 5s	132
A.1.24	Avaliação do classificador <i>KNN</i> usando como entradas as características extraídas com as grandezas <i>Estatísticas</i> e as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 5s	134

A.1.25	Avaliação do classificador <i>SVM</i> usando como entradas as características extraídas com as grandezas <i>Estatísticas</i> e as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 5s	135
A.1.26	Avaliação do classificador <i>RL</i> com características extraídas usando <i>HT</i> e <i>FFT</i> e <i>Estatísticas</i> em uma janela de 5s	136
A.1.27	Avaliação do classificador <i>KNN</i> com características <i>Estatísticas</i> em uma janela de 5s	138
A.1.28	Avaliação do classificador <i>SVM</i> usando como entradas apenas as características <i>Estatísticas</i> em uma janela de 5s	139
A.1.29	Avaliação do classificador <i>RL</i> usando como entradas apenas as características <i>Estatísticas</i> em uma janela de 5s	140
A.1.30	Avaliação do classificador <i>KNN</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 3s	142
A.1.31	Avaliação do classificador <i>SVM</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 3s	143
A.1.32	Avaliação do classificador <i>RL</i> com características extraídas usando <i>HT</i> e <i>FFT</i> em uma janela de 3s	144
A.1.33	Avaliação do classificador <i>KNN</i> usando como entradas as características <i>Estatísticas</i> e aquelas extraídas com as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 3s	146
A.1.34	Avaliação do classificador <i>SVM</i> usando como entradas as características <i>Estatísticas</i> e aquelas extraídas com as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 3s	147
A.1.35	Avaliação do classificador <i>RL</i> usando como entradas as características <i>Estatísticas</i> e aquelas extraídas com as técnicas <i>HT</i> e <i>FFT</i> em uma janela de 3s	148
A.1.36	Avaliação do classificador <i>KNN</i> usando como entradas as características <i>Estatísticas</i> em uma janela de 3s	150
A.1.37	Avaliação do classificador <i>SVM</i> usando como entradas as características <i>Estatísticas</i> em uma janela de 3s	151
A.1.38	Avaliação do classificador <i>RL</i> com características <i>Estatísticas</i> em uma janela de 3s	152

B Apêndice **154**

B.1	Resumo das Configurações Empregadas em Todos os Modelos	154
-----	---	-----

LISTA DE FIGURAS

1.1	O Universo Tecnológico de Motores Elétricos (Adaptado de "WEG Guia de Especificação de Motores Elétricos", 2016)	23
1.2	Exemplo de barra rompida em um <i>MIT</i> : (a) Rotor de um <i>MIT</i> com uma barra rompida, (b) Anel de curto de um <i>MIT</i> rachado (FISER et al., 2011)	26
1.3	Barra rompida na junção com o anel de curto (DIAS, 2006)	26
2.1	Rotor do tipo gaiola de esquilo (CHAPMAN; STEPHEN, 2013)	31
2.2	Visão explodida de um motor de indução com rotor do tipo gaiola de esquilo (Morais de Sousa, 2017)	33
2.3	Frequências de falhas normalmente encontradas em um motor de indução trifásico (LIANG; EDOMWANDEKHOE, 2017)	34
2.4	Representação de um rotor do tipo gaiola de esquilo (CHAPMAN; STEPHEN, 2013)	34
2.5	A produção de conjugado induzido em um motor de indução (CHAPMAN; STEPHEN, 2013)	37
2.6	Funcionamento da <i>DFT</i> , adaptado de (Steven W. Smith, 1999)	40
2.7	Função <i>Sigmoide</i> (NG, 2012)	48
2.8	Funcionamento do gradiente ascendente, adaptado de (HARRINGTON, 2015)	49
2.9	Classificação de duas classes, adaptado de (KUBAT, 2015)	50
2.10	Exemplo de classificação usando o <i>SVM</i> , adaptado de (KUBAT, 2015)	51
2.11	Agrupamento de dados de três tipos de flores de acordo com o tamanho da pétala, adaptado de: https://www.mathworks.com/help/stats/classification-using-nearest-neighbors.html	52
2.12	Exemplo de classificação utilizando o <i>KNN</i> , adaptado de (KUBAT, 2015)	53
2.13	Curva <i>ROC</i> padrão para a avaliação dos modelos	56
4.1	Metodologia proposta no trabalho	64
4.2	Aparato experimental montado para a avaliação de falhas no rotor	68
4.3	Simulação das barras rompidas no rotor para o MIT obtido com o fomento da FAPESP	68
5.1	A) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 10s) e B) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 10s)	72
5.2	A) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 5s) e B) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 5s)	73

5.3	A) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 3s) e B) Banda lateral esquerda para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 3s)	73
5.4	A) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 10s) e B) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 10s)	75
5.5	A) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 5s) e B) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 5s)	75
5.6	A) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 3s) e B) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 3s)	76
5.7	A) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 0,66% (janela de 1s) e B) Banda lateral esquerda (<i>2sf</i>), usando as técnicas <i>HT</i> e <i>FFT</i> , para um <i>MIT</i> avariado e operando com um escorregamento igual a 1,33% (janela de 1s)	76
5.8	Curvas <i>ROC</i> das melhores abordagens para uma janela de 10s	79
5.9	Curva <i>ROC</i> do melhor classificador para uma janela de 10s	79
5.10	Curvas <i>ROC</i> das melhores abordagens para uma janela de 5s	80
5.11	Curva <i>ROC</i> da melhor abordagem para uma janela de 5s	81
5.12	Curvas <i>ROC</i> das melhores abordagens para uma janela de 3s	82
5.13	Curva <i>ROC</i> do melhor classificador para uma janela de 3s	83
5.14	Curvas <i>ROC</i> das melhores abordagens para uma janela de 1s	84
5.15	Curva <i>ROC</i> da melhor abordagem para uma janela de 1s	85
5.16	Curvas <i>ROC</i> das melhores abordagens para uma janela de 0,5s	86
5.17	Curva <i>ROC</i> da melhor solução para uma janela de 0,5s	87
5.18	Curvas <i>ROC</i> das melhores abordagens para uma janela de 0,25s	88
5.19	Curva <i>ROC</i> da melhor solução para uma janela de 0,25s	89
A.1	Curvas <i>ROC</i> do classificador <i>KNN</i> para diferentes valores de <i>K</i>	102
A.2	Curvas <i>ROC</i> do classificador <i>SVM</i> para diferentes parâmetros de <i>C</i> e γ	103
A.3	Curva <i>ROC</i> gerada para a <i>RL</i>	104
A.4	Curvas <i>ROC</i> geradas para o classificador <i>KNN</i> usando <i>FFT</i> e <i>Estatísticas</i> como entradas em uma janela de 10s	106
A.5	Curvas <i>ROC</i> do classificador <i>SVM</i> usando ambas as características <i>FFT</i> e <i>Estatísticas</i> em uma janela amostral de 10s	107

A.6	Curvas <i>ROC</i> do classificador <i>RL</i> utilizando <i>FFT</i> e <i>Estatísticas</i>	109
A.7	Curvas <i>ROC</i> para o classificador <i>KNN</i> usando as técnicas <i>HT</i> e a <i>FFT</i>	110
A.8	Curvas <i>ROC</i> do classificador <i>SVM</i>	111
A.9	Curvas <i>ROC</i> do classificador <i>RL</i> utilizando <i>HT</i> e <i>FFT</i>	113
A.10	Curvas <i>ROC</i> do classificador <i>KNN</i>	114
A.11	Curvas <i>ROC</i> para o classificador <i>SVM</i> usando diferentes configurações de pa- râmetros	115
A.12	Curva <i>ROC</i> obtida para o classificador <i>RL</i> , usando as entradas processadas com as grandezas <i>Estatísticas</i> e as técnicas <i>HT</i> e <i>FFT</i>	117
A.13	Curvas <i>ROC</i> geradas com o classificador <i>KNN</i> usando apenas entradas no domínio do tempo	118
A.14	Curvas <i>ROC</i> geradas para o classificador <i>SVM</i> usando apenas as grandezas <i>Estatísticas</i> como entradas	120
A.15	Curva <i>ROC</i> do classificador <i>RL</i> utilizando entradas no domínio do tempo	121
A.16	Curvas <i>ROC</i> geradas com o classificador <i>KNN</i> , em uma janela de 5s, usando apenas entradas extraídas com a <i>FFT</i>	122
A.17	Curvas <i>ROC</i> do classificador <i>SVM</i> utilizando as entradas extraídas com a <i>FFT</i> , em uma janela de 5s	124
A.18	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>FFT</i> em uma janela de 5s	125
A.19	Curvas <i>ROC</i> obtidas para o classificador <i>KNN</i> , em uma janela de 5s, usando as técnicas <i>FFT</i> e <i>Estatísticas</i>	126
A.20	Curvas <i>ROC</i> do classificador <i>SVM</i> em uma janela de 5s, usando entradas extraídas com <i>FFT</i> e <i>Estatísticas</i>	128
A.21	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>FFT</i> e <i>Estatísticas</i>	129
A.22	Curvas <i>ROC</i> do classificador <i>KNN</i>	130
A.23	Curvas <i>ROC</i> do classificador <i>SVM</i> , em uma janela de 5s, usando as técnicas <i>HT</i> e <i>FFT</i>	132
A.24	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>HT</i> e <i>FFT</i>	133
A.25	Curvas <i>ROC</i> geradas para o classificador <i>KNN</i> em uma janela de 5s, usando características de entrada nos domínios do tempo e da frequência	134
A.26	Curvas <i>ROC</i> geradas do classificador <i>SVM</i> em uma janela de 5s, usando en- tradas nos domínios do tempo e da frequência	136
A.27	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>HT</i> , <i>FFT</i> e <i>Estatísticas</i>	137
A.28	Curvas <i>ROC</i> do classificador <i>KNN</i> , em uma janela de 5s, usando apenas en- tradas geradas no domínio do tempo	138
A.29	Curvas <i>ROC</i> do classificador <i>SVM</i> em uma janela de 5s, usando apenas entra- das no domínio do tempo	140
A.30	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>Estatísticas</i>	141
A.31	Curvas <i>ROC</i> geradas com o classificador <i>KNN</i> em uma janela de 3s	142

A.32	Curvas <i>ROC</i> do classificador <i>SVM</i>	144
A.33	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>HT</i> e <i>FFT</i>	145
A.34	Curvas <i>ROC</i> do classificador <i>KNN</i> na janela de 3s e entradas nos domínios do tempo e da frequência	146
A.35	Curvas <i>ROC</i> do classificador <i>SVM</i>	148
A.36	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>HT</i> <i>FFT</i> e <i>Estatísticas</i>	149
A.37	Curvas <i>ROC</i> do classificador <i>KNN</i> em uma janela de 3s, usando apenas entra- das no domínio do tempo	150
A.38	Curvas <i>ROC</i> do classificador <i>SVM</i>	152
A.39	Curva <i>ROC</i> do classificador <i>RL</i> utilizando <i>Estatísticas</i>	153

LISTA DE ABREVIATURAS

<i>MCSA</i>	<i>Motor current signature analyse</i>
<i>SVM</i>	<i>Support Vector Machine</i>
<i>KNN</i>	<i>K - nearest neighbor</i>
<i>RL</i>	<i>Regressão logística</i>
<i>MIT</i>	<i>Motor de indução Trifásico</i>
<i>FFT</i>	<i>Fast fourier transform</i>
<i>HT</i>	<i>Hilbert transformation</i>
<i>IA</i>	<i>Inteligência Artificial</i>
<i>TC</i>	<i>Transformador de Corrente</i>
<i>MUSIC</i>	<i>Multiple Signal Classifier</i>
<i>ESPRIT</i>	<i>Estimation of Signal Parameters via Rotational Invariance Technique</i>
<i>EMD</i>	<i>Empirical Mode Decomposition</i>
<i>DFT</i>	<i>Discret Fourier Trensform</i>
<i>CA</i>	<i>Corrente Alternada</i>
<i>CC</i>	<i>Corrente Contínua</i>
<i>ML</i>	<i>Machine Learning</i>
<i>NN</i>	<i>Neural Network</i>
<i>RPM</i>	<i>Rotações por minuto</i>
<i>EST</i>	<i>Estatística</i>
<i>ASC</i>	<i>Área Sobre a Curva</i>
<i>ROC</i>	<i>Receiver Operating Characteristic</i>
<i>IC</i>	<i>Inteligência Computacional</i>
<i>freq</i>	<i>Frequência</i>
<i>DE</i>	<i>Distância Euclidiana</i>
<i>DCUB</i>	<i>Distância Cúbica</i>
<i>DCOSS</i>	<i>Distância Cosseno</i>

LISTA DE SÍMBOLOS

V	<i>Tensão Elétrica</i>
η_{sinc}	<i>Velocidade de sincronismo do campo magnético</i>
f_s	<i>Frequência do sistema</i>
P	<i>Números de Polos</i>
e_{ind}	<i>Tensão induzida nas barras rotoricas</i>
v	<i>Velocidade da barra rotorica em relação ao campo magnético</i>
B	<i>Vetor densidade de fluxo magnético</i>
l	<i>Comprimento do condutor dentro do campo magnético</i>
ω	<i>Velocidade angular</i>
θ	<i>Angulo entre tensão induzida no rotor e corrente no rotor</i>
B_{lq}	<i>Soma dos campos resultantes em cada bobina</i>
B_s	<i>Campo magnético girante gerado pelas correntes da bobina do estator da máquina</i>
E_R	<i>Tensão induzida no rotor</i>
I_R	<i>Corrente fluindo nas barras do rotor da máquina</i>
k	<i>Fator de construção da máquina</i>
B_R	<i>Campo magnético do rotor</i>
η_m	<i>Velocidade mecânica do eixo do motor</i>
s	<i>Escorregamento</i>
VP	<i>Verdadeiro Positivo</i>
VN	<i>Verdadeiro Negativo</i>
FP	<i>Falso Positivo</i>
FN	<i>Falso Negativo</i>
S	<i>Sensibilidade</i>
E	<i>Especificidade</i>
Ef	<i>Eficiência</i>
TA	<i>Taxa de Acerto</i>
CCM	<i>Coefficiente de Correlação de Matthews</i>
ω_{sinc}	<i>Velocidade angular dos campos magnéticos</i>
ω_m	<i>Velocidade angular mecânica do eixo do motor</i>
$Cos_{freq}[in]$	<i>Função cosseno</i>
$Sen_{freq}[in]$	<i>Função seno</i>
$Re\bar{X}[Freq]$	<i>Amplitude da função cosseno</i>
$Im\bar{X}[Freq]$	<i>Amplitude da função seno</i>
C	<i>Parâmetro C</i>
γ	<i>Parâmetro gama</i>
f	<i>Frequência fundamental da rede elétrica</i>

LISTA DE SÍMBOLOS

e^{jz}	<i>Equação de Euler</i>
$e^{j\omega t}$	<i>Notação complexa</i>
$\hat{x}(t)$	<i>Transformada de Hilbert de um sinal real</i>
$\bar{x}(t)$	<i>Sinal analítico</i>
$Mod\bar{x}(t)$	<i>Módulo do sinal analítico</i>
σ	<i>Desvio Padrão</i>
m_1	<i>Média</i>
m_2	<i>Variância</i>
m_3	<i>Momento terceira ordem</i>
m_4	<i>Momento quarta ordem ordem</i>
m_6	<i>Momento sexta ordem</i>
c_1	<i>Cumulante primeira ordem</i>
c_2	<i>Cumulante segunda ordem</i>
c_3	<i>Cumulante terceira ordem</i>
c_4	<i>Cumulante quarta ordem</i>
Ass	<i>Assimetria</i>
$ \bar{x} $	<i>Média Absoluta</i>
x_p	<i>Valor de Pico Máximo</i>
x_r	<i>Valor SRV</i>
f_c	<i>Fator de Crista</i>
f_o	<i>Fator de Onda</i>
x_{rms}	<i>Valor RMS</i>
$h_\theta(x)$	<i>Função hipótese</i>
$g(z)$	<i>Função sigmoide</i>
$\nabla f(x, y)$	<i>Gradiente ascendente</i>
in	<i>Indexador</i>
N	<i>Número de amostras</i>
X	<i>Sinal de entrada</i>
$Re X$	<i>Parte real do sinal</i>
$Im X$	<i>Parte imaginária do sinal</i>
θ	<i>Pesos da função sigmoide e hipótese</i>
X	<i>Sinal de entrada</i>
x	<i>Novo exemplo em um conjunto de treinamento</i>
y	<i>Dados rotulados do conjunto de treinamento</i>

1.1 A RELEVÂNCIA DOS MOTORES DE INDUÇÃO TRIFÁSICOS NO CENÁRIO INDUSTRIAL

Os motores de indução trifásicos (*MIT*) são máquinas elétricas rotativas muito relevantes no cenário industrial atual, não apenas no Brasil, mas em muitos segmentos produtivos ao redor do mundo. Atualmente, os *MIT* são as máquinas mais usadas no acionamento de cargas, devido ao seu baixo custo, robustez e confiabilidade (LIANG; EDMONDWANDERHOE, 2017).

Tais motores são empregados em diversas áreas como por exemplo: fábricas de cerâmicas, mineração, siderúrgica e fabricação de metais, entre outras. Segundo (MORAIS DE SOUSA, 2017) e (MICERINO, 2012), os *MIT* são fortemente utilizados no acionamento de bombas, ventiladores, misturadores, sopradores e em esteiras transportadoras, entre outras inúmeras aplicações.

Segundo (SAUER et al., 2015), existem atualmente cerca de 12 milhões de motores em operação no Brasil, o que evidencia a necessidade em manter essas máquinas funcionando com eficiência e sem paradas. Por esse motivo, é importante estudos e pesquisas voltados para o diagnóstico de eventuais falhas nos motores de indução, a fim de evitar perdas não apenas para a máquina, mas também para o processo produtivo a ela associado. Segundo (SILVA, 2014), a detecção de falhas pode ajudar as indústrias na manutenção preditiva, pois auxilia no diagnóstico prévio evitando assim uma parada não programada e ainda segundo (GODOY, 2016), a agilidade na manutenção diminui o tempo de inatividade das máquinas ligadas aos processos industriais.

Segundo (MORAIS DE SOUSA, 2017) e (GARCIA, 2003), a construção robusta dos *MIT*, com rotor do tipo gaiola de esquilo, evita faíscas entre as suas partes fixa (estator) e móvel (rotor), característica única que permite alcançar um rendimento da ordem de 90%, além do fato de reduzir o número de manutenções, quando comparado aos casos que utilizam os denominados motores de corrente contínua.

Logo, fica evidente a relevância que essas máquinas possuem para todo o sistema econômico e industrial de um país. Segundo (GODOY, 2016), houve um incremento no investimento da manutenção dessas máquinas nos últimos 10 anos, de 16,48% para 18,32%, especialmente no que diz respeito a manutenção preditiva.

Cabe mencionar que existem diferentes tipos de motores elétricos, como mostra a figura 1.1. Neste sentido, os motores podem ser divididos em dois grandes grupos, quais sejam:

1. Aqueles acionados por corrente alternada (Motor CA); e

2. Um outro grupo das máquinas alimentadas com corrente contínua (Motor CC).

É possível visualizar, na mesma figura, que os *MIT* podem ser ainda síncronos ou assíncronos. Os assíncronos, ou acionados por indução, são também divididos em dois tipos: com rotor bobinado e outro construído por barras e dois anéis terminais de curto-circuito, objeto do presente trabalho. Este segundo tipo de motor, embora bastante confiável, está sujeito a falhas nas diversas partes da sua estrutura, como será melhor descrito na próxima seção.

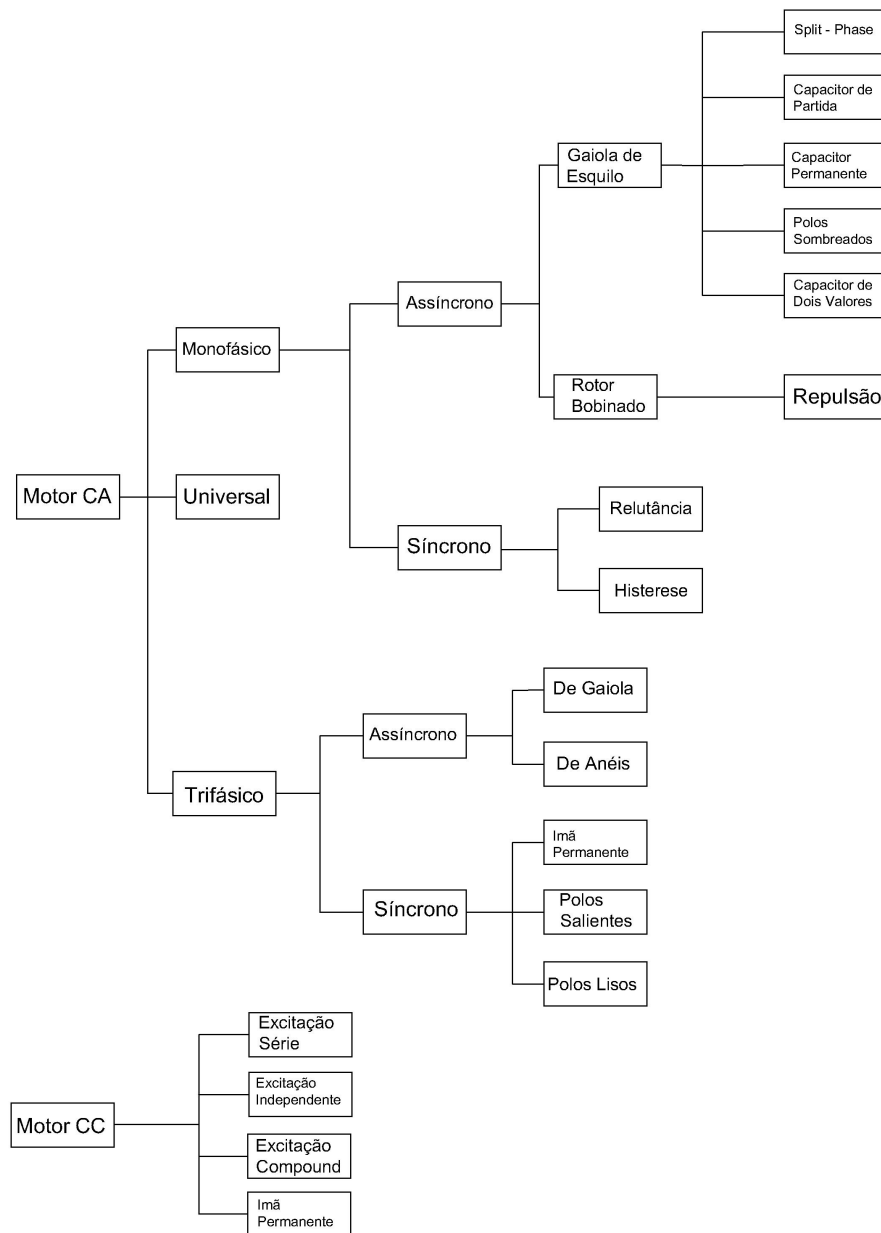


Figura 1.1: O Universo Tecnológico de Motores Elétricos (Adaptado de "WEG Guia de Especificação de Motores Elétricos", 2016)

1.2 PROBLEMA E QUESTÃO DE PESQUISA

O problema abordado neste trabalho envolve a detecção de barras quebradas na gaiola dos motores de indução trifásicos, especialmente na condição de baixo torque de carga e alimentado por um inversor de frequência. O fato do motor estar alimentado por um inversor de frequência torna essa avaliação ainda mais onerosa, em função de outras componentes harmônicas que surgem na corrente do motor.

Ademais, para um motor operando em baixa carga, ou baixo escorregamento, há uma limitação importante para o uso da tradicional técnica de análise da assinatura da corrente do estator, uma vez que as frequências harmônicas de interesse aparecem muito próximas à frequência fundamental da rede, o que inviabiliza, muitas vezes, o correto diagnóstico das reais condições do rotor.

Outro fator importante no estudo do presente problema de pesquisa, reside no fato de que diferentes abordagens de aprendizagem de máquina podem oferecer desempenhos distintos para determinadas características de entrada de um dado problema, logo, torna-se importante avaliar de forma comparativa não apenas as técnicas de processamento de sinais para extrair informações da corrente do motor, mas também o desempenho dos classificadores inteligentes capaz de indicar o melhor modelo para detectar o defeito na gaiola na menor janela de tempo amostral possível.

Vale mencionar que, a análise espectral da corrente, como descrito na literatura, necessita de uma janela amostral em torno de 100s, para uma frequência de amostragem igual a 10kHz, para a correta identificação das bandas de frequência responsáveis pela identificação da falha no rotor, logo, o seu uso implica em um armazenando da ordem de 1 milhão de amostras.

Muitas das soluções encontradas hoje no estado da arte empregam técnicas de processamento de sinais isoladas, ou mesmo com um único classificador, além do fato de processar os dados coletados em uma única janela amostral.

Dada a situação problema levantada, foi proposta a seguinte pergunta de pesquisa:

Como comparar a eficiência do diagnóstico de barras quebradas no rotor de um motor de indução, operando em baixo escorregamento e alimentado por um inversor, usando características da corrente do estator, nos domínios do tempo e da frequência, em conjunto com alguns classificadores inteligentes, a fim de encontrar a melhor abordagem?

1.3 OBJETIVO

1.3.1 OBJETIVO GERAL

O objetivo deste trabalho foi estudar e comparar técnicas de processamento digital de sinais para a extração de características do sinal de corrente de um motor de indução, e o

desempenho de diferentes classificadores inteligentes, a fim de avaliar a melhor abordagem para a detecção de barras rompidas na gaiola rotórica.

1.3.2 OBJETIVOS ESPECÍFICOS

Entre os objetivos específicos, pode-se destacar:

1. Extrair as características da corrente do motor no domínio da frequência, ou seja, as componentes harmônicas de interesse, usando de forma isolada e combinada, as técnicas *FFT* e *HT*;
2. Extrair as características estatísticas da corrente do motor, ou seja, as alterações observadas no domínio do tempo de forma isolada;
3. Combinar as características extraídas nos domínios do tempo e da frequência;
4. Implementar uma base de dados com todas as características extraídas em janelas de tempo amostral iguais a 10s, 5s, 3s, 1s, 0.5s e 0.25s;
5. Implementar os algoritmos baseados em classificadores e avaliar o desempenho obtido por eles, no caso a máquina de vetor de suporte, a técnica dos vizinhos mais próximos e a regressão logística;
6. Avaliar as abordagens implementadas para o motor operando em baixa carga, ou baixo escorregamento;
7. Avaliar as abordagens implementadas para o motor alimentado por um inversor de frequência; e
8. Comparar o desempenho obtido em cada caso (processamento e classificador) e que oferece uma maior acurácia para o diagnóstico da condição do rotor.

1.4 JUSTIFICATIVA E MOTIVAÇÃO

1.4.1 IMPORTÂNCIA EM RESOLVER O PROBLEMA DE BARRAS ROMPIDAS NO *MIT*

Segundo (CARDOSO, 1991), muito embora as falhas no rotor não sejam as mais frequentes, quando presentes, elas podem provocar danos e efeitos indesejáveis adicionais na máquina como a ocorrência de vibrações, um aumento na elevação de temperatura, a possibilidade de comprometer o seu isolamento e como consequência uma redução na sua vida útil. Quando não detectada no início, uma barra quebrada pode ainda ser projetada em direção ao estator e danificar todo o motor.

O rotor dos *MIT*, como apresentado anteriormente, é formado por barras transversais curto-circuitadas por dois anéis de curto (anéis terminais) em cada extremidade. Quanto ao rompimento de uma ou mais barras, ele ocorre geralmente na junção da barra com o anel terminal. A figura 1.2 mostra uma barra rompida (a) e um anel de curto trincado (b) e a figura 1.3 mostra um outro motor com uma barra rompida na junção com o anel.

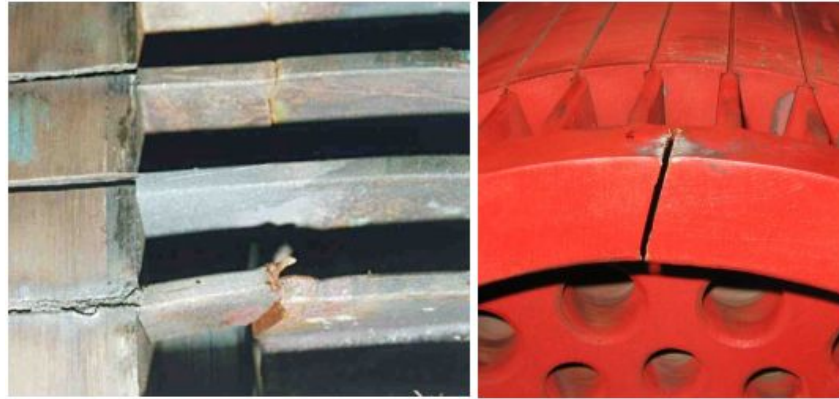


Figura 1.2: Exemplo de barra rompida em um MIT: (a) Rotor de um MIT com uma barra rompida, (b) Anel de curto de um MIT rachado (*FISER et al., 2011*)

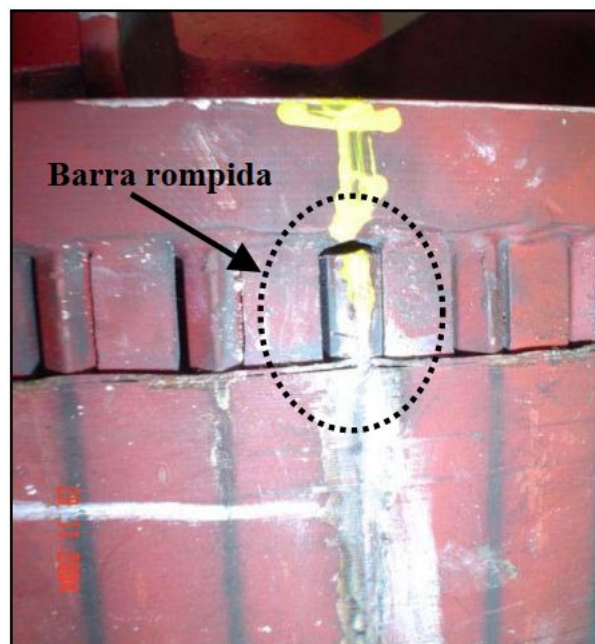


Figura 1.3: Barra rompida na junção com o anel de curto (*DIAS, 2006*)

Segundo (*CARDOSO, 1991*) e (*DIAS, 2006*) quando uma barra quebrada não é diagnosticada precocemente, ocorre uma propagação dessa avaria da seguinte maneira:

- Maior possibilidade de rompimento total, devido ao sobreaquecimento na região de falha;

- Potencial para iniciar uma erosão entre a barra e o núcleo rotórico, face as correntes difusas que passam a circular pelo seu núcleo; e
- As barras adjacentes as barras rompidas sofrem uma sobrecorrente bastante superior a sua corrente nominal, a fim de garantir o torque solicitado a máquina.

Ainda segundo (DIAS, 2006), e como já descrito, caso a fratura não seja detectada no seu início, a erosão causada pela barra quebrada poderá ser agravada e, pela ação de forças centrífugas, a barra poderá se soltar causando danos graves ao estator e a própria máquina. Segundo (KARMAKAR et al., 2016), as principais causas de barras rompidas em um motor de indução são as seguintes:

- Defeitos de fabricação;
- Esforços térmicos acima das suas especificações;
- Estresse mecânico causado por falhas nos rolamentos;
- Partidas consecutivas do motor na tensão nominal; e
- Fadiga da estrutura das barras do rotor.

Como revelado por (ZHANG et al., 2011), nas máquinas de grande porte, por exemplo, a incidência de barras quebradas é ainda maior. Desse modo é de extrema importância o estudo e o diagnóstico precoce deste tipo de falha no rotor dos *MIT*, especialmente devido aos elevados custos associados a manutenção e eventual parada da máquina.

O trabalho publicado por (THOMSON; FENGER, 2003) revela que, em uma dada aplicação industrial, o rompimento de barras em um motor de 2900 kW demandou reparos com custos da ordem de US\$ 160.000. No mesmo artigo, destaca-se outro caso de um motor de grande porte, acometido por barras rotóricas quebradas, em uma planta de gás, com prejuízos de aproximadamente US\$ 2 milhões de dólares.

Um artigo publicado no ano de 2007 (MCKINNON, 2007), revelou o caso de rompimento de barras no rotor de um motor de 3500 HP, com custos de manutenção da ordem de US\$ 90.000, incluindo US\$ 60.000 para o reparo do próprio motor e US\$ 30.000 relacionados a parada da produção. O autor revela ainda que caso o defeito não fosse detectado em uma fase inicial, tais custos poderiam alcançar valores próximos a US\$ 280.000. A origem do defeito foi associada a uma manutenção anterior precária realizada no rotor, especialmente na solda da barra na junção com o anel de curto-circuito.

Outros trabalhos publicados mais recentemente abordam a necessidade em se detectar o rompimento de barras no rotor de um motor de indução, especialmente nos casos em que a máquina opera em baixa carga, ou baixo escorregamento, (PUCHE-PANADERO et al., 2009), (LAALA; GUEDINI; ZOUZOU, 2011), (XU et al., 2013), (GYFTAKIS et

al., 2015), (DIAS; PEREIRA, 2018), e ainda a importância em diagnosticar o defeito quando o motor é alimentado por tensão não-senoidal, ou seja, acionado por um inversor de frequência (ROMERO-TRONCOSO et al., 2016) e (GODOY, 2016).

Face ao exposto, a motivação para o resolver o problema de barras rompidas reside no fato de que a detecção delas nos *MIT*, particularmente em baixo escorregamento e alimentado por um inversor, ainda encontra grande interesse na comunidade científica, frente a abrangência de uso dos motores nas aludidas condições operacionais.

1.4.2 JUSTIFICATIVA PARA O USO DAS TÉCNICAS EMPREGADAS NA DETECÇÃO DE BARRAS ROMPIDAS EM *MIT*

Neste trabalho foram escolhidas três abordagens de pré-processamento para a extração de características da corrente do estator do motor e três classificadores inteligentes para o diagnóstico final da condição do rotor.

Para a extração das características, foi escolhido a técnica *MCSA* (sigla em inglês para *Motor Current Signature Analise* ou análise da assinatura da corrente do estator), que usa a *FFT* (sigla em inglês para *Fast Fourier Transform* ou Transformada Rápida de Fourier) como ferramenta de aquisição de características, devido ao seu amplo uso na comunidade científica, além de ser um método de fácil implementação, pois basta um transformador de corrente acoplado a uma das fases do motor para coletar os dados.

A segunda técnica de extração de características utilizada foi a Transformada de Hilbert (*HT*) (sigla em inglês para *Hilbert Transform*). Esta segunda técnica foi escolhida pois permite melhorar a resolução em frequência do método *MCSA*.

Para a extração das características em uma condição do motor operando em baixa carga, ou baixo escorregamento, é necessário ter uma resolução de frequência de 0,01 Hz para uma janela amostral de 100s. Nesse trabalho, dada a limitação de hardware disponível e o interesse em avaliar o defeito em janelas amostrais inferiores a 100s, todos os experimentos foram realizados em uma janela amostral de 10s, para uma frequência de amostragem de 10kHz, logo com uma resolução em frequência de 0,1s. Esta resolução inferior a exigida foi possível apenas com o uso combinado das técnicas *HT* e *FFT*, pois como já mencionado, a *HT* permite melhorar a resolução em frequência do sinal no domínio da frequência, e tal condição é necessária para encontrar a frequência de interesse com o motor operando em carga baixa.

Uma terceira técnica consistiu em extrair algumas medidas estatísticas do sinal no domínio do tempo, devido ao seu baixo esforço computacional. Tais características estatísticas permitiram avaliar outros parâmetros da corrente de fase frente ao defeito no rotor.

Os três classificadores escolhidos foram a máquina de suporte vetorial *SVM* (sigla em inglês para *Support Vector Machine*), a técnica dos vizinhos mais próximos *KNN* (sigla

em inglês para *K - Nearest Neighbor*) e a regressão logística *RL*. Foram escolhidos o *KNN* e o *SVM* por conta dos trabalhos mais recentes que aplicaram essas duas técnicas na área de diagnóstico de máquinas elétricas (DIAS; PEREIRA, 2018). O classificador *RL* foi escolhido dada a sua facilidade de implementação.

1.5 LIMITAÇÃO DA PESQUISA

O presente estudo está direcionado para a detecção de barras quebradas em um motor de indução trifásico operando em regime permanente, logo, não foi objetivo deste trabalho investigar o referido defeito para a máquina na condição de partida. Adicionalmente, a pesquisa ora descrita procurou avaliar a ruptura completa da barra com o anel de curto, assim não foram estudados os efeitos de uma quebra parcial de barra na gaiola rotórica. A abordagem proposta procurou ainda comparar a eficiência do diagnóstico da falha para algumas das principais técnicas de processamento de sinais empregadas em métodos de diagnóstico não-invasivos e usando apenas o sinal de corrente do motor.

Finalmente, o trabalho investigou o uso de alguns classificadores inteligentes usados na literatura no campo das máquinas elétricas, para a detecção deste tipo de defeito, considerando ainda o motor alimentado por um inversor de frequência e operando em baixa carga. A aquisição da corrente do motor foi limitada em uma janela amostral menor ou igual a 10s, de modo a permitir que os modelos investigados possam ser oportunamente implementados em um hardware de baixo custo, além do fato de que trabalhos recentes procuram identificar o problema no rotor em tempos de processamento próximos a janela amostral proposta neste trabalho.

1.6 ORGANIZAÇÃO DO TRABALHO

Esta dissertação está organizada da seguinte forma:

- No capítulo 1 é apresentado uma introdução sobre o cenário no qual o motor de indução esta inserido, e também os objetivos, justificativa e problema de pesquisa desse trabalho;
- O capítulo 2 apresenta toda a fundamentação teórica para o diagnóstico de barras rompidas usada nesse trabalho;
- O capítulo 3 apresenta uma revisão da literatura sobre as principais técnicas usadas para o diagnostico de barras quebradas em motores de indução;
- O capítulo 4 apresenta a metodologia aplicada para o diagnóstico de barras rompidas nos *MIT*;
- O capítulo 5 mostra os resultados alcançados e;

- O capítulo 6 apresenta a conclusão desse trabalho.

FUNDAMENTAÇÃO TEÓRICA

2.1 OS MOTORES DE INDUÇÃO TRIFÁSICOS COM ROTOR GAIOLA DE ESQUILO

Os *MIT* com rotor gaiola de esquilo são máquinas rotativas alimentadas por corrente alternada. O uso de uma fonte de alimentação trifásica em um enrolamento trifásico do estator provoca a criação de um campo girante, e este induz uma força eletromotriz no enrolamento do rotor. Esta força induzida provoca a circulação de correntes nas barras da gaiola. Tal força é associada à onda de densidade de fluxo girante para produzir o torque.

Neste caso, o rotor apresenta uma velocidade de giro menor do que aquela produzida pelo campo magnético girante do estator da máquina. Portanto, haverá uma diferença entre as velocidades, ou a rotação, do campo magnético girante e o rotor da máquina, sendo essa diferença chamada de *escorregamento* (*s*). O rotor, portanto, "escorrega" em relação ao campo girante do estator, dando origem ao conjugado na máquina.

O rotor dos motores do tipo gaiola de esquilo são formados por barras condutoras encaixadas em ranhuras do núcleo rotórico e curto circuitadas em cada lado por anéis de curto, ou anéis terminais (Fitzgerald, A.E. and KINGSLEY, C. and Umans, 2006), como ilustrado na figura 2.1.

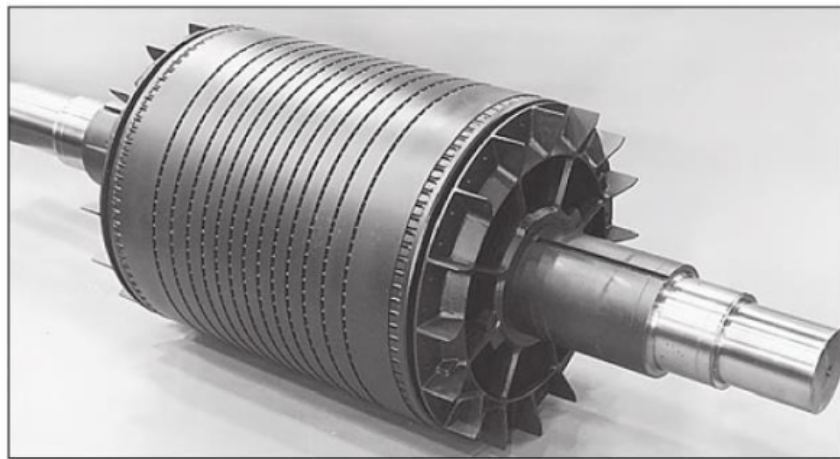


Figura 2.1: Rotor do tipo gaiola de esquilo (CHAPMAN; STEPHEN, 2013)

Segundo (ROÇAS, 2012), o *MIT* é uma máquina elétrica de corrente alternada, com o circuito elétrico de uma parte ligado a um sistema de alimentação, e a outra parte submetido à indução magnética criada pela circulação de correntes nos enrolamentos estatóricos. O autor afirma que o estator possui uma alimentação elétrica enquanto o rotor possui uma alimentação induzida pelo estator, logo, não há contato físico entre estator e rotor, mas sim um acoplamento magnético.

Segundo (CHAPMAN; STEPHEN, 2013), uma máquina de indução é diferente de

uma máquina de corrente contínua, uma vez que não há necessidade de uma corrente de campo para provocar o seu funcionamento, pois a tensão do estator é aquela que produz a corrente no rotor e o campo magnético do rotor é induzido nos seus enrolamentos, em vez de ser fornecida por uma conexão física de cabos.

Ainda segundo (CHAPMAN; STEPHEN, 2013), o *MIT* é usado preferencialmente como um motor, dado o fato de que este apresenta muitas desvantagens em relação ao seu uso como gerador.

2.2 A OCORRÊNCIA DE FALHAS NOS MOTORES DE INDUÇÃO TRIFÁSICOS

O estudo das possíveis falhas nos *MIT* vem se tornando um importante assunto nos últimos anos, devido ao seu uso bastante disseminado e também por serem capazes de acionar as mais diversas cargas em aplicações distintas, conforme já mencionado. Contudo, apesar da sua robustez, tais máquinas estão sujeitas aos mais diversos tipos de defeitos, e estes podem provocar prejuízos importantes aos processos produtivos, custos relacionados a parada total da própria máquina ou ainda a dispêndios de manutenção oriundos dessas avarias. De modo geral, o desempenho dessas máquinas pode ser afetado pelas seguintes classes de falhas (BELLINI et al., 2008); (FILIPPETTI; BELLINI; CAPOLINO, 2013):

- Falhas Elétricas;
- Falhas Mecânicas; e
- Falhas Externas.

Os *MIT* são compostos basicamente em três partes fundamentais, quais sejam: o estator, sendo esta a parte fixa da máquina, o rotor, formada pela parte móvel, ou eixo do motor, e os seus rolamentos. A figura 2.2 mostra as principais partes de um motor de indução como rotor gaiola de esquilo.

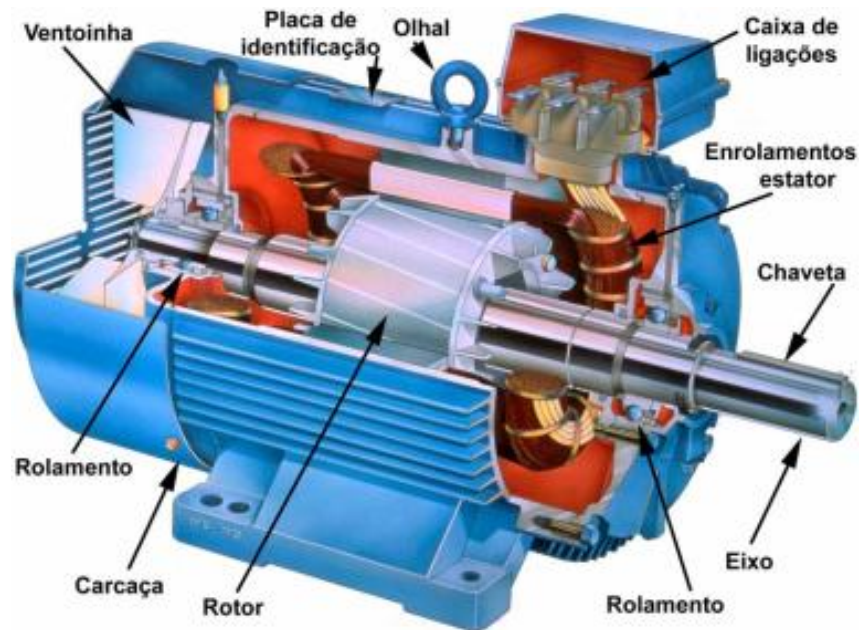


Figura 2.2: *Visão explodida de um motor de indução com rotor do tipo gaiola de esquilo (Moraís de Sousa, 2017)*

As partes da máquina estão sujeitas a falhas recorrentes a diversos tipos de problemas. (BONNETT; SOUKUP, 1992) ressalta que as falhas mais comuns acontecem nessas três principais regiões do motor. Em alguns trabalhos, como por exemplo em (PUCHE-PANADERO et al., 2009) e (LIANG; EDMOWANDEKHOE, 2017), os autores mostram um levantamento sobre a frequência de falhas que ocorrem nas referidas porções do *MIT*, de modo que cerca de 38% das falhas relatadas são ligadas ao estator, 10% ao rotor, 40% aos rolamentos e 12% em outras partes do equipamento.

É importante citar que as falhas no estator da máquina usualmente são ligadas a problemas de ordem elétrica, e no caso dos rotores, elas estão associadas a problemas elétricos e/ou mecânicos. (LIANG; EDMOWANDEKHOE, 2017) oferecem um estudo sobre a frequência de falhas em cada parte da máquina, como mostra a figura 2.3.

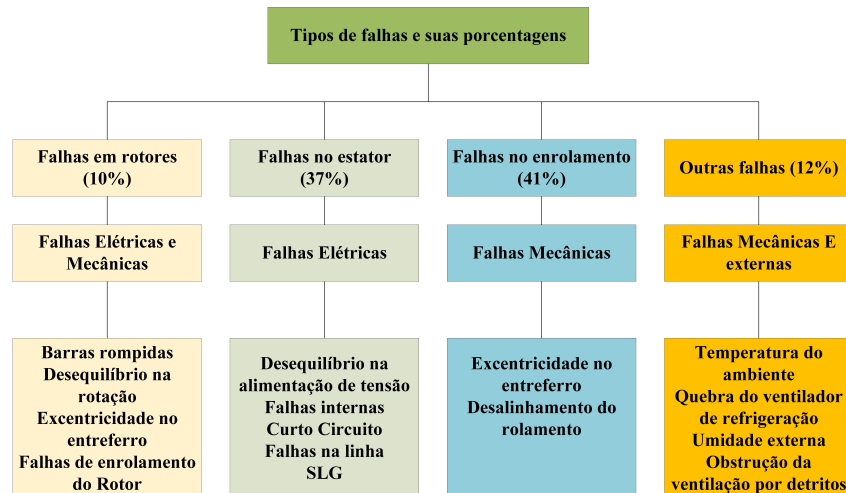


Figura 2.3: *Frequências de falhas normalmente encontradas em um motor de indução trifásico (LIANG; EDMOWANDEKHOE, 2017)*

Embora as falhas no rotor representam em média cerca de 10% do total de avarias que acometem os *MIT*, para motores de grande porte, por exemplo, essa incidência é ainda maior (DIAS; CHABU, 2014). Ademais, vale frisar que, caso esse defeito não seja tratado no início, efeitos secundários podem levar a perda não apenas do próprio rotor, mas também do estator e de todo o equipamento, caso a barra quebrada seja projetada para fora do seu núcleo. (VALLES-NOVO et al., 2014).

Neste sentido, torna-se fundamental avaliar de forma incipiente eventuais fraturas, ou o rompimento total, de uma ou mais barras na junção com o anel de curto-circuito da gaiola de um motor de indução, a fim de evitar custos adicionais para o reparo do motor e para a retomada da produção.

A figura 2.4 mostra um desenho esquemático retirado em (CHAPMAN; STEPHEN, 2013) de um rotor gaiola de esquilo, é possível visualizar as barras rotóricas, ou condutores, conectadas com o anel de curto, nas duas extremidades da barra.

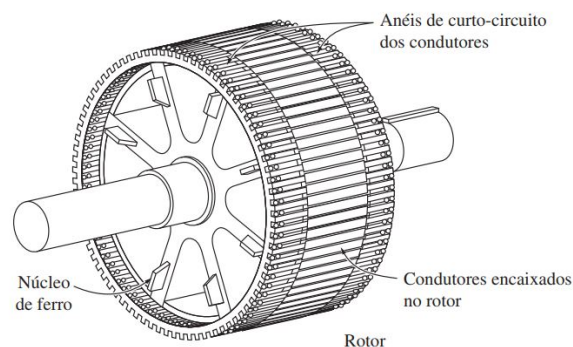


Figura 2.4: *Representação de um rotor do tipo gaiola de esquilo (CHAPMAN; STEPHEN, 2013)*

Deve-se ressaltar que, existem algumas causas que levam ao rompimento de barras

rotóricas com o anel de curto, quais sejam: elevadas correntes de partida, partidas consecutivas, defeitos de fabricação da máquina e insuficiência de ventilação.

Nas últimas décadas, diversas técnicas foram propostas para a detecção de falhas em rotores de um motor de indução trifásico, e como mostrado por (DIAS, 2006), tais técnicas pode ser resumidas em magnéticas e elétricas, ou ainda mecânicas (tabela 2.1). A análise espectral da corrente elétrica do estator é uma das mais usadas e ainda investigadas para avaliar o defeito no rotor, dada a sua facilidade de aquisição, muito embora seja necessário um hardware capaz de armazenar uma grande quantidade de dados para a extração de características responsáveis pelo diagnóstico da avaria. Uma análise mais apropriada das técnicas de processamento de sinais da corrente será apresentada no Capítulo 3.

Tabela 2.1: *Agrupamento das principais técnicas voltadas a detecção de falhas no rotor (DIAS, 2006)*

Técnicas	Descrição da técnica
Magnética e Elétricas	Análise espectral da corrente elétrica do estator
	Aplicação da transformada complexa espacial (Park)
	Análise do fluxo magnético
	Análise das descargas parciais
Mecânicas	Análise das vibrações
	Análise da velocidade
	Análise da temperatura
	Análise do torque

2.2.1 PRINCÍPIOS DE FUNCIONAMENTO DO MOTOR DE INDUÇÃO TRIFÁSICO

Como mencionado, o motor de indução, objeto deste estudo, é alimentado por tensões trifásicas normalmente equilibradas, de tal modo que essas tensões resultam em correntes que circulam pela bobina estatórica. Por sua vez, essas correntes criam um campo magnético B_e que gira no sentido anti-horário (CHAPMAN; STEPHEN, 2013). Esse campo "gira" devido ao posicionamento das bobinas, que por sua vez, defasa as tensões e correntes elétricas, variando o campo magnético produzido pelas correntes do estator.

A equação 2.1 mostra a velocidade de rotação do campo magnético resultante:

$$\eta_{sinc} = \frac{120f_s}{P} \quad (2.1)$$

Onde: η_{sinc} é a velocidade de sincronismo do campo magnético, f_s é a frequência do sistema aplicada ao estator em Hz , P é o número de polos da máquina.

A equação 2.2 mostra a tensão induzida numa dada barra *rotórica* onde $v \times B$ é um produto vetorial:

$$e_{ind} = (\mathbf{v} \times \mathbf{B}) \bullet \mathbf{l} \quad (2.2)$$

Onde:

- \mathbf{v} = velocidade da barra em relação ao campo magnético
- \mathbf{B} = vetor densidade de fluxo magnético
- \mathbf{l} = comprimento do condutor dentro do campo magnético

Um conceito principal para o funcionamento do *MIT* é que se o campo magnético gerado pelas correntes trifásicas defasadas em 120° girasse, como acontece de fato, o conjugado, que é a força de arranque e que foi induzido no rotor, resultaria numa certa "perseguição" constante com o campo magnético gerado pelas bobinas do estator da máquina, em um círculo.

Em mais detalhes, o campo magnético girante B_s induz uma tensão nas barras do rotor, que por sua vez, produz um fluxo de corrente que está defasada em relação a tensão, devido a indutância do rotor. Um campo magnético girante B_R é produzido pela corrente I_R , sendo esta defasada 90° graus elétricos, como mostra a figura. Assim, ocorre uma interação entre o campo B_R com a soma dos campos resultantes de cada bobina, o B_{liq} . É produzido, então, um conjugado anti-horário na máquina, como explica o autor (CHAPMAN; STEPHEN, 2013) e ilustrado na figura 2.5.

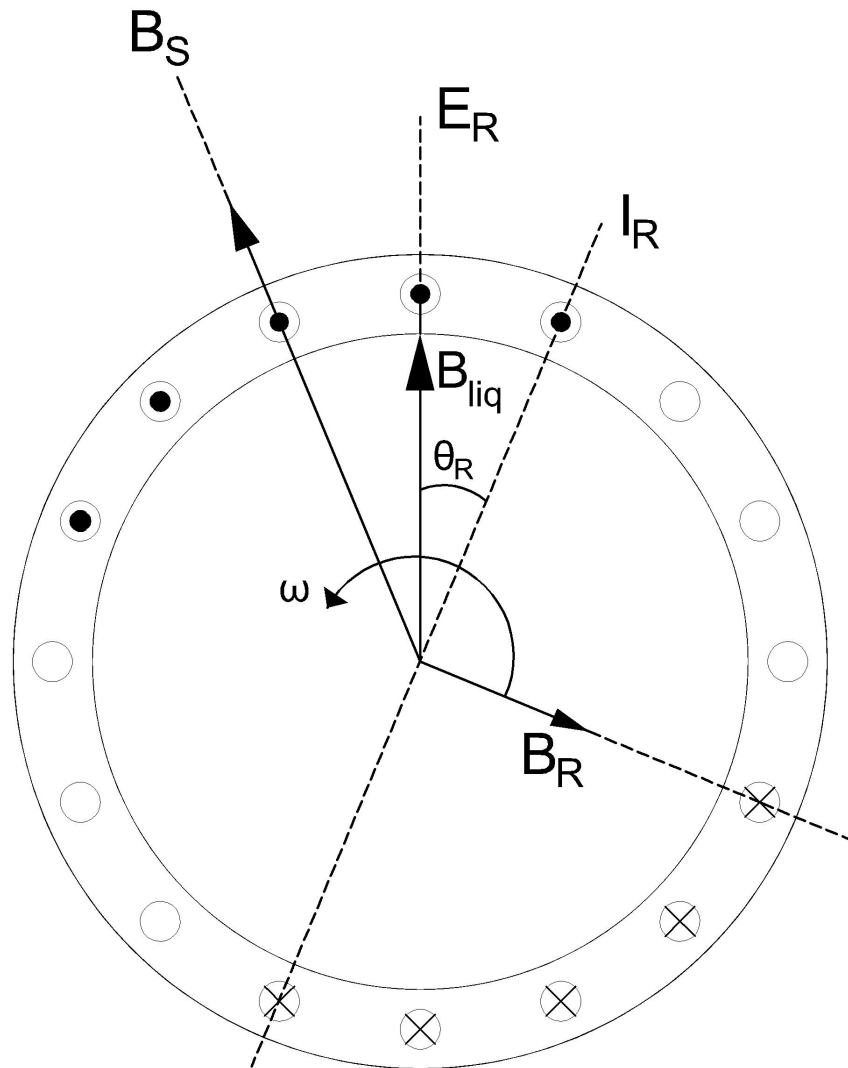


Figura 2.5: A produção de conjugado induzido em um motor de indução (CHAPMAN; STEPHEN, 2013)

- ω = Velocidade angular
- θ = Ângulo entre tensão induzida no rotor e corrente no rotor
- B_{lq} = Soma dos campos resultantes em cada bobina
- B_s = campo magnético girante gerado pelas correntes da bobina do estator da máquina
- E_R = Tensão induzida no rotor
- I_R = Corrente fluindo nas barras do rotor da máquina

Esse conjugado resultante, como já explicado é anti-horário e o rotor acelera nesse sentido. O fluxo de corrente do rotor, induzido pelo estator, também produz um campo

magnético no próprio rotor \mathbf{B}_R . A equação 2.3 revela o conjugado induzido na máquina, que é o produto escalar entre B_R e B_s .

$$\tau_{ind} = k \mathbf{B}_R \times \mathbf{B}_s \quad (2.3)$$

Onde k é um fator que depende da construção da máquina

Vale ressaltar que, caso o rotor estivesse em uma velocidade síncrona, não haveria variação do fluxo magnético em relação às espiras, ou barras do rotor, e portanto, o conjugado seria nulo com $e_{ind} = 0$.

2.3 O ESCORREGAMENTO DO MOTOR

O escorregamento é uma grandeza adimensional e é diretamente dependente da velocidade do rotor. Tal grandeza é definida pela diferença entre a velocidade de sincronismo, ou seja a velocidade do campo girante e a velocidade do eixo do rotor da máquina, sendo esta diferença dividida pela velocidade de sincronismo. Esse resultado é mostrado na equação 2.4.

Logo, o escorregamento é expresso como uma fração da velocidade síncrona. O variação do escorregamento está entre 0, na condição do rotor girando na mesma velocidade do campo girante, e 1, quando o rotor está em repouso. É possível expressar ainda o escorregamento em termos de velocidade angular, como visto na equação 2.5.

$$s = \frac{n_{sinc} - n_m}{n_{sinc}} (\times 100\%) \quad (2.4)$$

Onde:

n_{sinc} = velocidade síncrona do campo girante

n_m = velocidade mecânica do eixo do motor

$$\frac{\omega_{sinc} - \omega_m}{\omega_{sinc}} (\times 100\%) \quad (2.5)$$

ω_{sinc} = velocidade angular dos campos magnéticos

ω_m = velocidade angular mecânica do eixo do motor

As equações 2.6 e 2.7 são as velocidades mecânicas encontradas no eixo da máquina e podem ser expressadas em termos de velocidade síncrona e de escorregamento.

$$n_m = (1 - s)n_{sinc} \quad (2.6)$$

$$\omega_m = (1 - s)\omega_{sinc} \quad (2.7)$$

2.4 A ANÁLISE DE *Fourier*

A análise de *Fourier* foi originalmente criada pelo matemático francês **Jean Baptiste Joseph Fourier** (1768 - 1830) para resolver um problema de distribuição de temperaturas, e nas últimas décadas foi bastante empregada na área de processamento digital de sinais, especialmente para a decomposição de sinais.

Neste caso, o seu propósito é mostrar um espectro de frequências e informações normalmente não disponíveis no domínio do tempo. Em outras palavras, as informações contidas no sinal do domínio da frequência continuam as mesmas em comparação com o sinal do domínio do tempo, nada é perdido, o que difere, é a forma como o sinal é apresentado. Na análise de *Fourier* o sinal é decomposto e transformado em uma somatória de cossenos (parte real) e senos (parte imaginária). Existem quatro tipos básicos de sinais e por sua vez, uma transformada para cada um desses tipos (Steven W. Smith, 1999).

- Para sinais contínuos e periódicos - *Série de Fourier*
- Para sinais contínuos e não periódicos - *Transformada de Fourier*
- **Para sinais discretos e periódicos - Transformada Discreta de Fourier**
Discret Fourier Transform (DFT)
- Para sinais discretos e não periódicos - *Transformada Discreta de Fourier no Tempo*

Em (Steven W. Smith, 1999), é descrito que a *DTF* pode ser utilizada em sinais discretos e periódicos finitos no tempo, pois para implementá-la em um computador é necessário amostrar o sinal, ou seja, trabalhar com amostras finitas no tempo. Portanto, o processo para extrair o espectro de frequência do sinal no domínio do tempo é chamado de decomposição, análise ou simplesmente *DFT* (Steven W. Smith, 1999).

A figura 2.6 explica como funciona basicamente a *DFT*, onde $X[]$ representa o sinal de entrada, no domínio do tempo. Entre colchetes são alocados os valores que representam cada endereço (indexador) do vetor em questão. A letra N representa as amostras de

entrada, no qual cada indexador do vetor representa os seus pontos. Por exemplo, $X[4]$ representa o quarto valor do vetor X , ou seja, o ponto quatro que está entre 0 e $N - 1$.

A *DFT* transforma um sinal no domínio do tempo para o domínio da frequência, trocando as amostras de entrada, representadas pelo vetor N , e produzindo dois sinais de saída com o N dividido por dois ($N/2 + 1$). Esse sinal resultante possui duas partes, $ReX[]$ - parte real (amplitude dos cossenos) e $ImX[]$ - parte imaginária (amplitude dos senos). Cada uma dessas partes vai de 0 a $N/2 + 1$, ou seja, cada vetor de saída recebe metade do tamanho do vetor de entrada. Em outras palavras, a partir de um sinal no domínio do tempo, a *DFT* produz dois sinais no domínio da frequência ($ReX[]$ e $ImX[]$).

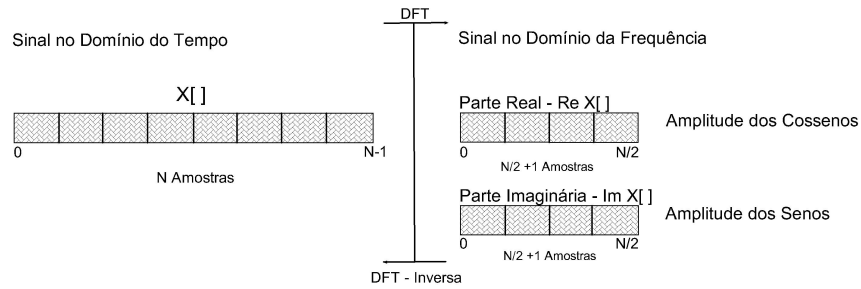


Figura 2.6: Funcionamento da *DFT*, adaptado de (Steven W. Smith, 1999)

As amplitudes dos cossenos e senos, como já discutido, se forem somadas, o resultado se aproximaria do sinal no domínio do tempo. A *DFT* inversa, a partir da somatória das amplitudes dos senos e cossenos, tem como resultado esse sinal, como mostra as equações abaixo (Steven W. Smith, 1999).

As equações 2.8 e 2.9 mostram as funções cosseno e seno, que são chamadas de **funções de base**

$$Cos_{freq}[in] = \cos\left(\frac{2\pi \cdot freq \cdot in}{N}\right) \quad (2.8)$$

$$Sen_{freq}[in] = \sin\left(\frac{2\pi \cdot freq \cdot in}{N}\right) \quad (2.9)$$

Onde in representa o indexador que corre todo o vetor e varia entre 0 a $N - 1$. A variável $freq$ é a frequência do sinal, variando de 0 até $N/2$ ciclos a cada vetor de pontos N . $Cos_{freq}[in]$ representa a função de base cosseno em que está contido a parte real ($Re X[freq]$) e $Sen_{freq}[in]$ representa a função de base seno em que está contido a parte imaginária ($Im X[freq]$) ambas do sinal no domínio do tempo.

A equação 2.10 representa o somatório das amplitudes das funções de base. Que retorna

o valor do sinal no domínio do tempo ($X[n]$).

$$X[n] = \sum_{freq=1}^{N/2} Re\bar{X}[Freq] \cos\left(\frac{2\pi \cdot freq \cdot n}{N}\right) + \sum_{freq=1}^{N/2} Im\bar{X}[Freq] \sin\left(\frac{2\pi \cdot freq \cdot n}{N}\right) \quad (2.10)$$

$Re\bar{X}[Freq]$ e $Im\bar{X}[Freq]$ são as amplitudes das funções cosseno e seno necessárias para a *DFT* inversa. As equações 2.11 e 2.12 mostram como são calculadas:

$$Re\bar{X}[Freq] = \frac{ReX[freq]}{N/2} \quad (2.11)$$

$$Im\bar{X}[Freq] = \frac{ImX[freq]}{N/2} \quad (2.12)$$

Onde: $ReX[freq]$ e $ImX[freq]$ são as partes reais e imaginárias do domínio da frequência.

As amplitudes em $freq = 0$ e $freq = N/2$ são casos especiais, como mostra as equações 2.13 e 2.14:

$$Re\bar{X}[Freq] = \frac{ReX[freq]}{N/2} \quad (2.13)$$

$$Im\bar{X}(Freq) = \frac{ImX[freq]}{N/2} \quad (2.14)$$

2.4.1 A TRANSFORMADA RÁPIDA DE FOURIER (*FFT*)

Para calcular a *DFT* existe basicamente três maneiras distintas:

- Matematicamente, resolvendo equações lineares simultâneas;
- Através do método de correlação;
- **Utilizando o algoritmo da *Transformada Rápida de Fourier* (*FFT*)**

A *FFT*, que será usada nesse estudo, difere dos outros métodos pois, é muito mais rápida. Ou seja, exige menos complexidade computacional, reduzindo o tamanho da complexidade de N^2 para aproximadamente $N \log N$ operações (N representa o tamanho do vetor) isso, reduz o número de operações drasticamente.

Ela foi originalmente implementada por *J.W. Cooley* and *J.W. Tukey* no estudo intitulado "*An algorithm for the machine calculation of complex Fourier Series*," (COOLEY; LEWIS; WELCH, 1967). Vale ressaltar, que o matemático *Karl Friedrich Gauss* (1777-1855), propôs uma introdução da ideia, contudo, na época, ainda não havia "estourado" a revolução computacional, portanto, não tendo utilidade na prática.

A ideia principal da *FFT* foi em implementar um algoritmo que calculasse a *DFT*, reduzindo o seu custo computacional, por isso ela tornou-se conhecida como um eficiente algoritmo para calcular a *DFT* (HEIDEMAN; JOHNSON; BURRUS, 1984).

Vale pontuar que o algoritmo da *FFT* é baseado na *DFT* complexa (na qual o sinal de entrada é um número complexo), contudo, é importante dizer, que ela é também utilizada para calcular a *DFT* com números *reais*, bastando transformá-los em números complexos, com uma parte imaginária formada por zeros $0j$ onde j representa a raiz quadrada de $\sqrt{-1}$.

A explicação para a complexidade computacional da *DFT* esta nas operações (multiplicações) envolvendo matrizes, isso revela um certo problema, operações com matrizes tornam o algoritmo muito lento. Por esta razão, o algoritmo proposto por *Cooley-Tukey* (*FFT*) é muito importante e necessário pois, não envolve multiplicações com matrizes, reduzindo assim, em muito, a sua complexidade computacional.

Ela funciona resumidamente em três passos, como revela o autor (Steven W. Smith, 1999):

1. O primeiro passo é a decomposição entrelaçada, que consiste em decompor o sinal de duas em duas partes, inicialmente divididos em par e ímpar. O algoritmo colhe sempre o primeiro e o terceiro ponto de cada parte do vetor que foi dividido, em primeiro instante, o vetor com os pontos pares e com os pontos ímpares. Por exemplo, um sinal com 8 pontos é dividido, tornando-se dois sinais com quatro pontos, que por sua vez transforma-se em quatro sinais com dois pontos e assim sucessivamente até que transforma-se em oito sinais com um ponto cada.
2. O segundo passo é encontrar o espectro de frequência em cada ponto. O sinal que foi multiplicado, transformado-se em vários sinais com um ponto apenas (espectro de frequência) que já está no domínio da frequência.
3. O último passo é combinar esses vários pontos com N espectros de frequência em um sinal apenas. Em outras palavras, é feito o reverso do que foi feito no primeiro passo, os pontos são *shintetizados* até que são transformados em um espectro de frequência (chamado de diagrama **borboleta**) com 8 pontos que é a saída do *FFT*.

Esse é resumidamente o funcionamento da *FFT*, o autor (Steven W. Smith, 1999) explica com detalhes mais ricos o seu funcionamento.

2.5 A TRANSFORMADA DE HILBERT (HT)

No ano de 1743 *Leonard Euler* (1707 - 1783) em seus estudos derivou a seguinte equação (STILLWELL, 2010):

$$e^{jz} = \cos(z) + j \sin(z) \quad (2.15)$$

Anos mais tarde, a partir desse estudo, o físico *Arthur E. Kennelly* e o cientista *Charles P. Steinmetz* (Karl L. Wildes, Nilo A. Lindgren, 1985) introduziram uma notação complexa das formas de ondas harmônicas usadas na área de engenharia elétrica, como revela a equação 2.16.

$$e^{j\omega t} = \cos(\omega t) + j \sin(\omega t) \quad (2.16)$$

Onde: j representa a parte imaginária de um número complexo, ω é a velocidade angular e t o tempo.

O matemático *David Hilbert* mostrou que a função $\sin(\omega t)$ é a transformada de *Hilbert* do $\cos(\omega t)$. Em outras palavras, ele demonstrou que há uma defasagem com um operador $\pm \frac{\pi}{2}$ de mudança de fase.

A transformada de *Hilbert* foi criada pelo matemático *David Hilbert*, nascido na Alemanha no ano de 1862. A sua definição é dada pela convolução entre um sinal real no domínio do tempo, $x(t)$ e a transformada de *Hilbert* do impulso $\frac{1}{\pi t}$ como mostra a equação.

A equação 2.17 mostra a transformada de *Hilbert* do sinal real $x(t)$ (PUCHE-PANADERO et al., 2009) e (XU et al., 2013):

$$\hat{x}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x_s(\tau)}{t - \tau} d\tau \quad (2.17)$$

Onde: $\tau = t$ (tempo)

A transformada de *Hilbert* adianta a fase do sinal em $\frac{\pi}{2}$ radianos ou 90° graus fazendo a convolução no tempo por $\frac{1}{\pi t}$ que é o mesmo que multiplicar por $j \cdot \text{sgn}(\omega)$ (SETTE, 2017). Isso não provoca alteração na amplitude, apenas desloca a fase, transformando uma função cosseno em seno, como revela (XU et al., 2013).

Assim, a *HT* não altera o domínio do sinal, pois este continuará em seu domínio original, mas sim, defasa o ângulo do sinal em todas as frequências, e segundo (SETTE, 2017) induz um atraso em todas as frequências do sinal no domínio do tempo, fazendo com que o valor médio quadrático, o chamado valor (*rms*) de um sinal, e o da respectiva transformada sejam iguais.

A *HT* possui diversas aplicações, seja no campo da engenharia elétrica, eletrônica, sismologia, vibrações mecânicas, medicina e etc. Abaixo alguns exemplos de aplicações (SETTE, 2017), (PUCHE-PANADERO et al., 2009).

- Análise de vibração em motores e rolamentos;
- Diagnósticos de falhas em rotores de motores de indução;
- Modulação e demodulação de sinais em sistemas de comunicação;
- Processamento e codificação de imagens;

- Eletro medicina; e
- Análise de sinais sismológicos.

O sinal real x_t e a sua HT \hat{x}_t estão relacionadas e criam o sinal analítico, segundo (JOHANSSON, 1999) a transformada de *Fourier* desse sinal analítico resulta em um espectro unilateral no domínio da frequência.

A equação 2.18 mostra o sinal analítico. Ele é criado pela soma do sinal real no tempo, pela sua parte imaginária.

$$\bar{x}(t) = x_s(t) + j\hat{x}(t) \quad (2.18)$$

Onde: $\bar{x}(t)$ é o sinal analítico, $x(t)$ é a parte real e $j\hat{x}(t)$ a parte imaginária do sinal estudado, ainda no domínio do tempo.

O modulo do sinal analítico é definido por (XU et al., 2013) na equação 2.19:

$$Mod\bar{x}(t) = \sqrt{x^2(t) + j\hat{x}^2(t)} \quad (2.19)$$

O modulo do sinal analítico é usado para extrair, através da *FFT*, o espectro de frequência, que por exemplo, foi usado nesse trabalho, para extrair componentes importantes, localizadas no espectro de frequência do sinal de corrente do estator de um *MIT* para diagnosticar barras rompidas, que será melhor explicado ao longo dos capítulos.

Ná prática a *HT* segundo (XU et al., 2013) e (PUCHE-PANADERO et al., 2009) pode ser usada para diagnóstico de barras rompidas e funciona nos seguintes passos:

1. Mensurar o sinal da corrente de uma das fases do estator da máquina $x(t)$;
2. Calcular a *HT* do sinal de corrente do estator $x(t)$; ficando como $\hat{x}(t)$, equação 2.17;
3. Extração do sinal analítico, que é feito pela equação 2.18;
4. Calculo do modulo do sinal analítico com o sinal real 2.19;
5. Remoção da componente *DC* do módulo do sinal calculado pela equação 2.19 subtraindo com a média do vetor do modulo, $Mod_{semdc}(t) = Mod(t) - \text{média}$;
6. Obtenção do espectro de frequência do Mod_{semdc} , usando a *FFT*, investigando assim uma certa componente indicadora de falhas no rotor da máquina no espectro de frequência.

2.6 MEDIDAS ESTATÍSTICAS EXTRAÍDAS DA CORRENTE DO MOTOR

A seguir são apresentadas, na tabela 2.2, as equações utilizadas neste trabalho para extrair as principais características estatísticas do sinal de corrente do *MIT*. Tais grandezas

foram escolhidas tomando por base o trabalho de (FERNANDEZ-TEMPRANO et al., 2013). Assim, as características mostradas na tabela 2.2 são obtidas a partir da corrente de uma fase da máquina, no domínio do tempo e após o processo de filtragem, uma vez que o motor foi alimentado por um inversor de frequência.

Tabela 2.2: *Medidas Estatísticas (FERNANDEZ-TEMPRANO et al., 2013)*

Medidas Estatísticas		
	Medidas	Equações
1	Desvio Padrão	$\sigma = \frac{\sum (x_i - \bar{x})^2}{n-1}$
2	Média	$m_1 = \frac{1}{n} \sum (x)$
3	Variância	$m_2 = \frac{1}{n} \sum (x_i - \bar{x})^2$
4	Momento 3ª Ordem	$m_3 = \frac{1}{n} \sum (x_i - \bar{x})^3$
5	Momento 4ª Ordem	$m_4 = \frac{1}{n} \sum (x_i - \bar{x})^4$
6	Momento 6ª Ordem	$m_6 = \frac{1}{(n \times m_2) \sum (x_i - \bar{x})^6}$
7	Cumulante 1ª Ordem	$c_1 = m_1$
8	Cumulante 2ª Ordem	$c_2 = m_2 - m_1^2$
9	Cumulante 3ª Ordem	$c_3 = m_3 - 3m_1m_2 + 2m_1^3$
10	Cumulante 4ª Ordem	$c_4 = m_4 + m_3m_1 - 3m_2^2 + 12m_2^2 - 6m_1^4$
11	Assimetria	$Ass = \frac{m_3}{(\sqrt{m_2})^3}$
12	Média Absoluta	$ \bar{x} = \frac{1}{n} \sum x $
13	Valor de Pico Máximo	$x_p = \max x $
14	Valor SRV	$x_r = (\frac{1}{n} \sum \sqrt{ x })^2$
15	Fator de Crista	$f_c = \frac{x_p}{x_{rms}}$
16	Fator de Onda	$f_o = \frac{x_{rms}}{ x }$
17	Valor RMS	$x_{rms} = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$

2.7 APRENDIZAGEM DE MÁQUINA

A aprendizagem de máquina, ou *Machine Learning (ML)*, é um dos ramos da inteligência artificial (*IA*) e é aplicada na solução de problemas em diversos ambientes produtivos, a exemplo do que ocorre para o diagnóstico de falhas nos *MIT*.

A *ML* é frequentemente empregada na extração de conhecimento ou na avaliação de possíveis padrões em um conjunto de dados. Segundo (BAKSHI; BAKSHI, 2018), o *ML* está se tornando comum na vida cotidiana e é usado, entre outras aplicações, em carros autônomos, buscas na internet, seleção de *email* com spam, predição de doenças e até mesmo no entendimento do genoma humano.

O conceito principal em *ML* é oferecer um algoritmo capaz de aprender apenas com um conjunto de dados. Eles são muitas vezes agrupados em uma matriz, de modo que as linhas representam as instâncias e as colunas as características, ou atributos, do conjunto de dados. A escolha da saída, ou atributo alvo, é definida pelo programador com o intuito de ser predita logo após a criação do modelo. Esse processo é chamado de aprendizado supervisionado, uma vez que o algoritmo aprende com um conjunto de dados, a partir de uma saída conhecida. Deve-se, neste caso, avaliar a distância entre a resposta atual e aquela desejada.

Neste sentido, a área de *ML* é basicamente dividida em duas abordagens, quais são: a aprendizagem supervisionada e a não supervisionada. Este trabalho comparou apenas o desempenho dos modelos treinados utilizando o aprendizado supervisionado, uma vez que a saída desejada é conhecida, no caso a condição física do rotor, entre saudável e avariado (barra quebrada).

2.7.0.1 Aprendizagem Supervisionada

A aprendizagem supervisionada é usada para prever uma saída (atributo alvo), a partir de um modelo já treinado com um conjunto separado da base de dados. Esta base de dados é normalmente separada em duas partes, uma para treinamento e outra para validação.

Deste modo, o algoritmo aprende com os dados de treinamento na qual a coluna alvo e os atributos de entrada são empregados para a criação de um modelo. Em seguida é usado um conjunto de dados de validação, que não foi utilizado na fase de treinamento, sem a coluna alvo, de modo que o modelo já criado seja empregado para prever uma nova saída. Os dados de validação ratificam o aprendizado alcançado pelo algoritmo. Esse caso é chamado de problema de aprendizagem supervisionada, classificatório ou problema de classificação. Em outras palavras, um problema de classificação procura, na maioria das vezes, prever, com o apoio de um conjunto de dados, saídas discretas ou binárias (NG, 2012).

No caso do presente estudo, há a necessidade de prever o estado do motor de indução, mais precisamente a condição estrutural da gaiola do rotor. A saída é binária, logo para duas condições, de modo a indicar valor igual a "0" para um rotor saudável, e valor igual a "1" para o rotor com uma barra rompida, portanto, trata-se de uma tarefa de classificação. O conjunto de dados é estruturado com as respectivas características e com a saída, coluna alvo, com os valores 0 e 1. Deste modo, o classificador deverá prever a saída (condição da máquina) a partir dos dados de validação, que foi predito através de um modelo já treinado. Existem alguns tipos de algoritmos utilizados para a tarefa de classificação em *ML* como por exemplo: Árvores de decisão, *Naive Bayes*, Redes *Bayesianas*, Redes Neurais, Máquina de vetor de suporte, Regressão logística, Método ou técnica dos vizinhos

mais próximos e etc.

A seguir são descritos alguns passos básicos para a implementação desses algoritmos baseados em *ML*:

1. Coleta dos dados;
2. Preparação dos dados (normalização);
3. Análise (quais *características* são realmente necessárias);
4. Treinamento do modelo; e
5. Validação.

Este trabalho utiliza apenas três algoritmos baseados em *ML* para a tarefa de classificar a condição do rotor do *MIT*, quais são: A máquina de vetor de suporte, a técnica ou método dos vizinhos mais próximos e a regressão logística. Eles foram escolhidos uma vez que são bastante empregados atualmente em problemas de classificação no campo das máquinas elétricas.

2.7.0.2 A Regressão Logística

A Regressão Logística *RL* é um modelo no qual a variável resposta pode assumir dois valores, ou seja, valores binários. É um classificador baseado em aprendizagem supervisionada e que a partir de um conjunto de dados, com uma saída prevista, é capaz de criar um modelo que permite a predição em função de um conjunto de dados de entrada. Assim, a *RL* é usada em situações nas quais é possível classificar modelos com saídas discretas.

Apesar do nome regressão ser usualmente empregado em problemas com saídas contínuas, ela é aplicada também em tarefas de classificação, pois usa o mesmo método da regressão linear para o ajuste da fronteira de decisão no conjunto de dados para a sua classificação. Portanto o que difere a regressão logística, por exemplo, da regressão linear, é o tipo de função usada para isso, na regressão linear utiliza-se a função que representa uma reta (hipótese), já na regressão logística aplica-se, portanto, uma função chamada de *sigmoide* (HARRINGTON, 2015).

Em (Claude Sammut, 2011), é definida a *RL* como um modelo que fornece um mecanismo capaz de aplicar a técnica de regressão linear em um problema de classificação. Em (NG, 2012), o autor revela que utilizar um algoritmo de regressão linear para saídas discretas torna a sua performance muito limitada. Tal fato tem relação com o uso da função *hipótese* (h_θ) utilizada na regressão linear, conforme descrita na equação 2.20.

Por outro lado, na regressão logística a função $h_\theta(x)$ torna-se diferente, como mostram as equações 2.21 e 2.22, representando a função *Sigmoide* ($g(\theta^T(x))$). A figura 2.7 representa a função descrita nas equações citadas.

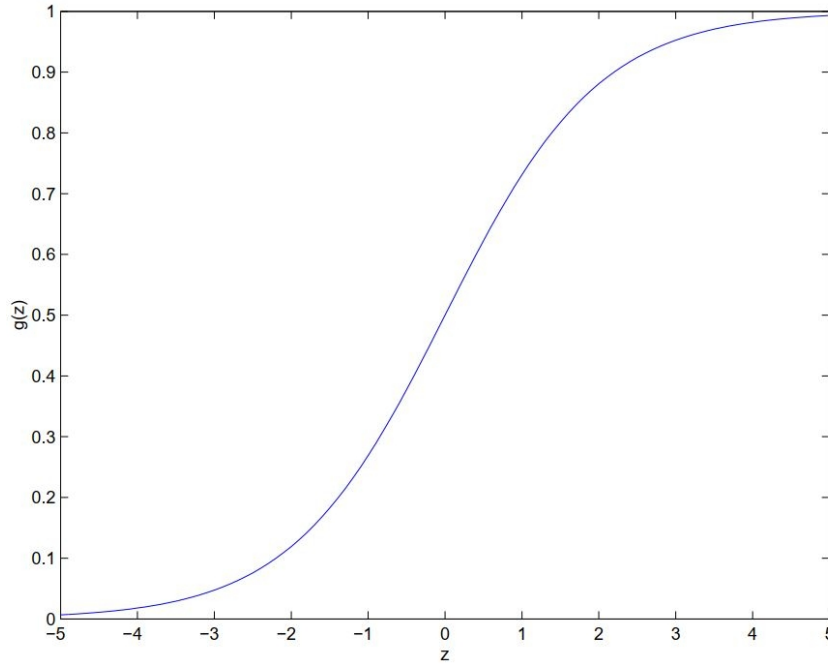


Figura 2.7: Função Sigmoide (NG, 2012)

Quando $g(z)$, na figura 2.7, se aproxima de zero, $z \rightarrow -\infty$ e quando $g(z)$ se aproxima de um, $z \rightarrow \infty$. Se z estiver com um valor acima de 0,5, $g(z) = 1$ e abaixo de 0,5, $g(z) = 0$. Isso nos mostra que a função $g(z) = h_\theta(x)$ estará entre valores de 0 a 1. (NG, 2012), (HARRINGTON, 2015).

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \quad (2.20)$$

$$h_\theta(x) = g(\theta^T(x)) = \frac{1}{1 + e^{-\theta^T(x)}} \quad (2.21)$$

$$g(z) = \frac{1}{1 + e^{-z}} \quad (2.22)$$

Onde:

θ representa os pesos, x_1 até x_n os valores de de entrada.

Para que a função se ajuste aos dados, torna-se necessário descobrir os valores dos 'pesos', ou coeficientes da equação *sigmoide* (θ). Esses coeficientes precisam ser otimizados de modo que a função se ajuste ao conjunto de dados, ou seja, minimizar os coeficientes. Para isso, é usado o método do *gradiente ascendente*, sendo este um método de otimização para encontrar os melhores parâmetros para do modelo.

O gradiente ascendente encontra a direção para o ponto máximo da função $f(x, y)$, e é definido pelas equações 2.23 e 2.24.

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \quad (2.23)$$

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial y} \quad (2.24)$$

O gradiente indica a direção necessária para se maximizar a função. A equação 2.23 aponta para a direção 'x' e a equação 2.24 aponta para a direção 'y'.

Em notação vetorial, a equação do gradiente ascendente pode ser escrita de acordo com a equação 2.25

$$\theta := \theta + \alpha \nabla_{\theta} f(\theta) \quad (2.25)$$

Como já comentado, o gradiente ascendente apenas indica a direção onde o ponto máximo da função se encontra, contudo, o valor de α é um parâmetro que indica o tamanho dos 'passos' que o algoritmo deverá caminhar. A figura 2.8 mostra como o algoritmo funciona. A cada passo o algoritmo se aproxima do ponto máximo da função, iniciando no P_0 e a cada iteração o gradiente fornece a direção e o α o tamanho do passo.

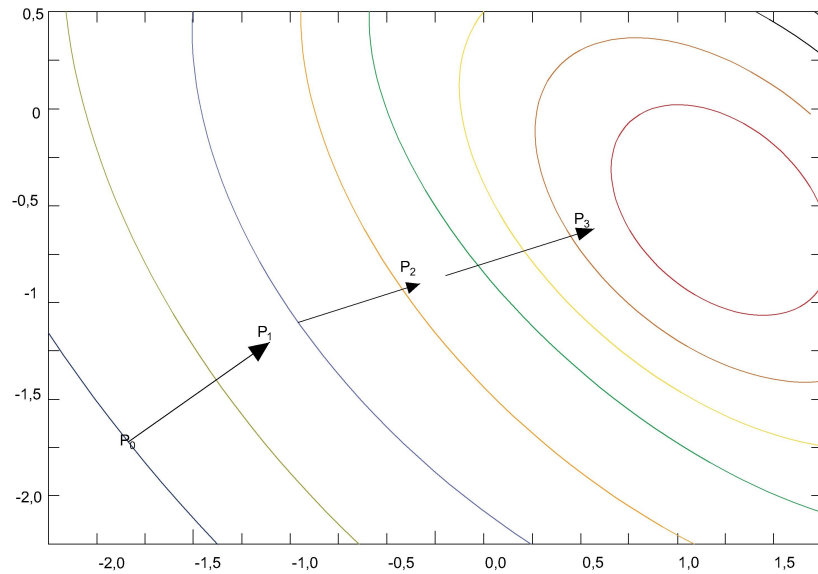


Figura 2.8: Funcionamento do gradiente ascendente, adaptado de (HARRINGTON, 2015)

O gradiente ascendente encontra, portanto, os melhores valores dos θ que são os 'pesos' da função *sigmoide* que melhor se ajustam ao conjunto de dados, a fim de classificá-los.

2.7.0.3 As Máquinas de Vetores de Suporte

As máquinas de vetores de suporte, ou *Support Vector Machine (SVM)*, são uma classe de algoritmos lineares usados principalmente para classificação e regressão (Claude Sammut, 2011). O SVM encontra um hiperplano de separação entre os dados, classificando assim as regiões entre as duas classes e criando uma grande margem entre o hiperplano e

dois vetores de suporte (Gordi Armaki; ROSHANFEKR, 2010).

Em (MARTIN-DIAZ et al., 2018), o autor define *SVM* como um tipo de modelo estatístico usado para converter classificadores lineares em modelos capazes de produzir limites de decisão não-lineares, e esse modelo acha o melhor hiperplano que separa as classes usando um problema de otimização atribuindo novas observações.

A figura 2.9 mostra três condições (três retas) para a classificação do conjunto de dados no gráfico de dispersão. É possível observar que todas as retas classificaram os dados de uma maneira satisfatória, contudo, há a necessidade de saber se esses resultados vão apresentar uma boa performance para novos exemplos, ou padrões de entrada (KUBAT, 2015).

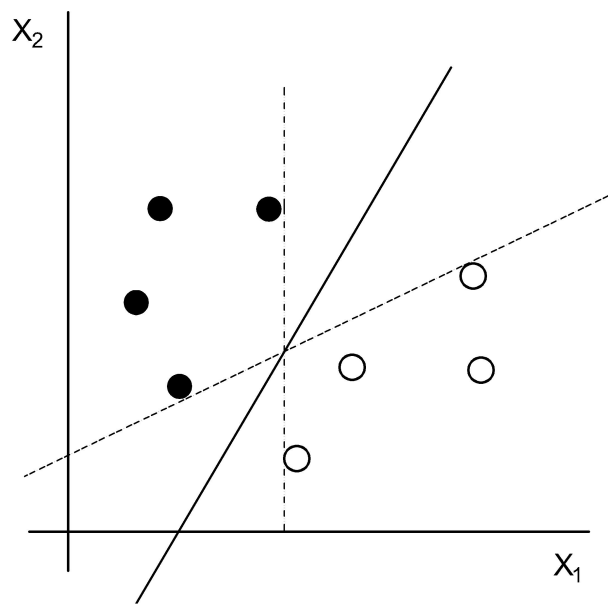


Figura 2.9: Classificação de duas classes, adaptado de (KUBAT, 2015)

É possível observar que a fronteira da classificação com as linhas pontilhadas se aproximam bastante das amostras, em ambas as classes. Nota-se, portanto, que a margem nesses dois casos de classificação é muito pequena. Por outro lado, a margem do classificador, com a linha mais escura é muito maior que as das outras duas classificações, isso resolve o problema para saber qual é a melhor classificação para novos exemplos, ou seja, melhora as chances do classificador ter uma boa performance para novos dados.

A figura 2.10 ilustra melhor o princípio de funcionamento de um classificador *SVM*. É possível observar que a linha mais escura oferece a melhor classificação para esse conjunto de dados, quando comparada aquelas linhas mais finas que estão paralelas a ela. Logo, o principal objetivo dessa técnica é identificar quais são os melhores vetores de suporte que maximizam a margem (KUBAT, 2015).

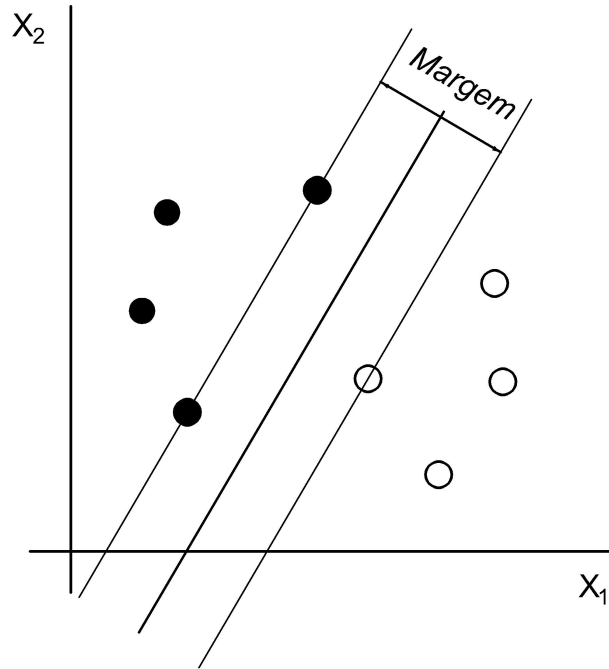


Figura 2.10: Exemplo de classificação usando o SVM, adaptado de (KUBAT, 2015)

O autor (ALPAYDIN, 2004) também mostra que uma ótima separação, ou seja, um hiperplano ideal é formado por margens maximizadas. Ainda em (ALPAYDIN, 2004), o autor diz que se a classificação não for linearmente separável, ao invés de utilizar funções não-lineares para separar as classes, é possível utilizar uma nova dimensão para o plano, ou seja, fazer uma transformação para um novo espaço com novas dimensões. Em outras palavras, o modelo que agora é linear no novo espaço, correspondeu a um modelo não-linear do espaço original, e tal comportamento é alcançado com o uso do chamado *Kernel*.

(ALPAYDIN, 2004) ainda ressalta que a ideia do *Kernel* é trocar o produto interno da função de base por uma função *Kernel* entre o vetor de suporte e a entrada do espaço original. Existe alguns tipos de função *kernel* e os avaliados nesse trabalho foram o *Gaussiano* e o *Linear*.

Existem alguns parâmetros de regularização usados na SVM, como por exemplo, o parâmetro **C**, que é um fator de penalidade do modelo e que evita uma classificação errada, ou um *overfitting*. Para um valor elevado de **C**, a otimização deverá escolher um hiperplano de separação de margem pequena e ao contrário, um hiperplano de separação de margem grande.

O parâmetro *Gamma* (γ), também utilizado nesse estudo, indica quais pontos serão considerados em relação à fronteira de separação, ou seja, com um γ baixo, pontos distantes da fronteira são considerados e com um γ alto o oposto é verdadeiro, ou seja, pontos mas próximos serão avaliados.

2.7.0.4 Método do Vizinho Mais Próximo

O algoritmo de classificação baseado nos vizinhos mais próximos (Nearest Neighbor – *KNN* em inglês) foi proposto inicialmente por (FUKUNAGA; NARENDRA, 1975). O *KNN* baseia-se na avaliação das amostras vizinhas, logo, o classificador determina o rótulo de classificação de um exemplo novo baseado nas amostras vizinhas de um conjunto de dados. Em (DIAS; PEREIRA, 2018), por exemplo, o autor diz que o *KNN* usa um método não paramétrico e atribui a cada padrão não rotulado ao *cluster* de seus vizinhos rotulados mais próximos.

Neste sentido, o *KNN* é um algoritmo de aprendizado não paramétrico. O funcionamento básico consiste em categorizar os exemplos de um conjunto de dados, com um pressuposto de que os dados similares estão concentrados em uma região no espaço de dispersão mais próximo. A figura 2.11 mostra o agrupamento natural em um plano de dispersão para algumas características relacionadas a espécies de flores. O *KNN* considera, então, a proximidade dos dados relevantes para a predição de novos exemplos.

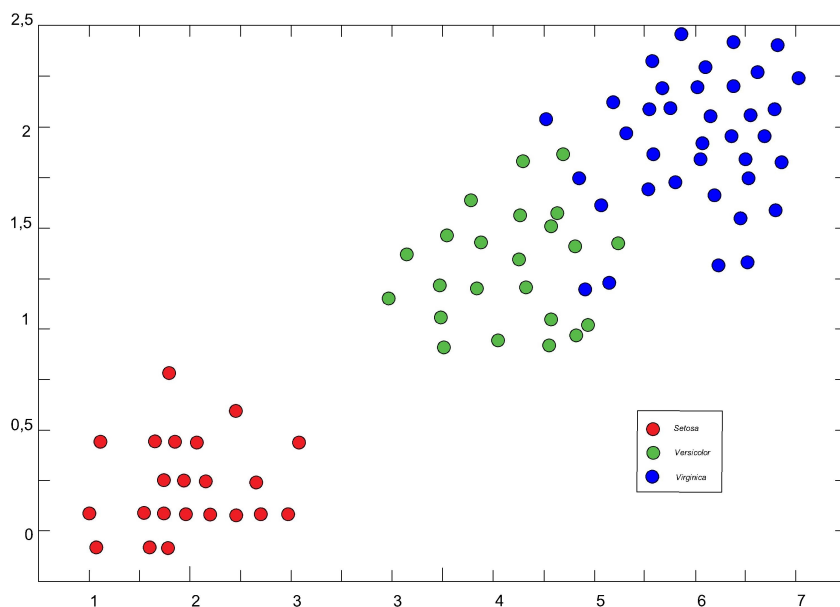


Figura 2.11: Agrupamento de dados de três tipos de flores de acordo com o tamanho da pétala, adaptado de: <https://www.mathworks.com/help/stats/classification-using-nearest-neighbors.html>

Deste modo, quanto mais próximas as características estão uma da outra, maior será a sua similaridade, ou seja maior a chance delas pertencerem a mesma classe. O *KNN* categoriza novos exemplos pela sua proximidade com as classes já rotuladas, portanto, ele calcula a distância entre o novo exemplo e os seus vizinhos mais próximos, atribuindo um rótulo da classe que estiver mais perto. O *K* representa a quantidade de vizinhos mais próximos.

Na figura 2.12, adaptado de (KUBAT, 2015), existem alguns novos exemplos para a

classificação entre duas classes já rotuladas em positivas (+) e negativas (-). Os novos exemplos são indicados pelos círculos pretos (●), dos quais o *KNN* deverá classificá-los. Para esses novos exemplos, as amostras 1 e 3 estão posicionadas próximas às suas respectivas classes '+' e '-', ou seja, há muitos dados da classes + ao redor do novo exemplo 1 e, de forma análoga, muitas amostras da classe '-' ao redor do novo exemplo 3, tornando essa classificação deles bastante fácil, ou direta, mesmo usando apenas 1 vizinho mais próximo. Por outro lado, o novo exemplo 2 está posicionado na fronteira dos valores positivos e negativos, portanto, nesse último caso, a classificação não é confiável.

Como princípio de funcionamento para esse algoritmo, é sempre calculada a distância métrica do novo exemplo, ou nova amostra, em relação a todos os dados do conjunto de treinamento. No caso do novo exemplo 4, ilustrado na figura 2.12, é calculada a distância dele em relação a todos os outros dados utilizados na fase de treinamento. Se o k for igual a 1, então o novo exemplo é rotulado com o rótulo mais próximo a ele, assim, percebe-se que este será rotulado erroneamente.

Agora, escolhendo $K = 3$, como foi adotado na figura, o classificador selecionará os três vizinhos que possuem a menor distância em relação ao novo exemplo e optando por uma votação majoritária será capaz de escolher o seu rótulo **majoritário** entre os 3 vizinhos, e nesse caso a classe positiva será a vencedora. O novo exemplo 4 será então positivo.

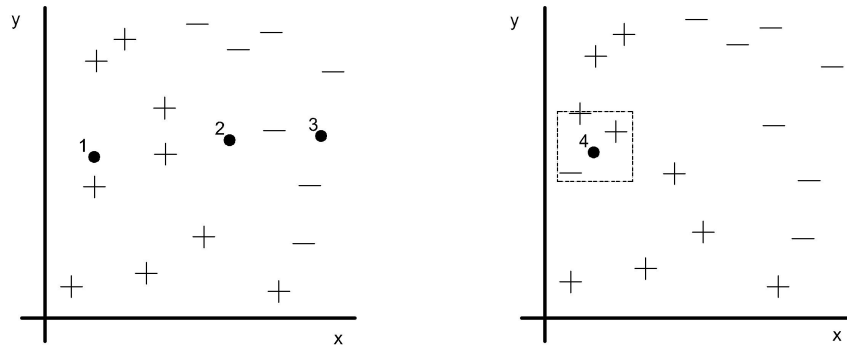


Figura 2.12: Exemplo de classificação utilizando o *KNN*, adaptado de (KUBAT, 2015)

A similaridade entre o novo exemplo e o seu vizinho mais próximo pode ser calculada utilizando alguns tipos de distâncias e a mais comum é a *Euclidiana*. Nesse trabalho foram testadas três tipos de distâncias métricas, definidas pelas equações 2.26, 2.27 e 2.28.

$$\text{Distância Euclidiana } DE = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.26)$$

$$\text{Distância Minkowski (Cúbica) } DCUB = \sqrt[n]{\sum_{i=1}^n (x_i - y_i)^n} \quad (2.27)$$

$$\text{Distância Cosseno } DCOSS = 1 - \frac{\sum_{i=1}^n (x_i y_i)}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (2.28)$$

Onde x representa o novo exemplo e y os dados já rotulados do conjunto de treinamento.

2.8 MEDIDAS PARA A AVALIAÇÃO DE DESEMPENHO DOS CLASSIFICADORES

Nesta seção são apresentados os critérios e parâmetros escolhidos para avaliar o desempenho dos classificadores nos experimentos de validação. Os chamados casos verdadeiros positivos e verdadeiros negativos são os acertos dos classificadores e os falsos positivos e falsos negativos são os erros obtidos durante o processo de classificação. Com o cálculo dessas medidas é possível montar a chamada matriz de confusão.

A tabela 2.3 mostra o modelo da matriz de confusão utilizada nessa seção. Essa matriz é usada para a avaliação de modelos estatísticos e comumente utilizada para a avaliação de algoritmos de classificação usados na área de aprendizagem de máquina. Ela determina e compara valores reais e previstos através dos verdadeiros positivos (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN). A diagonal principal da matriz mostra o número de predições corretas do classificador (VP e VN), enquanto a outra diagonal, a secundária, mostra o número de predições incorretas (FP e FN).

As previsões verdadeiras estão contidas na classe positiva (VP), ou seja, com o motor operando com uma barra rompida na sua gaiola rotórica, para a classe igual a "1", e as previsões negativas (VN), ou seja, com o motor operando sem nenhum defeito na gaiola rotórica, indica a classe "0". A primeira coluna da matriz de confusão apresentada na tabela 2.3 representa a soma dos valores classificados como VP , com aqueles identificados como FN . A outra coluna representa a soma dos dados classificados como FP com aqueles avaliados como VN . A primeira linha mostra a soma dos valores VP e FP , e a segunda linha identifica a soma de valores classificados como FN e VN . A letra N é o total geral de testes realizados.

Tabela 2.3: Modelo da Matriz de confusão usada nesse trabalho

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	VP	FP	$VP + FP$
	Negativo	FN	VN	$FN + VN$
Total		$VP + FN$	$FP + VN$	N

A tabela 2.4 mostra as métricas para a avaliação dos modelos usados nesse trabalho. Cada parâmetro é explicado nos itens abaixo (ZHU; ZENG; WANG, 2010) e (SANTOS, 2017).

1. A **sensibilidade (S)** mostra o quanto o modelo é sensível para reconhecer a proporção de VP e indicar a capacidade do modelo em mostrar casos que realmente existem como verdadeiros positivos. Em outras palavras, a sensibilidade é capaz de mostrar a proporção de acertos de cada classificador ao indicar barras rompidas no rotor da máquina.
2. A **especificidade (E)** mostra a proporção de VN e indica a capacidade do modelo em mostrar casos que realmente existem como verdadeiros negativos, logo, evidencia a proporção de acerto em classificar corretamente os casos nos quais o motor está realmente saudável.
3. Já a **eficiência (Ef)** é a média aritmética entre a especificidade e sensibilidade, e neste caso vale mencionar que esses dois parâmetros variam em direções opostas.
4. A **taxa de acerto** ou *acurácia* (**TA**) é a proporção de acertos ou predições corretas, contudo, não leva em consideração se os acertos foram positivos ou negativos.
5. O **coeficiente de correlação de Matthews (CCM)** é uma métrica de avaliação de qualidade para classificações binárias e indica uma correlação entre as classes binárias. Retorna um valor entre $[-1$ e $1]$. O valor -1 indica uma **classificação errada**, ou seja imprópria, uma discordância geral entre o previsto e o observado. O valor 1 indica uma **classificação perfeita** e o valor 0 uma **classificação aleatória**.

Tabela 2.4: Parâmetros de avaliação dos modelos

Parâmetros		
	Parâmetros	Equação
1	Sensibilidade (S)	$\frac{VP}{(VP+FN)}$
2	Especificidade (E)	$\frac{VN}{(VN+FP)}$
3	Eficiência (Ef)	$\frac{(S+E)}{2}$
4	Taxa de Acerto (TA)	$\left(\frac{VP+VN}{(N)}\right) \cdot 100$
5	Coeficiente de Correlação de Matthews (CCM)	$\frac{(VP \cdot VN) - (FP \cdot FN)}{\sqrt{(VP+FP) \cdot (VP+FN) \cdot (VN+FP) \cdot (VN+FN)}}$

2.8.1 A CURVA ROC

Segundo (PRATI; BATISTA; MONARD, 2008), a curva *ROC* (*Receiver Operating Characteristic*) é usada para selecionar, ou avaliar modelos de classificação, sendo recentemente aplicada na área de aprendizagem de máquina e mineração de dados. Em (PARK;

(GOO; JO, 2004), o autor ressalta que a curva *ROC* é definida pela **sensibilidade** versus o valor da **especificidade-1**. A figura 2.13 representa as curvas ROC obtidas nos extremos para a avaliação dos modelos. As coordenadas (0,1), por exemplo, representam um **classificador ideal** com uma área sobre a curva (*ASC*) igual a 1, ao passo que as coordenadas (0,0) representam um **classificador aleatório**, ou seja, o modelo neste caso classificou de uma forma aleatória e, portanto, não é capaz de discriminar as diferentes classes. Para uma área equivalente a 0,5, no caso as coordenadas (1,0) representam um modelo que sempre faz previsões erradas um **classificador invertido** com área equivalente a 0.

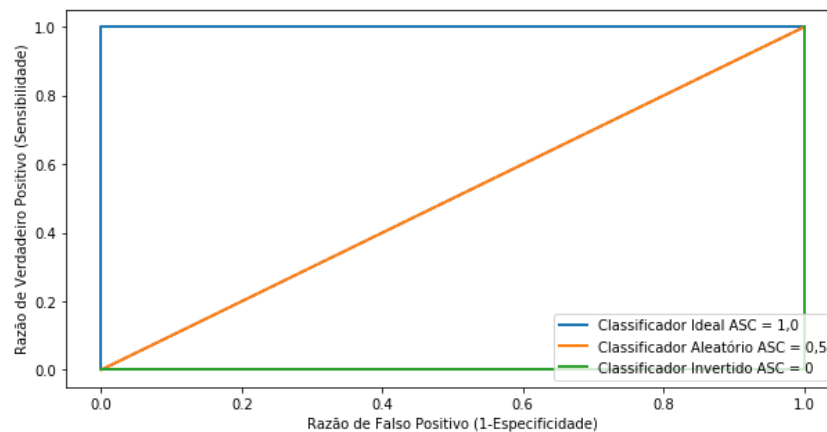


Figura 2.13: Curva ROC padrão para a avaliação dos modelos

A *ASC* é uma métrica importante para analisar o desempenho dos modelos estudados através da curva *ROC*. A tabela 2.5 mostra uma classificação comum utilizada para avaliar a *ASC* de determinados modelos (ZHU; ZENG; WANG, 2010).

Tabela 2.5: Modelo para a avaliação da *ASC* adaptado de (ZHU; ZENG; WANG, 2010)

Range da <i>ASC</i>	Avaliação
$0,9 \leq ASC < 1,0$	Previsão Excelente
$0,8 \leq ASC < 0,9$	Previsão Boa
$0,7 \leq ASC < 0,8$	Previsão Fraca
$0,6 \leq ASC < 0,7$	Previsão Ruim
$0,5 \geq ASC$	Previsão Aleatória

REVISÃO DA LITERATURA

As abordagens existentes para o diagnóstico de barras rompidas nos *MIT* são divididas basicamente em dois grupos: as soluções invasivas e não-invasivas. Em (LIANG; EDMOWANDEKHOE, 2017) é apresentada uma tabela sobre as principais vantagens e desvantagens das duas abordagens (tabela 3.1).

Tabela 3.1: *Vantagens e Desvantagens das abordagens invasivas e não invasivas (LIANG; EDMOWANDEKHOE, 2017)*

Invasiva		Não-invasiva	
Vantagens	Desvantagens	Vantagens	Desvantagens
Possui uma maior acurácia e resultados mais confiáveis	Instalação de sensores é requerida	Não necessita de instalação de sensores no interior da máquina	É uma abordagem indireta
Uma abordagem direta	Difícil Instalação	Fácil Instalação	Requer o uso de um especialista
Facilidade de análise dos resultados	Custo alto	Abordagem mais econômica	Requer uma abordagem mais analítica
	Sofre interferência	Não sofre interferência	

Uma técnica, ou solução invasiva, resulta em introduzir um sensor, ou sonda, no interior do motor para a detecção de assimetrias ou perturbações no campo magnético resultante no entreferro da máquina, quando da ocorrência de barras defeituosas, tal como descrito em (DIAS; CHABU, 2014) e (DIAS; PEREIRA, 2018).

Por sua vez, a abordagem não-invasiva muitas vezes emprega apenas o sinal da corrente de fase do motor para a detecção do defeito, sendo mais simples e mais comum no meio industrial, uma vez que não necessita da parada da máquina para a sua instalação (Di Tommaso; MICELI; GALLUZZO, 2011). Nas próximas seções será apresentada uma revisão da literatura acerca das principais técnicas empregadas para o processamento do sinal de corrente do motor de indução, no sentido de encontrar padrões entre um rotor saudável e uma gaiola danificada, bem como dos classificadores inteligentes usados neste tipo de problema e avaliados no presente estudo. Vale destacar que este trabalho usou uma técnica de coleta de dados não-invasiva, uma vez que apenas a corrente de uma das fases da máquina foi coletada.

3.1 AS TÉCNICAS DE PROCESSAMENTO DE SINAIS USADAS NO DIAGNÓSTICO DE DEFEITOS EM MIT

As soluções encontradas na literatura para o diagnóstico de barras quebradas nos *MIT* muitas vezes necessitam de um pré-processamento dos dados coletados dos sensores, quer seja uma abordagem de natureza invasiva, como uma sonda instalada próxima ao entreferro da máquina, para avaliar perturbações de fluxo, ou mesmo uma solução não-invasiva, como aquela que faz uso de um transformador de corrente acoplado em uma das fases do estator para ler o sinal de corrente.

Neste sentido, muitos estudos abordam as seguintes técnicas de processamento de sinais como mais relevantes para o pré-processamento do sinal da corrente do motor, com a finalidade de diagnosticar o defeito na gaiola rotórica: a *MCSA* (*motor current signature analysis*) que utiliza a *FFT*, a fim de extrair as componentes laterais da frequência fundamental da rede e que são os indicadores de barras rompidas, logo representa uma técnica voltada para o tratamento dos dados no domínio da frequência (FISER et al., 2011), (ISHIKAWA; SHINAGAWA; KURITA, 2013) (LEE et al., 2015).

Outra técnica empregada neste tipo de problema é a *HT*, em conjunto com a *FFT*, de modo a melhorar a resolução em frequência do sinal medido (LAALA; GUEDINI; ZOUZOU, 2011), (PUCHE-PANADERO et al., 2009) e (XU et al., 2013). Neste caso, a banda de frequência de interesse, ou indicador de falha, encontra-se mais distante da frequência fundamental da rede, o que facilita o diagnóstico do defeito especialmente para a condição de baixa carga, ou escorregamento, do motor. Alguns trabalhos adicionais, por exemplo, usaram a transformada discreta *Wavelet* (sigla *DWT* em inglês) para detectar o defeito no rotor particularmente na condição de partida da máquina (BOUZIDA et al., 2011), (Khadim Moin Siddiqui; GIRI, 2012) e (KIA; HENAO; CAPOLINO, 2009).

Adicionalmente, outros pesquisadores procuraram explorar as chamadas técnicas de subespaço ou análise espectral de alta resolução. Uma delas é a classificação de sinais múltiplos (sigla *MUSIC* em inglês), para a detecção de barras quebradas para a condição de partida do motor e também alimentado por tensão não-senoidal (ROMERO-TRONCOSO et al., 2014), (ROMERO-TRONCOSO et al., 2016) e (NAHA et al., 2016). Outros trabalhos usaram a estimação de parâmetros do sinal via técnica de invariância rotacional (sigla *ESPRIT* em inglês), capaz de detectar o defeito em conjunto com a técnica *HT* (TRACHI et al., 2016) e (XU et al., 2013).

No domínio do tempo foram avaliados outros métodos para a extração de características da corrente do motor, entre elas, pode-se destacar a decomposição de modo empírico (sigla *EMD* em inglês), capaz de decompor um sinal complexo original, estacionário ou não-estacionário, em uma série finita de conjuntos de dados chamados funções de modo intrínseco. Tomando por base tais funções, é possível encontrar determinados padrões na corrente do motor para a condição saudável e também para o caso de rotor avariado

(VALLES-NOVO et al., 2014).

É possível ainda utilizar uma análise de parâmetros estatísticos do sinal de corrente, como os apresentados na tabela 2.2, no sentido de encontrar padrões no sinal coletado do motor para a condição de barras quebradas, também no domínio do tempo (??) e (FERNANDEZ-TEMPRANO et al., 2013).

As técnicas que operam com modelos matemáticos capazes de extrair características do sinal no domínio da frequência são usadas para determinar frequências e as suas respectivas amplitudes, para fins de análise e posterior classificação da falha. A técnica *EMD* e a análise estatística do sinal de corrente, por exemplo, oferecem uma série de grandezas no chamado domínio do tempo, de forma a observar determinados padrões entre um sinal saudável e aquele defeituoso para o caso de um rotor danificado.

Muito embora a aplicação direta da técnica *FFT* ofereça algumas desvantagens para a detecção de barras quebradas, ela ainda é uma das mais utilizadas dada a facilidade de obtenção do sinal de corrente do motor, uma vez que apenas uma fase é necessária. De outro lado, a técnica *HT* permite melhorar a resolução em frequência do sinal, antes da aplicação da *FFT*, e também oferece uma implementação computacional bastante interessante, frente a outras técnicas acima relacionadas.

Finalmente, algumas métricas estatísticas, como já evidenciado por alguns autores, também permitem distinguir parâmetros e distorções importantes entre um sinal oriundo de uma máquina defeituosa e outra com um rotor saudável.

Deste modo, o presente estudo explorou as técnicas *FFT* e *HT*, bem como algumas métricas estatísticas, como serão melhor descritas a seguir, de modo a avaliar o melhor conjunto de características capazes de contribuir para o diagnóstico de barras quebradas nos *MIT*, usando diferentes classificadores inteligentes.

3.1.1 O USO DA TÉCNICA *MCSA* NA DETECÇÃO DE BARRAS QUEBRADAS

Como já mencionado, uma das abordagens não-invasivas mais utilizadas até os dias de hoje para a detecção de barras rompidas é a análise da assinatura da corrente do estator *MCSA*, usando a *FFT*, a partir do monitoramento da corrente de uma das fases da máquina com um sensor do tipo transformador de corrente (TC).

Esta técnica avalia a existência de uma frequência harmônica lateral, localizada à esquerda da frequência fundamental da rede (em 60Hz), e cuja amplitude pode indicar não apenas a condição de falha do rotor, mas também a sua possível severidade (EMATSU, 2008). Segundo (BELLINI; FRANCESCHINI; TASSONI, 2006), (SOUSA et al., 2017), (NAHA et al., 2017) e (KARMAKAR et al., 2016), as barras rompidas no rotor provocam uma série de frequências harmônicas, ou componentes laterais calculadas pela equação 3.1 a seguir.

$$f_{BRB} = (1 \pm 2ks)fs \quad (3.1)$$

Onde: f é a frequência fundamental da rede, s é o escorregamento (diferença percentual da velocidade de giro do rotor com a velocidade do campo magnético girante do motor) e k é uma constante relacionada ao número da harmônica.

As componentes harmônicas são induzidas no estator da máquina pelo rotor avariado, e, como descrito, são indicadores de barras quebradas na gaiola rotórica. Segundo (BELLINI; FRANCESCHINI; TASSONI, 2006), (BELLINI et al., 2002), (EMATSU, 2008) e (ANTONINO-DAVIU et al., 2006), a componente da esquerda, ou seja aquela obtida pela parcela $(1 - 2ks)fs$, é produzida pela assimetria causada pelo rotor danificado. Por outro lado, a componente da direita, ou seja, a componente $(1 + 2ks)fs$, tem a sua causa mais relacionada a oscilações de velocidade e torque na máquina. Vale ressaltar que, segundo (EMATSU, 2008), essas componentes muitas vezes já existem na máquina e são causadas por imperfeições na sua fabricação e outras características construtivas do motor, porém, no caso de barras rompidas, essas imperfeições causam um acréscimo na amplitude das harmônicas laterais.

O método *MCSA* se tornou uma das principais e mais usadas técnicas para o diagnóstico de assimetrias (CULBERT; RHODES, 2007) nos *MIT*, e segundo (KARMAKAR et al., 2016), tal técnica permite realizar o diagnóstico não apenas de barras quebradas, mas também identificar outras assimetrias no rotor, como por exemplo alterações relacionadas aos rolamentos e eventual desbalanceamento do eixo. De outro modo, a técnica *MCSA* exige algumas condições importantes para o seu uso, a fim de garantir a eficiência desejada no diagnóstico de barras quebradas. Segundo (EMATSU, 2008), essas premissas podem ser resumidas da seguinte maneira:

- A rotação do eixo da máquina precisa ser constante e conhecida durante a amostragem do sinal de corrente;
- Não pode haver variações na frequência fundamental do estator;
- A carga deve ser constante; e
- A carga aplicada ao eixo deve ter amplitude suficiente para que as bandas laterais estejam suficientemente distantes da frequência fundamental.

Face ao acima exposto, para que seja possível distinguir as componentes laterais à frequência fundamental, especialmente aquele banda lateral esquerda responsável pela detecção da falha no rotor, é necessária uma carga, ou um escorregamento, capaz de

distanciar aquelas frequências da frequência da rede. Como visto na equação 3.1, as componentes laterais estão diretamente associadas ao escorregamento do eixo.

Nota-se, portanto, que a aplicação da *MCSA* requer uma medição bastante precisa da rotação do eixo, a fim de se conhecer o seu escorregamento real. Para os motores de grande porte, por exemplo, com dezenas ou centenas de quilowatts, a detecção do defeito é ainda mais onerosa, pois tais máquinas operam normalmente em baixo escorregamento, mesmo para a condição de carga nominal (LAALA; GUEDINI; ZOUZOU, 2011), (XU et al., 2013), (THOMSON; FENGER, 2003) e (GYFTAKIS et al., 2015). Logo, nos motores maiores torna-se ainda mais difícil localizar as aludidas bandas de frequência laterais.

3.2 TÉCNICAS DE APRENDIZAGEM DE MÁQUINA PARA A DETECÇÃO DE BARRAS ROMPIDAS NO *MIT*

As técnicas de processamento digital de sinais são comumente empregadas no diagnóstico de falhas em *MIT*, em especial, para extrair características do sinal coletado da máquina, como corrente, fluxo magnético, vibrações, entre outros. Tais informações podem ser agrupadas no sentido de construir uma base de conhecimento sobre as condições operacionais do motor, como a condição de um rotor saudável e outro avariado.

Segundo (LIANG; EDMOWANDEKHOE, 2017), atualmente existem três categorias fundamentais aplicadas na detecção de falhas em motores de indução:

1. Aquela fundamentada na extração da assinatura da corrente, ou tensão, do motor;
2. Uma segunda baseada em um modelo matemático do comportamento do motor; e
3. Uma terceira apoiada na construção de uma base de conhecimento das condições operacionais da máquina.

O avanço da computação permitiu o desenvolvimento de modelos inteligentes de apoio à decisão, robustos e eficientes, particularmente orientados a partir de uma base de conhecimento de informações coletadas do motor. Deste modo, surgiram técnicas baseadas em *machine learning* (*ML*) aplicadas no monitoramento e diagnóstico de falhas nos *MIT*.

Alguns autores utilizaram algumas técnicas baseadas em *IC* para a detecção de diversos tipos de falhas. Por exemplo, em (LAALA; GUEDINI; ZOUZOU, 2011) e (Moraes de Sousa, 2017), os autores utilizaram um modelo de inferência *fuzzy* para detectar barras quebradas no rotor operando em baixo escorregamento.

Em (DIAS; PEREIRA, 2018), os autores compararam o uso de uma rede neural do tipo perceptron multicamadas (NN) com a técnica dos vizinhos mais próximos (K - nearest neighbor, sigla *KNN* em inglês) e a técnica conhecida como máquina de suporte vetorial (Support Vector Machine, sigla *SVM* em inglês), para classificar as características extraídas a partir de um sensor de efeito hall instalado no interior do motor, usando a *FFT*

como ferramenta de pré-processamento. Ademais, os estudos de (DIAS; PEREIRA, 2018) e (FERNANDEZ-TEMPRANO et al., 2013) avaliaram algumas características estatísticas do sinal da sonda, também usados como entradas para os classificadores.

Em (MARTIN-DIAZ et al., 2018), os autores usaram diferentes técnicas inteligentes, como por exemplo, o *SVM*, a rede neural e o *K-NN* para a identificação de falhas em motores de indução sob diferentes condições de operação. Os autores em (WIDODO; YANG; HAN, 2007) descreveram o uso da *SVM* para o diagnóstico de falhas em motores de indução e usaram a chamada análise de componentes independentes para a extração das características e redução das mesmas.

Neste sentido, a tabela 3.2 resume alguns dos trabalhos mais recentes no campo do diagnóstico de barras quebradas nos *MIT*, e que empregam não apenas as técnicas de processamento de sinais escolhidas neste trabalho, mas também as abordagens inteligentes como suporte à decisão para avaliar as condições do rotor e diferentes janelas amostrais, como também é objetivo deste estudo.

Tabela 3.2: *Sumário de trabalhos recentes que empregam técnicas de processamento de sinais e técnicas de ML para o diagnóstico de barras quebradas*

Referências	Abordagem	Inversor	Extração de características	Técnica Computacional	Janela amostral	Carga baixa
(LAALA; GUEDINI; ZOUZOU, 2011)	Não invasiva	-	$HT + FFT$	Lógica Fuzzy	-	Sim
(Morais de Sousa, 2017)	Invasiva	Não	FFT	Lógica Fuzzy	4s	Sim
(FERNANDEZ-TEMPRANO et al., 2013)	Não invasiva	Sim	características estatísticas e FFT	Linear Discriminant Analysis	10s	Sim
(DIAS; PEREIRA, 2018)	Invasiva	Não	características estatísticas e FFT	Rede Neural Perceptron Multicamadas, KNN e SVM	4s	Sim
(WIDODO; YANG; HAN, 2007)	Não invasiva	Não	características estatísticas e frequência	SVM	3.27s	-
(MARTIN-DIAZ et al., 2018)	Não invasiva	Sim	características estatísticas e frequência	Naive Bayes, KNN, Bagging, AdaBoost, SVM e Rede Neural	10s	Não
(PUCHE-PANADERO et al., 2009)	Não invasiva	Não	$HT + FFT, FFT$	-	100s	Sim
(XU et al., 2013)	Não invasiva	Não	$FFT, HT + ESPRIT$ e $HT + MUSIC$	-	10s, 4s e 100s	Sim
(ROMERO-TRONCOSO et al., 2014)	Não invasiva	Sim	$MUSIC$ e $STFT$	-	10s e 3s	Sim
Este Trabalho	Não invasiva	Sim	características estatísticas, FFT , $HT + FFT$	SVM, KNN e RL	10s, 5s, 3s, 1s, 0, 5s e 0, 25s	Sim

METODOLOGIA PROPOSTA

4.1 CARACTERIZAÇÃO METODOLÓGICA

A metodologia de pesquisa proposta neste trabalho foi estruturada como um estudo exploratório na literatura, com a finalidade de avaliar e conhecer em mais detalhes as técnicas de processamento de sinais e os principais classificadores empregados na detecção de barras totalmente quebradas na junção com o anel de curto-circuito de um motor de indução. O trabalho pode ainda ser caracterizado como uma pesquisa aplicada, dada a sua linha de investigação experimental e também de natureza quantitativa, pois há uma relação direta entre os dados experimentais coletados e os resultados numéricos advindos de cada modelo implementado e testado no curso da pesquisa.

4.2 PROCEDIMENTOS PARA A CONDUÇÃO DOS EXPERIMENTOS

A seguir são apresentadas as principais etapas usadas na abordagem proposta neste trabalho, para o diagnóstico de barras rompidas em um motor de indução trifásico. A figura 4.1 mostra todas as etapas usadas na metodologia.

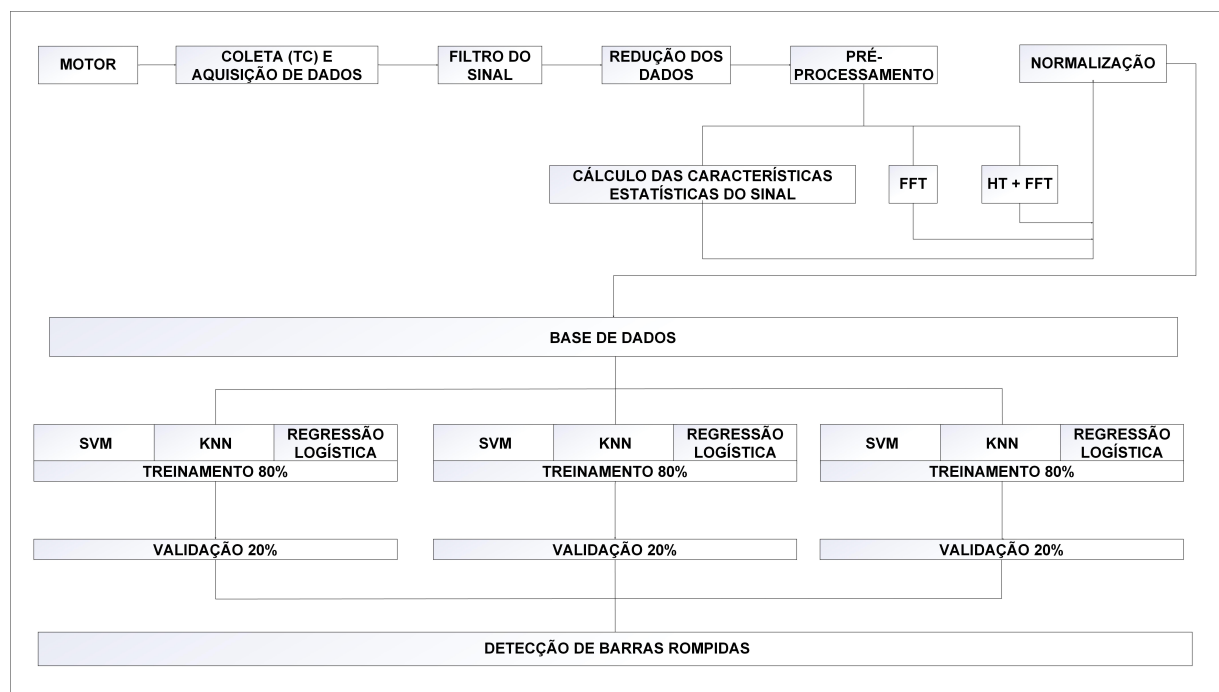


Figura 4.1: Metodologia proposta no trabalho

A primeira etapa consistiu em coletar o sinal de corrente elétrica oriundo do estator da máquina, com o rotor saudável e também para o caso do rotor com uma barra rompida. Nessa etapa um conjunto de $N = 100000$ amostras do sinal da corrente do estator foi

coletado com uma frequência amostral de $10kHz$ e uma resolução em frequência de $0,1Hz$ (janela amostral de 10s). Um transformador de corrente (TC) foi acoplado em uma das fases da máquina para a coleta dos dados. Esses dados foram agrupados e usados para extrair características necessárias, nos domínios do tempo e da frequência, para posterior avaliação da condição operacional do motor.

Em uma segunda etapa, utilizou-se um filtro digital do tipo *Butterworth* 6ª ordem para a redução de eventuais harmônicas, que poderiam prejudicar a extração das frequências de interesse na avaliação do defeito, uma vez que o motor foi alimentado por um inversor e, neste caso, há uma maior predominância de frequências múltiplas a fundamental. O filtro digital foi implementado na linguagem de programação *Python*.

Uma terceira etapa consistiu em reduzir os dados originais em janelas amostrais inferiores, tais como 0.25s, 0.5s, 1s, 3s e 5s utilizando o *MATLAB*, a fim de avaliar a eficiência da abordagem proposta mesmo para um conjunto inferior de informações do motor. Logo, a finalidade neste caso foi avaliar se as características extraídas nos domínios do tempo e da frequência, para um vetor de dados menor, seriam suficientes para o classificador avaliar com acuidade as condições estruturais do rotor. Cabe mencionar que uma janela amostral inferior proporciona um menor custo computacional para o pré-processamento do sinal. Uma contribuição interessante desse trabalho foi avaliar a eficiência das técnicas combinadas para uma janela amostral bastante inferior aquela empregada no tradicional método *MCSA*, uma vez que este último necessita de 100s para a extração das bandas de frequência laterais.

A etapa seguinte consistiu em realizar as três principais etapas do pré-processamento do sinal de corrente do estator, para posterior comparação, quais sejam: extração das características estatísticas, aplicação da Transformada Rápida de *Fourier* (*FFT*) e uso da Transformada de *Hilbert* e Transformada de *Fourier* (*HT* e *FFT*) combinadas.

Todos os dados foram normalizados, a fim de garantir os mesmos limites mínimos e máximos como entrada para os classificadores escolhidos. Após a normalização, foi montada uma base de dados dividida em dois conjuntos, um primeiro para treinamento, correspondendo a 80% de toda a base de dados montada e outros 20% responsáveis pela validação de cada modelo. Os algoritmos classificam as características e entregam uma saída binária, com valor igual a 1 representando o rotor com uma barra rompida e valor igual a 0 para o rotor na condição saudável. Vale salientar que, no processo de treinamento foi utilizada uma validação cruzada com 10 *K-folds*.

Os dados de entrada foram pré-processados para a extração do espectro de frequências e da componente indicadora de barras rompidas, como mostra a tabela 4.1, nas quais são extraídas a frequência, a amplitude e o escorregamento da componente $(1 - 2s)f$, bem como a componente $2sf$, utilizando a *FFT* e a combinação das técnicas *HT* e *FFT* respectivamente. Como mencionado, algumas características estatísticas do sinal de corrente também foram calculados no domínio do tempo.

Cabe ressaltar que, a amplitude das componentes extraídas apenas com o uso da *FFT* e/ou com as técnicas *HT* e *FFT* são diretamente proporcionais a severidade do defeito.

Há alguns trabalhos, como aqueles publicados por (DIAS; PEREIRA, 2018), (GARDEL et al., 2012) e (FERNANDEZ-TEMPRANO et al., 2013), que mostram uma performance melhor dos classificadores com o uso combinado de características estatísticas com as características de frequência, na predição das respostas para o problema de barras rotóricas rompidas.

Tabela 4.1: *CARACTERÍSTICAS EXTRAÍDAS EM CADA DOMÍNIO*

Etapa de pré-processamento	Número de características extraídas do sinal	Características extraídas e usadas como entradas para os modelos inteligentes
FFT	3	Frequência da banda lateral esquerda (1-2s) f , amplitude da frequência lateral esquerda e escorregamento (s)
HT+FFT	3	Frequência da banda 2sf, amplitude da respectiva frequência e escorregamento (s)
Dados estatísticos	17	Todas as características descritas na tabela 1

A base de dados criada, a partir da extração das características utilizando as três abordagens citadas, foi montada para diferentes tensões aplicadas no freio eletromagnético do motor, que são: 40V, 50V, 60V e 70V, uma vez que tais valores correspondem a um baixo escorregamento do eixo do motor, entre 0,66% a 1,33%. Deste modo, foram extraídas seis características no domínio da frequência e dezessete grandezas estatísticas conforme descrito na tabela 4.1. Foram coletadas 100000 amostras para cada condição de tensão no disco do freio, a fim de extrair cada característica do sinal da corrente.

No tocante aos ensaios, foram realizados 70 experimentos na condição saudável e outros 70 casos para o rotor danificado. Para cada experimento foram aplicadas as quatro condições de tensão ao freio do motor, logo, tais ensaios totalizam 280 linhas de dados para a classe 0 (saudável) e outras 280 linhas para a classe 1 (uma barra rompida). O banco de dados, por fim, resulta em uma matriz de 560 por 23 características.

4.3 MONTAGEM DO APARATO EXPERIMENTAL

Para a avaliação da metodologia proposta neste trabalho, foi montado um aparato experimental em laboratório, formado por um motor de indução de 7,5kW, 220V, com 4 polos magnéticos, sendo este dotado de um freio eletromagnético de Foucault responsável por aplicar diferentes cargas no eixo da máquina e um ventilador usado para o resfriamento do freio. As especificações deste conjunto do motor são descritas na tabela 4.2.

Tabela 4.2: *ESPECIFICAÇÕES DO CONJUNTO MOTOR*

Motor Experimental	Motor	Freio	Ventilador Freio
Fabricante	Equacional	Equacional	WEG
Modelo	EA2.180SE	EFC-180	W22
Tensão/Velocidade (V/RPM)	220/1740	240/1740	220/3420
Potência (W)	7500	240	3000
Corrente Nominal (A)	25,2	1	-
Quantidade de Barras	38	-	-
Número de Polos	4	-	-
Número de ranhuras do estator	48	-	-

Cabe destacar que o motor usado nesse trabalho foi especialmente projetado para o estudo de barras quebradas, sendo este adquirido por meio de fomento de pesquisa regular da *FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo)* – processo nº 2016/02525-1. Para o acionamento da máquina foi usado um inversor de frequência da marca *WEG*, pois um dos objetivos do estudo é avaliar o defeito de barras quebradas com o motor alimentado com tensão não-senoidal.

Para a coleta e processamento dos dados foi usado um PC Intel core i5, com frequência de clock em 1,8G Hz e 6 GB de memória *RAM*, além de um osciloscópio digital da marca *Hantek* (modelo *HT6022BE*), com 20 MHz de largura de banda e uma taxa de frequência máxima de 48 MS/s.

O software *MATLAB* foi usado para o processamento de sinais e aplicação dos algoritmos baseados em *ML*, o software *ORIGIN* e a linguagem de programação *Python* foram usados para a geração de alguns gráficos e o *AUTOCAD* utilizado para desenhar algumas figuras desse trabalho. O aparato completo empregado nos ensaios pode ser visto na figura 4.2.

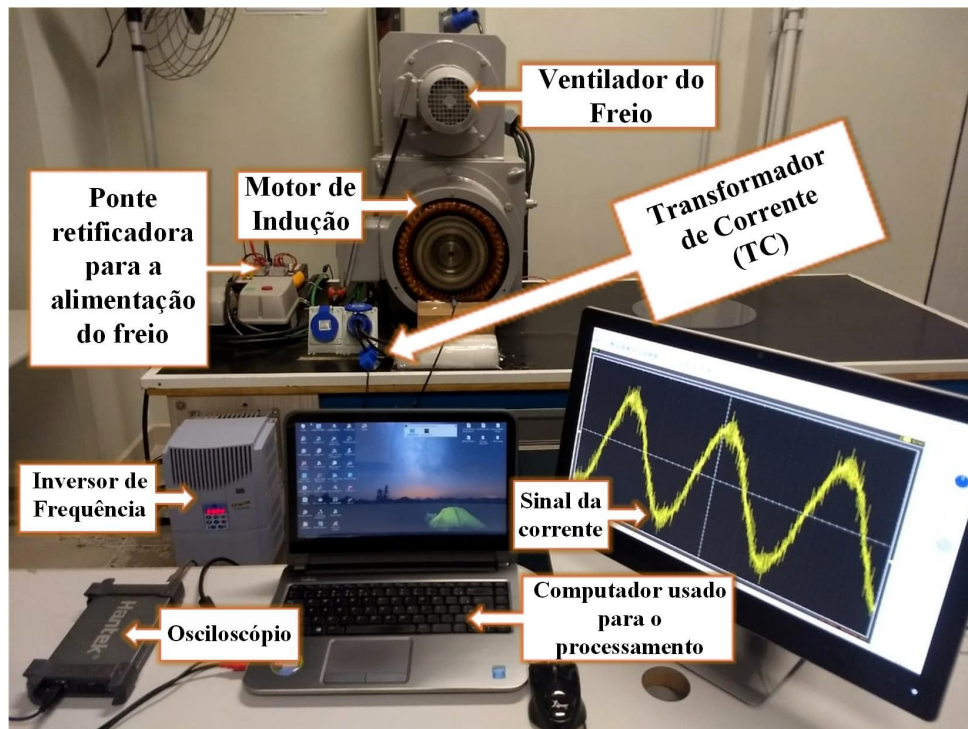


Figura 4.2: *Aparato experimental montado para a avaliação de falhas no rotor*

4.4 DEFINIÇÃO DOS CENÁRIOS OPERACIONAIS PARA O MOTOR

Como mencionado anteriormente, o motor objeto dos ensaios experimentais, foi especialmente fabricado para investigar barras quebradas com o anel de curto do rotor. A sua construção permite soltar, de um dos lados do rotor, algumas porcas associadas aos parafusos em cada barra e assim simular a desconexão destas com o anel, como mostra a figura 4.3.



Figura 4.3: *Simulação das barras rompidas no rotor para o MIT obtido com o fomento da FAPESP*

Ao desconectar duas porcas, por exemplo, uma na posição inferior, isto é, no anel, e

outra no lado superior do rotor, ou seja, na extremidade da barra, é possível simular o rompimento de uma barra. Deste modo, os experimentos foram realizados em laboratório e os resultados foram obtidos para duas condições da máquina, na situação saudável e outra simulando uma barra quebrada.

Para fins de comparação de desempenho da combinação das técnicas de pré-processamento e dos classificadores, a janela amostral foi testada em diferentes condições. Os cenários estudados estão descritos na tabela 4.3. Foram realizados experimentos em cinco cenários diferentes, sendo possível comparar o uso das características de frequência, extraídas com as técnicas *FFT* e com a combinação da *HT* e *FFT*, bem como aquelas calculadas apenas com parâmetros estatísticos. Adicionalmente, foram combinadas algumas características estatísticas com outras no domínio da frequência.

Para os primeiro e segundo cenários foram utilizadas vinte características como entrada para os três classificadores. Para os terceiro e quarto cenários foram utilizadas três características em cada classificador. E por fim, para o quinto cenário foram utilizadas 17 características.

Neste estudo foram avaliadas as seguintes janelas amostrais: 10s, 5s, 3s, 1s, 0,5s e 0,25s. Cabe frisar que, tais janelas foram escolhidas em função da necessidade em se avaliar o defeito no rotor em tempo real, uma vez que a detecção de defeito em tempos bastante reduzidos permite diagnosticar a condição do rotor não apenas em carga variável, mas também em situações nas quais o motor opera com velocidade variável. Neste último caso, por exemplo, torna-se necessário identificar a condição do rotor em poucos segundos e não apenas quando a máquina alcançou o regime permanente.

Para as janelas amostrais em 10s e 5s foi possível identificar as características de frequência utilizando a *FFT*. No caso das janelas em 10s, 5s e 3s foram extraídas as características utilizando as técnicas *HT* e *FFT*. Para as características estatísticas, foi possível calcular as grandezas em todas as janelas amostrais.

Desse modo, não foi possível extrair a componente $2sf$ nas janelas amostrais entre 0,25s e 1s, devido a sua baixa resolução em frequência. De forma análoga, não foi possível extrair a componente $(1 - 2s)f$ em janelas compreendidas no intervalo de 0,25s e 3s. Todas as características foram extraídas utilizando o software *MATLAB* e tais grandezas serviram como entradas para os classificadores testados.

Tabela 4.3: *CENÁRIOS USADOS PARA COMPARAÇÃO DE DESEMPENHO DE CADA SOLUÇÃO INTELIGENTE*

Cenários	Descrição
1	Características de frequência extraídas com <i>FFT</i> para 10s e 5s
2	Parâmetros estatísticos combinados com características de frequência, extraídas com <i>FFT</i> , para janelas de tempo em 10s e 5s
3	Características de frequência extraídas com <i>HT + FFT</i> para 10s, 5s, 3s
4	Parâmetros estatísticos combinados com características de frequência, extraídas com <i>HT + FFT</i> , para janelas de tempo em 10s, 5s, 3s
5	Parâmetros estatísticos para 10s, 5s, 3s, 1s, 0,5s e 0,25s

RESULTADOS

Neste capítulo serão apresentadas alguns exemplos de limitações de cada abordagem de pré-processamento (*FFT* e *HT* com *FFT*), a fim de demonstrar as dificuldades em extrair as componentes harmônicas responsáveis por identificar o defeito na gaiola, especialmente na condição de baixo escorregamento.

A seção 5.1 apresenta algumas condições de acionamento do motor e os limites nos quais ainda é possível, ou não, identificar as frequências de interesse.

Essa capacidade é avaliada, variando o escorregamento (baixo escorregamento) e a janela amostral e com isso, consegue-se saber até que ponto as técnicas são eficazes afim de mostrar a componente indicadora do problema.

As seções seguintes mostram os resultados advindos dos experimentos conduzidos no presente estudo. Vale lembrar que a base de dados foi construída com 560 linhas divididas em 80% e 20%, sendo treinamento e validação respectivamente. Assim, o conjunto de validação representou um total de 112 linhas, com 56 linhas para a classe 0 (motor saudável) e outras 56 linhas para a classe 1 (uma barra rompida).

5.1 LIMITAÇÕES DAS ABORDAGENS UTILIZADAS NO PRÉ-PROCESSAMENTO DO SINAL

5.1.1 EXEMPLO DO ESPECTRO DE FREQUÊNCIAS EXTRAÍDO COM A *MCSA* PARA DIFERENTES JANELAS AMOSTRAIS

As figuras 5.1A e 5.1B mostram o espectro de frequência do sinal da corrente do estator extraído com a *FFT* para uma janela amostral de 10s. Neste caso são apresentadas duas condições de carga para um motor de indução trifásico avariado, ou seja, com uma barra rotórica rompida, a primeira para um baixo escorregamento, em torno de 0,66% e a segunda para o motor operando com um escorregamento igual a 1,33%. O motor neste exemplo tem um escorregamento nominal igual a 3,33%.

Na figura 5.1A é possível observar, de forma bastante suave, a componente de frequência $((1-2s)f_s)$ localizada à esquerda da frequência fundamental (60 Hz), o que caracteriza, para um escorregamento baixo, a dificuldade em se detectar a frequência responsável por indicar a falha no rotor, usando apenas a técnica *FFT*. Quando a banda está bastante próxima à frequência da rede, e com baixa resolução, ocorre a chamada dispersão espectral, o que equivale a dizer que a referida harmônica está sendo suprimida pelas componentes mais próximas a 60 Hz (XU et al., 2013). Tal fenômeno dificulta ainda mais a identificação da harmônica e, portanto, do defeito no rotor.

Por sua vez, a figura 5.1B mostra a componente harmônica lateral esquerda de forma

mais evidente, quando comparada ao caso anterior, pois nesta condição o motor foi alimentado com uma carga maior, logo com um escorregamento capaz de distanciar a frequência de interesse da componente fundamental da rede.

Vale ressaltar que, a técnica *MCSA* necessita de um volume de dados importante para visualizar as bandas laterais do defeito, algo em torno de 1 milhão de amostras, a fim de garantir a resolução em frequência necessária para distinguir as harmônicas de interesse da frequência fundamental da rede. Assim, este método normalmente exige uma janela amostral em torno de 100s para uma frequência de 10 kHz. Neste trabalho, foi avaliado o uso da *MCSA* em janelas menores, iguais a 10s e 5s, como detalhado no Capítulo 4.

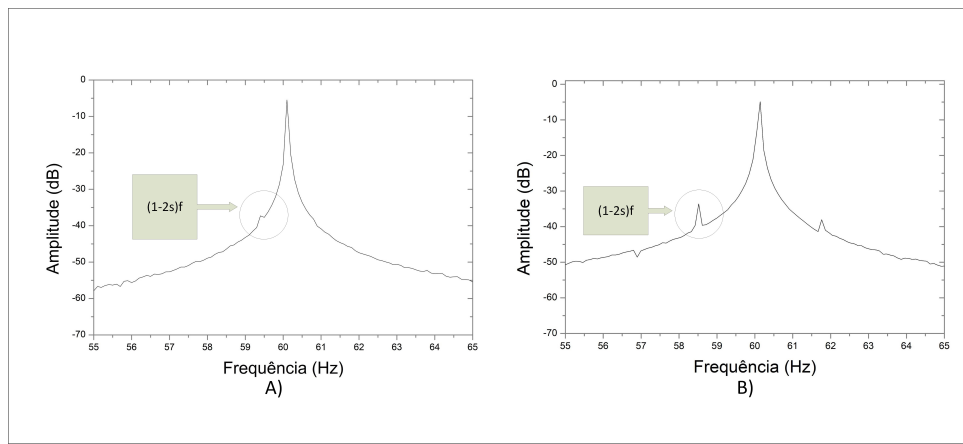


Figura 5.1: A) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 10s) e B) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 10s)

As figuras 5.2 A e 5.2 B mostram o espectro de frequência extraído com a *FFT* para uma janela amostral de 5s, nas mesmas condições de carga apresentadas nas figuras 5.1A e 5.1B. Nota-se, para a primeira figura, que não foi possível identificar a componente lateral esquerda, mas apenas quando a carga foi incrementada no eixo do motor (figura 5.2 B). Trata-se de uma avaliação com resolução em frequência ainda menor do que aquela processada para uma janela igual a 10s, além do fato da banda lateral estar bastante próxima da frequência fundamental.

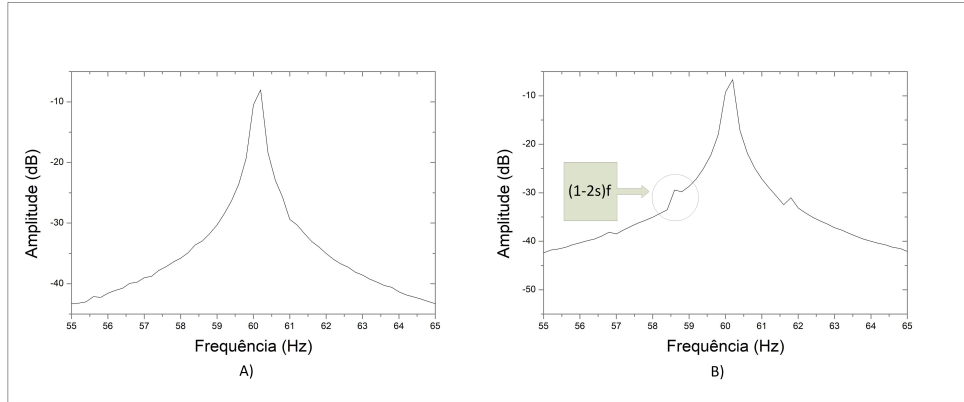


Figura 5.2: A) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 5s) e B) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 5s)

Por fim, as figuras 5.3 A e 5.3 B mostram o espectro de frequência extraído com a *FFT* para uma janela amostral ainda menor, igual a 3s, nas mesmas condições de carga dos dois casos anteriores. Neste último, mesmo para uma carga maior, não foi possível identificar de forma precisa as bandas laterais quando da ocorrência do defeito. Logo, fica evidente que a janela amostral interfere de forma bastante significativa na extração das frequências de interesse para a detecção de barras rompidas nos motores de indução, usando a conhecida técnica *MCSA*. Então, o limite dessa técnica encontra-se na janela de 3s.

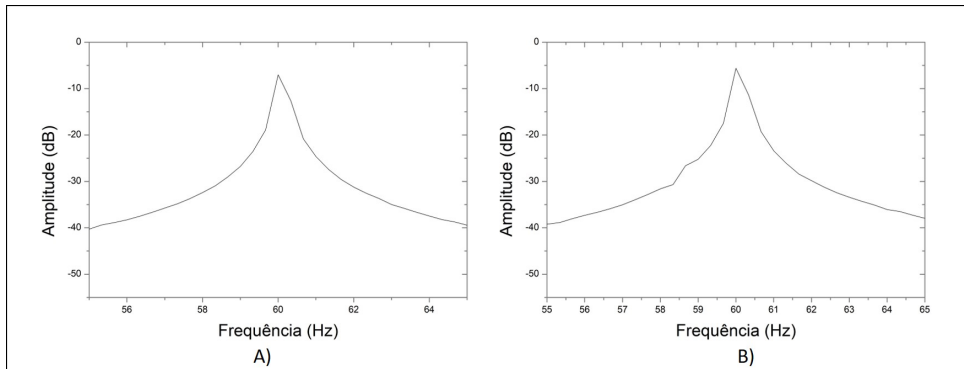


Figura 5.3: A) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 3s) e B) Banda lateral esquerda para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 3s)

5.1.2 A TRANSFORMADA DE HILBERT NO APOIO A DETECÇÃO DE BARRAS QUEBRADAS

Como descrito na seção anterior, existem dificuldades no uso da *MCSA* como a necessidade de uma resolução em frequência de $0,01\text{Hz}$, para uma janela amostral de $100s$ (1 milhão de amostras em 10 kHz), a fim de alcançar as condições suficientes para distinguir o que é a componente lateral e o que é a frequência fundamental em caso de quebra de barras, especialmente em baixo escorregamento (TRACHI et al., 2016).

Neste sentido, o uso combinado das técnicas *HT* e *FFT* permitiu obter um melhor desempenho na avaliação das condições do rotor, quando comparada a aplicação isolada da *MCSA*, em condições de baixa carga (PUCHE-PANADERO et al., 2009).

A grandeza a ser medida continua sendo a corrente de fase do estator, mas neste caso as duas técnicas (*HT* + *FFT*) são empregadas para extrair uma outra componente de frequência, localizada na posição $2sf$. Em (LAALA; GUEDINI; ZOUZOU, 2011), os autores aplicaram as técnicas *HT* e *FFT* para extrair a componente $2sf$ da corrente do estator, e um modelo de inferência *fuzzy* para detectar barras quebradas no rotor para o motor operando em baixo escorregamento. Essa técnica realiza um espalhamento do sinal, de modo a deixar a componente lateral mais distante da fundamental evitando assim a dispersão espectral.

5.1.3 EXEMPLO DO ESPECTRO DE FREQUÊNCIA EXTRAÍDO COM AS TÉCNICAS *HT* E *FFT* PARA DIFERENTES JANELAS AMOSTRAIS

As figuras 5.4 A e 5.4 B mostram um espectro de frequências do sinal de corrente do estator, para um motor de indução com uma barra rompida, extraído com as técnicas *HT* e *FFT*, em uma janela amostral de $10s$.

A figura 5.4 A mostra o espectro de frequências para o motor operando com um escorregamento da ordem de $0,66\%$ e a banda lateral esquerda $2sf$ responsável por indicar a falha no rotor. Fica evidente que essa frequência está bastante distante da frequência fundamental em 60 Hz .

De forma similar, a figura 5.4 B mostra o espectro extraído para um escorregamento da ordem de $1,33\%$ e, neste caso, a banda $2sf$ fica localizada mais distante da frequência zero, o que facilita a sua identificação, além do fato de permanecer ainda bastante distante da componente em 60 Hz .

Como ilustrado nas duas figuras, a componente $2sf$ é um sinal em baixa frequência e o que caracteriza o rompimento de barras neste caso é o acréscimo da amplitude desta harmônica. Assim, nota-se que esse sinal possui a maior amplitude entre as frequências de 0 e 10Hz (PUCHE-PANADERO et al., 2009).

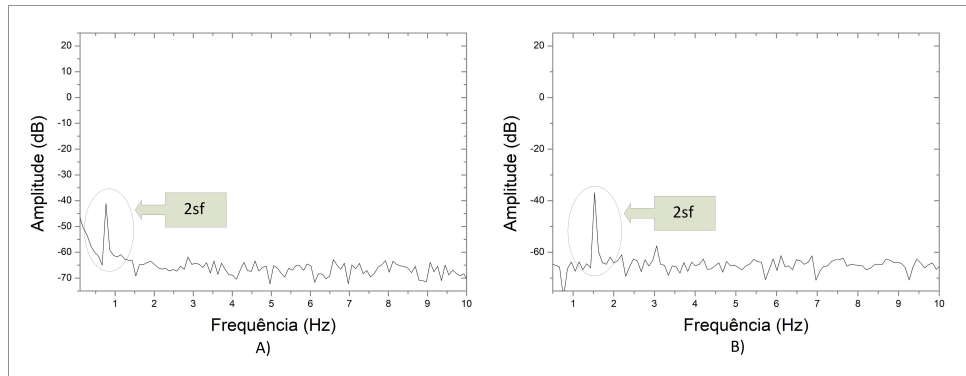


Figura 5.4: A) Banda lateral esquerda ($2sf$), usando as técnicas *HT* e *FFT*, para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 10s) e B) Banda lateral esquerda ($2sf$), usando as técnicas *HT* e *FFT*, para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 10s)

De forma semelhante, esse sinal foi avaliado para as janelas amostrais em 5s e 3s. As figuras 5.5 A e 5.5 B apresentam o espectro de frequências do sinal da corrente extraído com as técnicas *HT* e *FFT*, para uma janela de 5s, e neste caso a resolução em frequência foi reduzida. Diferentemente do que ocorreu para o uso isolado da técnica *FFT*, na mesma janela amostral, aqui foi possível extrair as características da componente $2sf$ em ambos os escorregamentos.

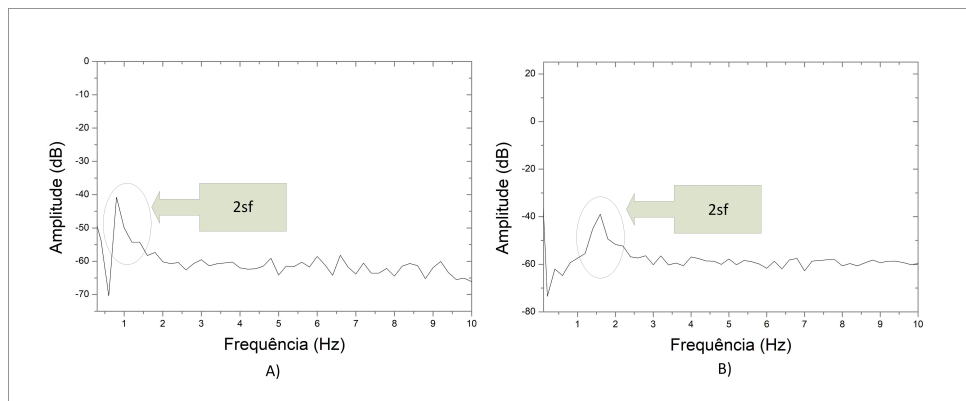


Figura 5.5: A) Banda lateral esquerda ($2sf$), usando as técnicas *HT* e *FFT*, para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 5s) e B) Banda lateral esquerda ($2sf$), usando as técnicas *HT* e *FFT*, para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 5s)

As figuras 5.6 A e 5.6 B mostram o espectro de frequências extraído com as duas técnicas em uma janela de 3s. As duas figuras mostram que, mesmo reduzindo ainda mais a resolução em frequência, é possível identificar a banda lateral responsável pelo defeito no rotor, a componente $2sf$.

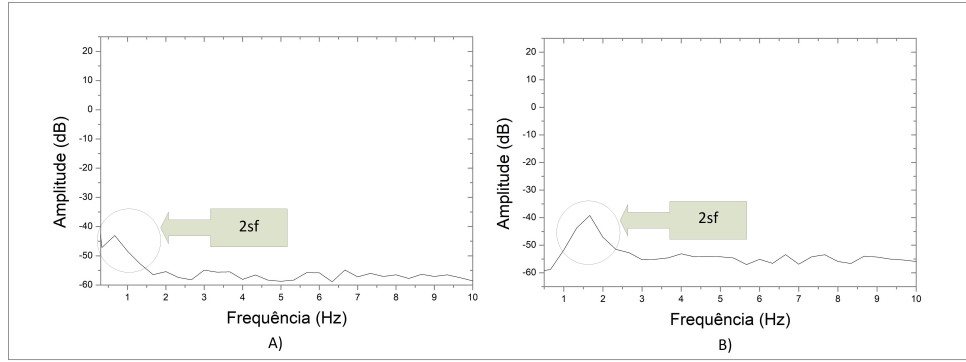


Figura 5.6: A) Banda lateral esquerda (2sf), usando as técnicas HT e FFT, para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 3s) e B) Banda lateral esquerda (2sf), usando as técnicas HT e FFT, para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 3s)

Finalmente, as figuras 5.7 A e 5.7 B mostram o espectro de frequências extraído para uma janela igual a 1s. Neste caso, fica evidente a dificuldade em encontrar o sinal em baixa frequência, então, o limite dessa técnica encontra-se na janela de 1s. Logo, neste trabalho as duas técnicas combinadas foram empregadas nas janelas iguais a 10s, 5s e 3s, como detalhado no Capítulo 4.

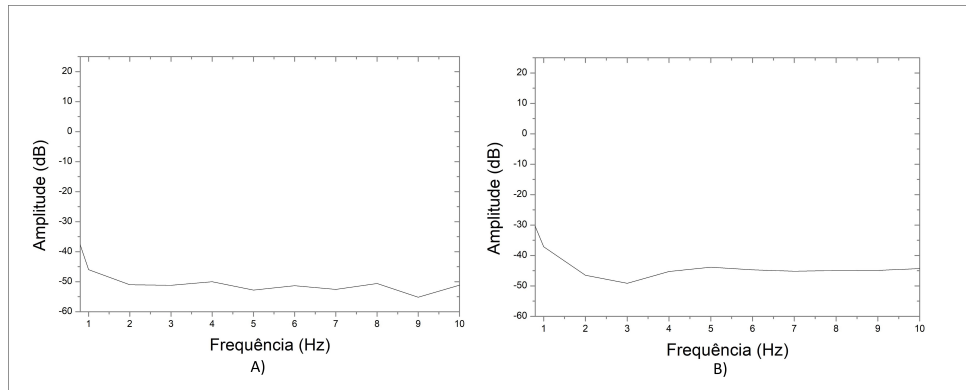


Figura 5.7: A) Banda lateral esquerda (2sf), usando as técnicas HT e FFT, para um MIT avariado e operando com um escorregamento igual a 0,66% (janela de 1s) e B) Banda lateral esquerda (2sf), usando as técnicas HT e FFT, para um MIT avariado e operando com um escorregamento igual a 1,33% (janela de 1s)

5.2 AVALIAÇÃO DOS MELHORES PARÂMETROS PARA OS CLASSIFICADORES

Para o classificador *KNN* foram utilizados 1, 3, 5, 7, 9 e 10 vizinhos mais próximos, baseados no estudo de (MARTIN-DIAZ et al., 2018). As distâncias métricas foram escolhidas de acordo com as melhores taxas de acerto obtidas na fase de treinamento. A

tabela 5.1 mostra os cenários avaliados neste trabalho com as suas respectivas distâncias métricas. Para o classificador *SVM* foi utilizada a função *kernel* gaussiana na maioria dos cenários e a função linear apenas para os casos nas quais as técnicas *HT*, *FFT* e *EST* foram usadas em uma janela de 3s. O parâmetro *C* de regularização do modelo *SVM* variou-se no intervalo entre 1 a 440 e o parâmetro γ (**gama**) foi testado no intervalo de 0 a 4.6, também tomando por base os modelos propostos por (MARTIN-DIAZ et al., 2018).

Tabela 5.1: Distâncias utilizadas de acordo com cada cenário (*DE* distância euclidiana, *DCUB* distância cúbica e *DCOSS* distância cosseno)

	Cenários	Distância Métrica Utilizada
1	<i>FFT</i>	<i>DCUB</i> em 10s, <i>DCUB</i> em 5s
2	<i>FFT + EST</i>	<i>DE</i> em 10s, <i>DE</i> em 5s
3	<i>HT + FFT</i>	<i>DE</i> em 10s, <i>DE</i> em 5s e <i>DCUB</i> em 3s
4	<i>HT+FFT+EST</i>	<i>DE</i> em 10s, <i>DCUB</i> em 5s e <i>DCOSS</i> em 3s
5	<i>EST</i>	<i>DE</i> em 10s, <i>DE</i> em 5s, <i>DCOSS</i> em 3s, <i>DCOSS</i> em 1s, <i>DE</i> em 0,5s e <i>DCUB</i> em 0,25s

5.3 AVALIAÇÃO DAS MELHORES ABORDAGENS PARA CADA JANELA AMOSTRAL

Nesta seção é realizada uma comparação entre os cenários estudados, ou seja, entre os três melhores modelos investigados para cada janela amostral, a fim de escolher o melhor modelo em cada caso. Cabe mencionar que, as janelas amostrais iguais a 1s, 0.5s e 0.25s também foram testadas, mas apenas com as entradas processadas no domínio do tempo, razão pela qual não foram apresentadas em detalhes. De todo o modo, as melhores abordagens encontradas para aquelas janelas também são apresentadas e discutidas a seguir. Foi escolhido assim o melhor modelo *KNN*, *SVM* e o único modelo *RL* para cada um dos cenários.

No *Apêndice A* são apresentados os resultados obtidos para outras configurações das técnicas investigadas neste trabalho, ou seja, aqueles que obtiveram desempenho inferior aos detalhados nesta seção. Para cada cenário foram escolhidos, entre as variações dos parâmetros dos modelos *KNN* (Vizinhos mais próximos e distância métrica), *SVM* (Função *Kernel*, parâmetro *C* e γ) e *RL* (que não sofreu parametrizações) os que obtiveram os maiores valores de acordo com as métricas de avaliação (ver tabela 2.4).

No *Apêndice B* é apresentado um resumo acerca das técnicas empregadas para a

extração de características do sinal da corrente, bem como das diferentes configurações de algoritmos de aprendizagem de máquina utilizadas neste estudo.

5.3.1 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 10s

A tabela 5.2 apresentam os 5 melhores modelos, obtidos entre os 15 testados (Ver *Apêndice B*), para a janela de 10s.

Tabela 5.2: *Melhores abordagens para uma janela de 10s*

Melhores Classificadores	S	E	Ef	CCM	TA	ASC
KNN - Cúbico - 3 vizinhos - FFT	0,75	0,821	0,786	0,573	78,571	0,78
SVM - Gaussiano - $C = 440$ e $\gamma = 1,1$ - <i>FFT</i> e <i>Estatística</i>	0,875	0,928	0,902	0,805	90,179	0,90
SVM - Gaussiano - $C = 1$ e $\gamma = 0,43$ - <i>HT</i> e <i>FFT</i>	0,768	0,982	0,875	0,768	87,5	0,87
KNN - Euclidiano - 3 vizinhos - HT e FFT e Estatística	0,893	0,893	0,893	0,786	89,285	0,89
Regressão Logística - Estatística	0,804	0,911	0,857	0,718	85,714	0,86

A tabela 5.3 mostra a matriz de confusão do melhor classificador para a janela de 10s.

Tabela 5.3: *Matriz de confusão do melhor classificador SVM utilizando FFT e Estatísticas como ferramenta de extração das características em uma janela de 10s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	49	4	53
	Negative	7	52	59
Total		56	56	112

A figura 5.8 mostra as curvas *ROC* geradas para os cinco melhores modelos.

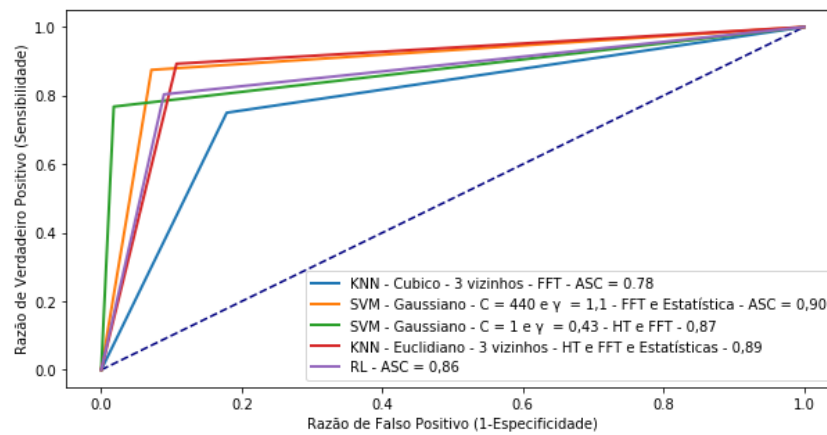


Figura 5.8: *Curvas ROC das melhores abordagens para uma janela de 10s*

A figura 5.9 destaca o melhor classificador entre os 5 avaliados, no caso o *SVM* com $C = 440$ e $\gamma = 1,1$, utilizando como entradas as grandezas *Estatísticas* e aquelas extraídas com a técnica *FFT*. Esta abordagem obteve uma *ASC* igual a 0.9, logo uma previsão excelente (ver tabela 2.5), uma sensibilidade em torno de 0.75 e uma especificidade em 0.821, portanto, isso mostra que o classificador obteve, na predição correta do motor saudável, mais acertos comparados com a predição correta do motor para uma barra rompida. Uma eficiência de 0.786, um *CCM* de 0.805, próximo de um, uma classificação perfeita e uma taxa de acerto de 90.179. Segundo a tabela 5.3, o classificador obteve uma maior tendência, mesmo que pequena, em detectar falsos negativos do que falsos positivos, ou seja, mesmo para situações nas quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

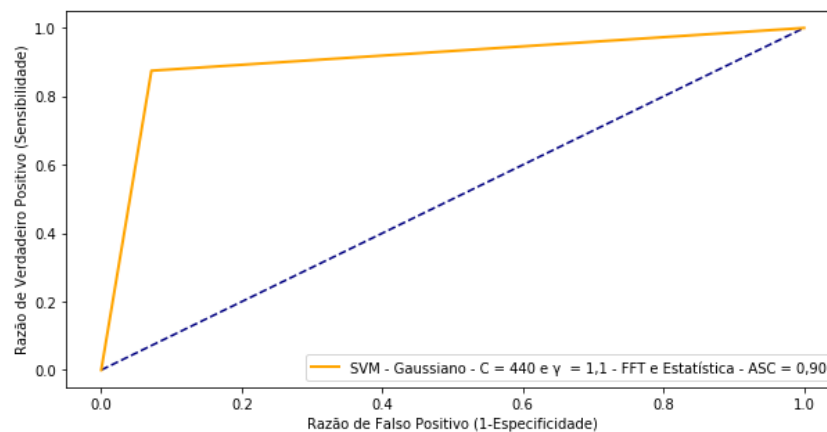


Figura 5.9: *Curva ROC do melhor classificador para uma janela de 10s*

5.3.2 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 5S

A tabela 5.4 apresenta os 5 melhores modelos, entre os 15 escolhidos, para a janela amostral de 5s.

Tabela 5.4: *Melhores abordagens para uma janela de 5s*

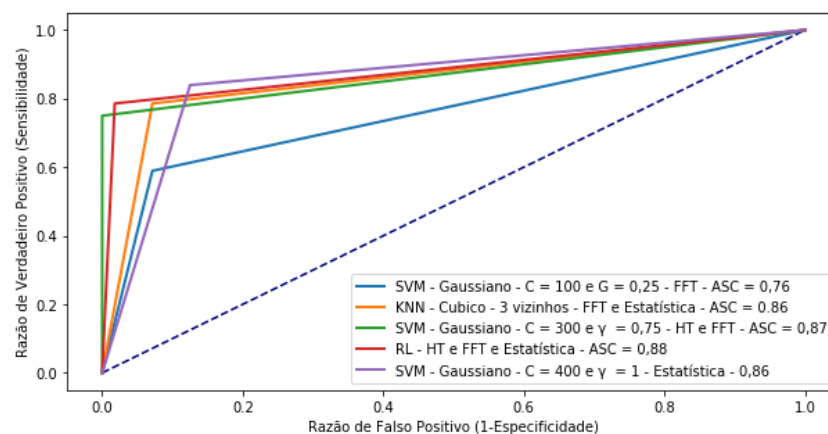
Melhores Classificadores	S	E	Ef	CCM	TA	ASC
KNN - Cúbico - 3 vizinhos - FFT	0,589	0,928	0,759	0,550	75,893	0,76
SVM - Gaussiano - $C = 440$ e $\gamma = 1,1$ - <i>FFT e Estatística</i>	0,786	0,928	0,857	0,722	85,714	0,86
SVM - Gaussiano - $C = 1$ e $\gamma = 0,43$ - <i>HT e FFT</i>	0,75	1	0,875	0,774	87,5	0,87
Regressão Logística - HT e FFT e Estatística	0,785	0,982	0,884	0,783	88,393	0,88
KNN - Euclidiano - 3 vizinhos - FFT e Estatística	0,839	0,875	0,8571	0,715	85,714	0,86

A tabela 5.5 mostra a matriz de confusão do melhor classificador para a janela de 5s.

Tabela 5.5: *Matriz de confusão do melhor classificador a RL utilizando a FFT e as grandezas estatísticas como entradas em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	4	43
	Negative	17	52	69
Total		56	56	112

A figura 5.10 mostra as curvas ROC geradas para os 5 melhores modelos.

**Figura 5.10:** *Curvas ROC das melhores abordagens para uma janela de 5s*

A figura 5.11 mostra a curva *ROC* gerada para o melhor classificador entre os 5 avaliados. Neste caso, a *RL* obteve o melhor desempenho, com uma $ASC = 0.88$, previsão boa (ver tabela 2.5), para as entradas extraídas com as técnicas *HT* e *FFT*, além das grandezas *Estatísticas*.

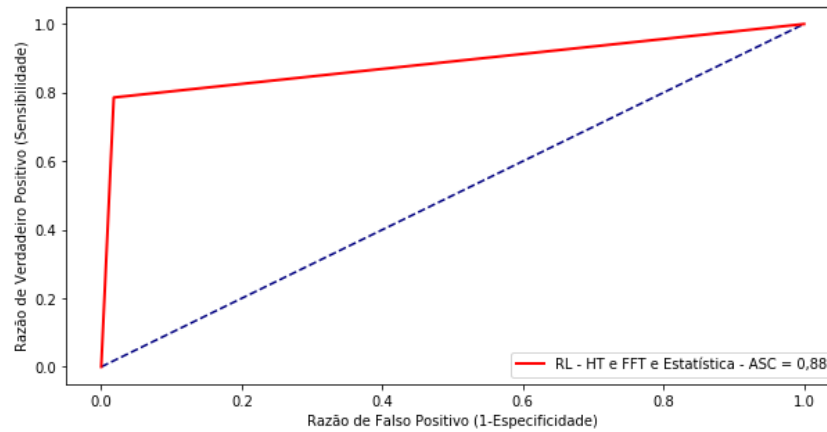


Figura 5.11: Curva *ROC* da melhor abordagem para uma janela de 5s

De forma semelhante ao caso anterior, a figura 5.11 destaca o melhor classificador entre os 5 avaliados, no caso a *RL*, utilizando como entradas as grandezas *Estatísticas* e aquelas extraídas com a técnica *HT* e *FFT*. Esta abordagem obteve uma sensibilidade em torno de 0.785 e uma especificidade em 0.982, portanto, isso mostra que o classificador obteve mais acertos na predição correta do motor saudável, comparado com a predição correta para o motor com uma barra rompida. Este caso alcançou uma eficiência de 0.884, um *CCM* de 0.783, próximo de um, assim uma classificação perfeita e uma taxa de acerto de 88.393. Segundo a tabela 5.5 o classificador, assim como no caso anterior, obteve uma maior tendência em detectar falsos negativos, 17 do que falsos positivos, 4, ou seja, também para situações nas quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

5.3.3 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 3s

Neste caso foram avaliados os melhores classificadores para a janela amostral de 3s, considerando também a melhor combinação das técnicas de processamento de sinais na extração das características de entrada. A tabela 5.6 apresenta as três melhores abordagens, entre as 9 testadas.

Tabela 5.6: *Melhores abordagens para uma janela de 3s*

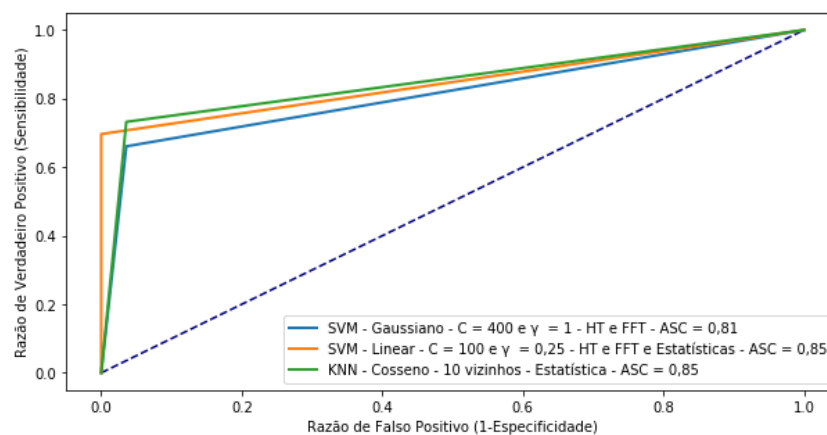
Melhores Classificadores	S	E	Ef	CCM	TA	ASC
SVM - Gaussiano - $C = 400$ e $\gamma = 1$ - <i>HT</i> e <i>FFT</i>	0,661	0,964	0,812	0,656	81,25	0,81
SVM - Linear - $C = 100$ e $\gamma = 0,25$ - <i>HT</i> e <i>FFT</i> e <i>Estatística</i>	0,696	1	0,848	0,731	84,821	0,85
KNN - Cosseno - 10 vizinhos - <i>Estatística</i>	0,732	0,964	0,848	0,716	84,821	0,85

A tabela 5.7 mostra a matriz de confusão do melhor classificador para a janela de 3s.

Tabela 5.7: *Matriz de confusão do classificador SVM utilizando HT e FFT e Estatísticas como ferramenta de extração das características em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	0	39
	Negative	17	56	73
Total		56	56	112

A figura 5.12 mostra as curvas *ROC* obtidas para os 3 melhores modelos.

**Figura 5.12:** *Curvas ROC das melhores abordagens para uma janela de 3s*

A figura 5.13 mostra a curva *ROC* para a melhor abordagem inteligente encontrada nesse caso, que foi o *SVM* com $C = 100$ e $\gamma = 0,25$, utilizando características extraídas

com as técnicas *HT* e *FFT*, além das informações extraídas no domínio do tempo com o uso das grandezas *Estatísticas*.

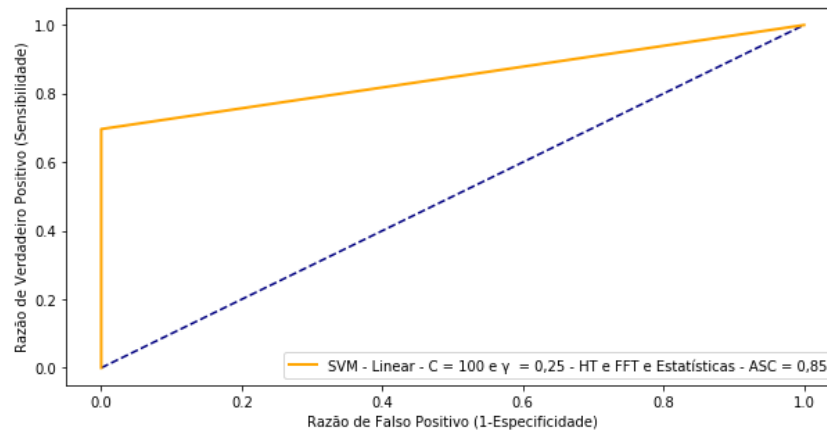


Figura 5.13: Curva ROC do melhor classificador para uma janela de 3s

A figura 5.13 destaca o melhor classificador entre os 3 avaliados para esta janela amostral, no caso a *RL*, utilizando como entradas as grandezas *Estatísticas* e aquelas extraídas com a técnica *HT* e *FFT*. Esta abordagem obteve uma sensibilidade em torno de 0.696 e uma especificidade em 1, portanto, isso mostra que o classificador obteve mais acertos na predição correta do motor saudável, comparado com a predição correta do motor com uma barra rompida. Este caso alcançou uma eficiência de 0.848, um *CCM* de 0.731, próximo de um, uma classificação perfeita e uma taxa de acerto de 84.821. Segundo a tabela 5.7 o classificador, assim como no caso anterior, obteve uma maior tendência em detectar falsos negativos(17), do que falsos positivos(0). Como nos casos anteriores, para condições nas quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

5.3.4 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 1S

Como já mencionado, as soluções avaliadas para as janelas de 10s, 5s e 3s também foram testadas em janelas inferiores, como 1s, 0.5s e 0.25s, uma vez que um volume menor de dados discretizados pode contribuir para uma redução não apenas em tempo de processamento de um sistema inteligente, mas também para a aplicação do sistema de diagnóstico em tempo real. Neste sentido, a tabela 5.8 apresenta a melhor abordagem entre as 3 testadas, para uma janela de 1s.

Tabela 5.8: *Melhores abordagens para uma janela de 1s*

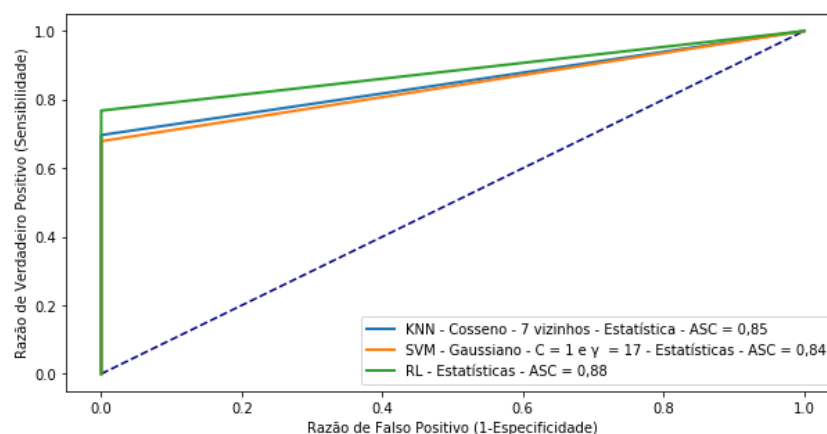
Melhores Classificadores	S	E	Ef	CCM	TA	ASC
KNN - Cosseno - 7 vizinhos - <i>Estatística</i>	0,696	1	0,848	0,731	84,821	0,85
SVM - Gaussiano - $C = 1$ e $\gamma = 17$ - <i>Estatística</i>	0,678	1	0,839	0,7166	83,928	0,84
Regressão Logística - <i>Estatística</i>	0,768	1	0,884	0,789	88,393	0,88

A tabela 5.9 mostra a matriz de confusão do melhor classificador para a janela de 1s.

Tabela 5.9: *Matriz de confusão do classificador RL utilizando Estatísticas como ferramenta de extração das características em uma janela de 1s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	43	0	43
	Negative	13	56	69
Total		56	56	112

A figura 5.14 mostra as curvas *ROC* geradas para os 3 melhores modelos.

**Figura 5.14:** *Curvas ROC das melhores abordagens para uma janela de 1s*

A figura 5.15 destaca o melhor resultado, ou seja, a curva *ROC* obtida com o uso da regressão logística, em uma janela de 1s, utilizando como entradas apenas as características *Estatísticas*. Tal abordagem encontrou uma $ASC = 0,88$, previsão boa.

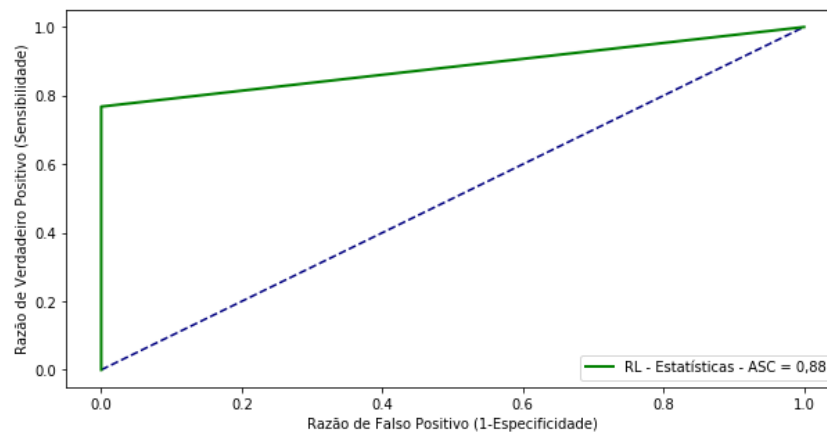


Figura 5.15: Curva ROC da melhor abordagem para uma janela de 1s

A figura 5.15 destaca o melhor classificador entre os 3 avaliados para esta janela de 1s, no caso a *RL*, utilizando como entradas apenas as grandezas *Estatísticas*. Esta abordagem obteve uma sensibilidade em torno de 0.77 e uma especificidade de 1, portanto, isso mostra que o classificador acertou todos os casos na predição do motor saudável. Foi obtida uma eficiência de 0.88, um *CCM* de 0.789, próximo de um, assim avaliado como uma classificação perfeita e uma taxa de acerto de 88.39. Segundo a tabela 5.9 o classificador, assim como nos outros casos, obteve uma maior tendência em detectar falsos negativos (13 no total), do que falsos positivos (0, nenhum caso). Logo, a exemplo dos casos anteriores, para situações nas quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

5.3.5 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 0,5S

Neste caso são apresentadas as melhores abordagens encontradas para uma janela amostral de 0,5s. A tabela 5.10 apresenta o melhor modelo entre os três escolhidos anteriormente para a janela de 0,5s.

Tabela 5.10: Melhores abordagens para uma janela de 0,5s

Melhores Classificadores	S	E	Ef	CCM	TA	ASC
KNN - Euclidiano - 7 vizinhos - Estatística	0,696	0,911	0,803	0,621	80,357	0,8
SVM - Gaussiano - $C = 1$ e $\gamma = 4,2$ - Estatística	0,7143	0,982	0,848	0,723	84,821	0,85
Regressão Logística - Estatística	0,732	0,803	0,768	0,537	76,786	0,77

A tabela 5.11 mostra a matriz de confusão do melhor classificador para a janela de 0.5s.

Tabela 5.11: *Matriz de confusão do classificador SVM utilizando Estatísticas como ferramenta de extração das características em uma janela de 0.5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	40	1	41
	Negative	16	55	71
Total		56	56	112

A figura 5.16 mostra as curvas *ROC* geradas para os 3 melhores casos testados.

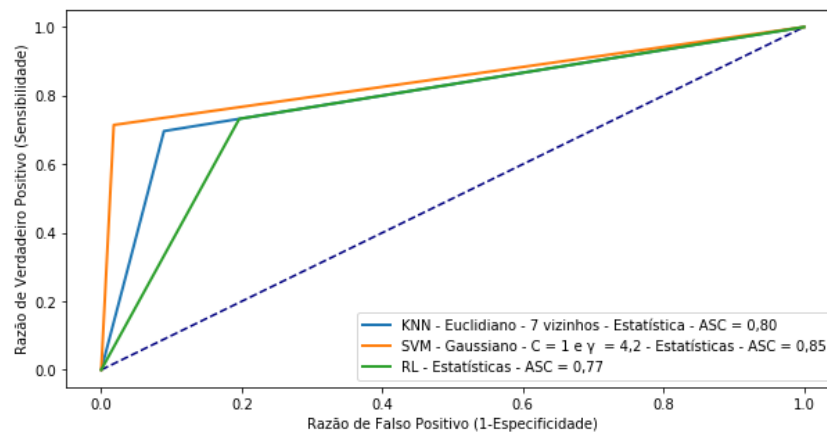


Figura 5.16: *Curvas ROC das melhores abordagens para uma janela de 0,5s*

A figura 5.17 mostra a curva *ROC* da melhor solução encontrada, no caso o classificador *SVM* ajustado com os valores de $C = 1$ e $\gamma = 4,2$, para a janela de 0,5s, utilizando como entradas apenas as características *Estatísticas*, a exemplo do caso anterior avaliado em uma janela amostral de 1s. Essa abordagem obteve uma $ASC = 0,85$, previsão boa (ver tabela 2.5), portanto, um pouco inferior ao caso processado em 1s ($ASC = 0,88$).

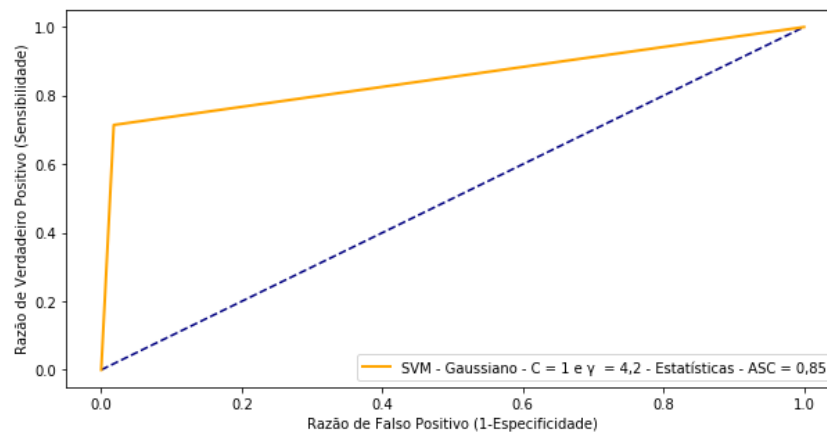


Figura 5.17: Curva ROC da melhor solução para uma janela de 0,5s

A figura 5.17 destaca o melhor classificador entre os 3 avaliados, para uma janela amostral de 0,5s, no caso o SVM com $C = 1$ e $\gamma = 4.2$, utilizando como entradas as grandezas *Estatísticas*. Esta abordagem obteve uma sensibilidade em torno de 0.71 e uma especificidade em 0.98, portanto, assim como as outras avaliações, mostra que o classificador obteve, na predição correta do motor saudável, mais acertos, comparado com a predição correta do motor para uma barra rompida. Uma eficiência de 0.85, um CCM de 0.72, próximo de um, indicam uma classificação perfeita e uma taxa de acerto de 84.82. Segundo a tabela 5.11, o classificador obteve uma maior tendência em detectar falsos negativos (16), do que falsos positivos (1), ou seja, mesmo para situações nas quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

5.3.6 MELHORES ABORDAGENS OBTIDAS PARA UMA JANELA DE 0,25S

Um último estudo foi realizado para uma janela amostral bastante reduzida, com apenas 0,25s. Deste modo, a tabela 5.12 apresenta o melhor modelo entre os 3 escolhidos, para a janela de 0,25s.

Tabela 5.12: Melhores abordagens para uma janela de 0,25s

Melhores Classificadores	S	E	Ef	CCM	TA	ASC
KNN - Cúbico - 5 vizinhos - <i>Estatística</i>	0,643	0,786	0,714	0,433	71,428	0,71
SVM - Gaussiano - $C = 1$ e $\gamma = 4,2$ - <i>Estatística</i>	0,661	0,875	0,768	0,548	76,785	0,77
Regressão Logística - <i>Estatística</i>	0,571	0,661	0,616	0,233	61,607	0,61

A tabela 5.13 mostra a matriz de confusão do melhor classificador para a janela de 0.25s.

Tabela 5.13: *Matriz de confusão do classificador SVM utilizando Estatísticas como ferramenta de extração das características em uma janela de 0.5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	37	7	44
	Negative	19	49	68
Total		56	56	112

A figura 5.18 mostra as curvas *ROC* geradas para os 3 melhores modelos.

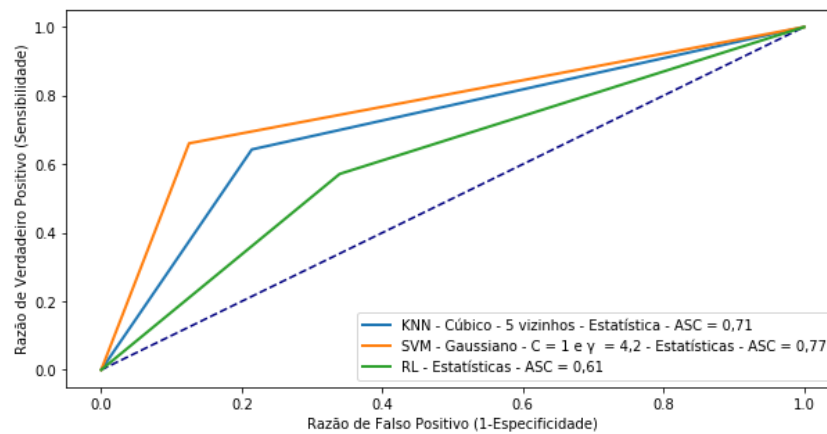


Figura 5.18: *Curvas ROC das melhores abordagens para uma janela de 0,25s*

A figura 5.19 mostra a curva *ROC* da melhor solução, no caso um classificador *SVM* configurado com $C = 1$ e $\gamma = 4,2$, na janela de 0,25s e utilizando como entradas apenas as características *Estatísticas*. Nota-se que nesta janela o valor da *ASC*, que foi de 0.77, previsão fraca (ver tabela 2.5), mesmo para o melhor modelo, foi bastante inferior aos casos analisados nas janelas de 1s e 0,5s.

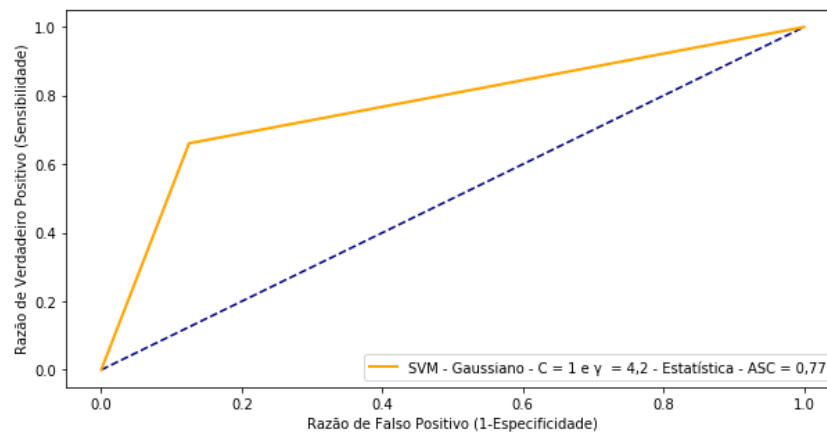


Figura 5.19: Curva ROC da melhor solução para uma janela de 0,25s

A figura 5.19 destaca o melhor classificador entre os 3 avaliados, para uma janela de 0,25s, no caso o SVM com $C = 1$ e $\gamma = 4.2$, utilizando como entradas as grandezas *Estatísticas*. Esta abordagem obteve uma sensibilidade em torno de 0.66 e uma especificidade em 0.87, portanto, assim como nas outras avaliações, mostra que o classificador obteve, na predição correta do motor saudável, mais acertos, comparado com a predição correta do motor com uma barra rompida. Uma eficiência de 0.76, um CCM de 0.55, próximo de zero, uma classificação aleatória e uma taxa de acerto de 76.78. Segundo a tabela 5.11, o classificador obteve uma maior tendência em detectar falsos negativos(19), do que falsos positivos(7), ou seja, mesmo para os casos nos quais a gaiola apresenta uma barra quebrada o modelo indicaria uma condição de rotor saudável.

Com base no exposto, a tabela 5.14 resume as melhores abordagens encontradas para cada janela amostral.

Tabela 5.14: *Parâmetros dos melhores modelos em cada janela amostral*

Janelas	Melhores classificadores com os seus parâmetros	S	E	Ef	CCM	TA	ASC
10s	<i>SVM - Gaussiano - $C = 440$ e $\gamma = 1,1$ - FFT e Estatística</i>	0,875	0,928	0,902	0,805	90,179	0,90
5s	<i>Regressão Logística - HT e FFT e Estatística</i>	0,785	0,982	0,884	0,783	88,393	0,88
3s	<i>SVM - Linear - $C = 100$ e $\gamma = 0,25$ - HT e FFT e Estatística</i>	0,696	1	0,848	0,731	84,821	0,85
1s	<i>Regressão Logística - Estatística</i>	0,768	1	0,884	0,789	88,393	0,88
0,5s	<i>SVM - Gaussiano - $C = 1$ e $\gamma = 4,2$ - Estatística</i>	0,7143	0,982	0,848	0,723	84,821	0,85
0,25s	<i>SVM - Gaussiano - $C = 1$ e $\gamma = 4,2$ - Estatística</i>	0,661	0,875	0,768	0,548	76,785	0,77

Na comparação entre os melhores modelos, ou abordagens, nota-se que a janela amostral de 10s permitiu alcançar uma melhor performance durante os testes realizados para a validação dos resultados, com a combinação das entradas nos domínios do tempo e da frequência. De todo o modo, foi possível constatar ainda que as melhores abordagens encontradas nas janelas de 5s e 1s também encontraram valores de desempenho bastante próximos, muito embora a solução testada na janela de 5s tenha obtido uma maior sensibilidade durante a validação dos dados, face ao valor observado na janela de 1s. Neste último caso, poderiam ser reduzidos os casos de falso negativo na avaliação da avaria na gaiola rotórica.

Finalmente, os resultados encontrados entre as melhores abordagens, para qualquer janela amostral, mostram que as entradas extraídas com as grandezas estatísticas são fundamentais para incrementar a performance dos classificadores inteligentes. A seguir são apresentadas algumas considerações adicionais acerca dos modelos testados no presente trabalho.

Este trabalho apresentou um estudo comparativo entre as principais técnicas de processamento de sinais empregadas na extração de características da corrente do estator de um motor de indução trifásico, do tipo gaiola de esquilo, e técnicas inteligentes capazes de detectar e classificar o rompimento de barras que fazem parte da gaiola rotórica. Como descrito no Capítulo 1, tais falhas podem provocar danos e efeitos indesejáveis não apenas para o motor, mas também para o estator e toda a máquina.

Para a extração das características no domínio da frequência foram testadas as Transformadas de *Fourier* e *Hilbert* e para a avaliação de outras características no domínio do tempo foram usadas algumas métricas estatísticas, bastante conhecidas na literatura que tratam sobre o defeito no rotor dos motores de indução. Todas as características foram utilizadas no sentido de construir uma base de conhecimento representativa para o rotor com uma barra quebrada e também na condição saudável.

O estudo comparou o desempenho de diversas abordagens, formadas por um conjunto de entradas distintas e diferentes classificadores inteligentes, como o *KNN*, o *SVM* e a *RL*, uma vez que tais modelos são bastante empregados na avaliação das condições operacionais das máquinas elétricas. Os resultados encontrados permitiram observar que, dentre os cenários investigados, não apenas a característica de entrada foi significativa para incrementar a performance de uma determinada abordagem, mas também o ajuste dos parâmetros selecionados para cada classificador interferiu de maneira decisiva no desempenho final de cada solução.

Como exemplo, pode-se citar que a janela amostral em 10s ofereceu o melhor resultado com o uso do classificador *SVM*, mas tomando como entradas as características extraídas com a técnica *FFT* mais as grandezas estatísticas, além de considerar o classificador ajustado com os parâmetros $C = 440$, $\gamma = 1,1$ e função *Kernel* Gaussiano. Deste modo, outras configurações para o mesmo classificador, também na janela de 10s, obtiveram resultados inferiores.

Adicionalmente, vale mencionar que, para aquelas abordagens que permitiram extrair características no domínio da frequência, como nas janelas em 10s, 5s, e 3s, os melhores resultados foram obtidos quando da combinação de tais características com aquelas oriundas das grandezas estatísticas, ou seja, o uso conjunto de informações extraídas dos domínios do tempo e da frequência incrementa bastante o desempenho da solução final.

Cumprir notar ainda que, a melhor solução encontrada no presente trabalho foi capaz de detectar o defeito no rotor em 90% dos casos, para o motor operando em baixa carga, ou baixo escorregamento, para uma janela amostral de 10s, o que seria bastante improvável alcançar com o uso da tradicional técnica *MCSA*, uma vez que este método exige ao menos 100s para extrair as características de interesse, conforme descrito no Capítulo 3. Por outro lado, novos estudos devem ser realizados no sentido de incrementar ainda mais a performance das soluções ora investigadas, como:

1. Avaliar outras técnicas para a extração de características da corrente do estator, como aquelas conhecidas como modelos de sub-espço;
2. Avaliar o uso de técnicas de redução capazes de selecionar as melhores características;
3. Avaliar o desempenho das técnicas testadas e outras que operam nos domínios do tempo e frequência, para o motor operando em velocidade e/ou carga variável; e
4. Investigar o desempenho das abordagens inteligentes para identificar e classificar a severidade do defeito no rotor, ou seja, o número de barras quebradas.

Trabalhos Oriundos da Dissertação

- R.C.Jesus, C.G. Dias, A.C.Lira e L.C. Silva "**Um Estudo Comparativo entre Técnicas de Aprendizado de Máquina e Processamento de Sinais para Detectar Barras Quebradas em um Motor de Indução**", 2018 *13th IEEE International Conference on Industry Applications (INDUSCON)*, São Paulo, SP, Brasil.
- R.C.Jesus, C.G. Dias, A.C.Lira, J. Marcelo e K.Rodrigues "**Uma Abordagem no Domínio da Frequência para o Diagnóstico de Barras Rompidas em Motores de Indução Usando a Máquina de Suporte Vetorial**", 2018 *3º Seminário em Tecnologia da Informação Inteligente (SETII)*, São Paulo, SP, Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

ALPAYDIN, E. *Introduction to Machine Learning*. Third edit. [S.l.: s.n.], 2004. 613 p. ISBN 0-262-01211-1. Citado na pág. 51.

ANTONINO-DAVIU, J. et al. Validation of a new method for the diagnosis of rotor bar failures via wavelet transform in industrial induction machines. *IEEE Trans. Ind. Appl.*, v. 42, n. 4, p. 990–996, jul 2006. ISSN 0093-9994. Disponível em: <http://ieeexplore.ieee.org/document/1658328/>. Citado na pág. 60.

BAKSHI, K.; BAKSHI, K. Considerations for artificial intelligence and machine learning: Approaches and use cases. In: *2018 IEEE Aerosp. Conf.* IEEE, 2018. p. 1–9. ISBN 978-1-5386-2014-4. Disponível em: <https://ieeexplore.ieee.org/document/8396488/>. Citado na pág. 45.

BELLINI, A. et al. On-field experience with online diagnosis of large induction motors cage failures using MCSA. *IEEE Trans. Ind. Appl.*, v. 38, n. 4, p. 1045–1053, jul 2002. ISSN 0093-9994. Disponível em: <http://ieeexplore.ieee.org/document/1019961/>. Citado na pág. 60.

BELLINI, A.; FRANCESCHINI, G.; TASSONI, C. Monitoring of induction machines by maximum covariance method for frequency tracking. In: *Conf. Rec. 2004 IEEE Ind. Appl. Conf. 2004. 39th IAS Annu. Meet.* IEEE, 2006. v. 2, n. 1, p. 743–749. ISBN 0-7803-8486-5. ISSN 00939994. Disponível em: <http://ieeexplore.ieee.org/document/1348497/>. Citado na pág. 59, 60.

BELLINI, A. et al. High Frequency Resolution Techniques for Rotor Fault Detection of Induction Machines. *IEEE Trans. Ind. Electron.*, v. 55, n. 12, p. 4200–4209, dec 2008. ISSN 0278-0046. Disponível em: <http://ieeexplore.ieee.org/document/4663714/>. Citado na pág. 32.

BONNETT, A.; SOUKUP, G. Cause and analysis of stator and rotor failures in 3-phase squirrel cage induction motors. In: *Conf. Rec. 1991 Annu. Pulp Pap. Ind. Tech. Conf.* IEEE, 1992. v. 28, n. 4, p. 22–42. Disponível em: <http://ieeexplore.ieee.org/document/239667/>. Citado na pág. 33.

BOUZIDA, A. et al. Fault Diagnosis in Industrial Induction Machines Through Discrete Wavelet Transform. *IEEE Trans. Ind. Electron.*, v. 58, n. 9, p. 4385–4395, sep 2011. ISSN 0278-0046. Disponível em: http://ieeexplore.ieee.org/xpl/freeabs{_}all.jsp?arnumber=5645682 <http://ieeexplore.ieee.org/document/56456>. Citado na pág. 58.

CARDOSO, A. J. M. *Diagnóstico de Avarias em Motores de Indução Trifásicos*. [S.l.: s.n.], 1991. Citado na pág. 25, 26.

CHAPMAN, C.; STEPHEN, J. *Fundamentos de Máquinas Elétricas*. 5ª edição. ed. Porto Alegre: [s.n.], 2013. 684 p. ISBN 0073529540/9780073529547. Citado na pág. 15, 31, 32, 34, 35, 36, 37.

Claude Sammut, G. I. W. *Encyclopedia of Machine Learning*. Sydney, Australia: Springer Science & Business Media, 2011. 1031 p. ISBN 0387307680, 9780387307688. Disponível em: <http://link.springer.com/10.1007/978-0-387-30164-8>. Citado na pág. 47, 49.

COOLEY, J.; LEWIS, P.; WELCH, P. Historical Notes on the Fast Fourier Transform. *IEEE Trans. Audio Electroacoust.*, v. 55, n. 2, p. 11675–1677, jun 1967. ISSN 0018-9278. Disponível em: <<http://ieeexplore.ieee.org/document/1161904/>>. Citado na pág. 41.

CULBERT, I.; RHODES, W. Using Current Signature Analysis Technology to Reliably Detect Cage Winding Defects in Squirrel Cage Induction Motors. In: *Rec. Conf. Pap. Ind. Appl. Soc. 52nd Annu. Pet. Chem. Ind. Conf.* IEEE, 2007. v. 43, n. 2, p. 95–101. ISBN 0-7803-9272-8. ISSN 00939994. Disponível em: <<http://ieeexplore.ieee.org/document/1524544/>>. Citado na pág. 60.

Di Tommaso, A. O.; MICELI, R.; GALLUZZO, G. R. Monitoring and diagnosis of failures in squirrel-cage induction motors due to cracked or broken bars. In: *8th IEEE Symp. Diagnostics Electr. Mach. Power Electron. Drives.* IEEE, 2011. p. 39–44. ISBN 978-1-4244-9301-2. Disponível em: <<http://ieeexplore.ieee.org/document/6063599/>>. Citado na pág. 57.

DIAS, C. G. *Proposta de um novo método para a detecção de barras rompidas em motores de indução com rotor em gaiola.* 1–204 p. Tese (Tese de Doutorado) — Escola Politécnica da Universidade de São Paulo, 2006. Citado na pág. 15, 26, 27, 35.

DIAS, C. G.; CHABU, I. E. Spectral Analysis Using a Hall Effect Sensor for Diagnosing Broken Bars in Large Induction Motors. *IEEE Trans. Instrum. Meas.*, v. 63, n. 12, p. 2890–2902, dec 2014. ISSN 0018-9456. Disponível em: <<http://ieeexplore.ieee.org/document/6842611/>>. Citado na pág. 34, 57.

DIAS, C. G.; PEREIRA, F. H. Broken Rotor Bars Detection in Induction Motors Running at Very Low Slip Using a Hall Effect Sensor. *IEEE Sens. J.*, v. 18, n. 11, p. 4602–4613, jun 2018. ISSN 1530-437X. Disponível em: <<https://ieeexplore.ieee.org/document/8338052/>>. Citado na pág. 28, 29, 52, 57, 61, 62, 63, 66.

EMATSU, M. Y. *Análise espectral da corrente de partida de motores de indução para detecção de falhas nas barras do rotor.* 1–100 p. Tese (Dissertação de Mestrado) — Universidade Federal de Santa Catarina, 2008. Disponível em: <<http://repositorio.ufsc.br/xmlui/handle/123456789/91388>>. Citado na pág. 59, 60.

FERNANDEZ-TEMPRANO, M. et al. Broken bar condition monitoring of an induction motor under different supplies using a linear discriminant analysis. In: *2013 9th IEEE Int. Symp. Diagnostics Electr. Mach. Power Electron. Drives.* IEEE, 2013. p. 162–168. ISBN 978-1-4799-0025-1. Disponível em: <<http://ieeexplore.ieee.org/document/6645712/>>. Citado na pág. 45, 59, 62, 63, 66.

FILIPPETTI, F.; BELLINI, A.; CAPOLINO, G. Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives. In: *2013 IEEE Work. Electr. Mach. Des. Control Diagnosis.* IEEE, 2013. p. 196–209. ISBN 978-1-4673-5658-9. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6525180> <http://ieeexplore.ieee.org/document/6525180/>>. Citado na pág. 32.

FISER, R. et al. Diagnostic system for on-line detection of rotor faults in induction motor drives. In: *8th IEEE Symp. Diagnostics Electr. Mach. Power Electron. Drives.* IEEE, 2011. p. 77–83. ISBN 978-1-4244-9301-2. Disponível em: <<http://ieeexplore.ieee.org/document/6063605/>>. Citado na pág. 15, 26, 58.

Fitzgerald, A.E. and KINGSLEY, C. and Umans, S. *Máquinas Elétricas com Introdução à Eletrônica de Potência*. 6ª edição. ed. [s.n.], 2006. 648 p. ISBN 9788560031047. Disponível em: <<https://books.google.com.br/books?id=zsH5FZYJhb4C>>. Citado na pág. 31.

FUKUNAGA, K.; NARENDRA, P. A Branch and Bound Algorithm for Computing k-Nearest Neighbors. *IEEE Trans. Comput.*, C-24, n. 7, p. 750–753, jul 1975. ISSN 0018-9340. Disponível em: <http://ieeexplore.ieee.org/xpl/freeabs{_}all.jsp?arnumber=1672890 <http://ieeexplore.ieee.org/document/16728>>. Citado na pág. 52.

GARCIA, A. G. P. *Impacto da Lei de Eficiência Energética para Motores Elétricos no Potencial de Conservação de Energia na Indústria*. 139 p. Tese (Dissertação de Mestrado) — Universidade Federal do Rio de Janeiro, 2003. Disponível em: <<http://www.ppe.ufrj.br/index.php/pt/publicacoes/dissertacoes/2003/1154-impacto-da-lei-de-eficiencia-energetica-para-motores-eletricos-no-potencial-de-conservacao-de-energia-na-industria>>. Citado na pág. 22.

GARDEL, P. et al. Neural network broken bar detection using time domain and current spectrum data. In: *2012 XXth Int. Conf. Electr. Mach.* IEEE, 2012. p. 2492–2497. ISBN 978-1-4673-0142-8. Disponível em: <<http://ieeexplore.ieee.org/document/6350234/>>. Citado na pág. 66.

GODOY, W. F. *Multiclassificador Inteligente de Falhas no Domínio do Tempo em Motores de Indução Trifásicos Alimentados por Inversores de Frequência*. 164 p. Tese (Tese de Doutorado) — Escola de Engenharia de São Carlos da Universidade de São Paulo, 2016. Citado na pág. 22, 28.

Gordi Armaki, M.; ROSHANFEKR, R. A new approach for fault detection of broken rotor bars in induction motor based on support vector machine. In: *2010 18th Iran. Conf. Electr. Eng.* IEEE, 2010. p. 732–738. ISBN 978-1-4244-6760-0. Disponível em: <<http://ieeexplore.ieee.org/document/5506976/>>. Citado na pág. 50.

GYFTAKIS, K. et al. Comparative Experimental Investigation of the Broken Bar Fault Detectability in Induction Motors. *IEEE Trans. Ind. Appl.*, v. 52, n. 2, p. 1–1, 2015. ISSN 0093-9994. Disponível em: <<http://ieeexplore.ieee.org/document/7347401/>>. Citado na pág. 28, 61.

HARRINGTON, P. *Machine Learning in Action*. Shelter Island, NY: Manning Publications Company, 2012, 2015. 382 p. ISBN 1617290181, 9781617290183. Citado na pág. 15, 47, 48, 49.

HEIDEMAN, M.; JOHNSON, D.; BURRUS, C. Gauss and the history of the fast fourier transform. *IEEE ASSP Mag.*, v. 1, n. 4, p. 14–21, oct 1984. ISSN 0740-7467. Disponível em: <<http://ieeexplore.ieee.org/document/1162257/>>. Citado na pág. 42.

ISHIKAWA, T.; SHINAGAWA, S.; KURITA, N. Analysis and failure diagnosis of squirrel-cage induction motor with broken rotor. In: *15th Int. Conf. Electr. Mach. Syst.* [S.l.: s.n.], 2013. v. 2, n. 6, p. 292–297. Citado na pág. 58.

JOHANSSON, M. *The Hilbert Transform*. 1–36 p. Tese (Master Thesis) — Växjö University, Suecia, 1999. Disponível em: <<http://www.fuchs-braun.com/media/d9140c7b3d5004fbffff8007ffffff0.pdf>>. Citado na pág. 44.

Karl L. Wildes, Nilo A. Lindgren, N. L. *A Century of Electrical Engineering and Computer Science at MIT, 1882-1982*. [S.l.]: MIT Press, 1985. 423 p. ISBN 0262231190.

Citado na pág. 43.

KARMAKAR, S. et al. *Induction Motor Fault Diagnosis*. Singapore: Springer Singapore, 2016. 161 p. (Power Systems). ISBN 978-981-10-0623-4. Disponível em: <http://link.springer.com/10.1007/978-981-10-0624-1>. Citado na pág. 27, 59, 60.

Khadim Moin Siddiqui; GIRI, V. Broken rotor bar fault detection in induction motors using Wavelet Transform. In: *2012 Int. Conf. Comput. Electron. Electr. Technol.* IEEE, 2012. v. 2, n. 4, p. 1–6. ISBN 978-1-4673-0212-8. Disponível em: http://ieeexplore.ieee.org/xpls/abs/_all.jsp?arnumber=1665592 <http://ieeexplore.ieee.org/document/62037>. Citado na pág. 58.

KIA, S.; HENAO, H.; CAPOLINO, G.-A. Diagnosis of Broken-Bar Fault in Induction Machines Using Discrete Wavelet Transform Without Slip Estimation. *IEEE Trans. Ind. Appl.*, v. 45, n. 4, p. 1395–1404, jul 2009. Disponível em: <http://ieeexplore.ieee.org/document/4957017>. Citado na pág. 58.

KUBAT, M. *An Introduction to Machine Learning*. Coral Gables, FL, USA ISBN: Springer International Publishing, 2015. 296 p. ISSN 13455354. ISBN 978-3-319-20009-5. Disponível em: <http://link.springer.com/10.1007/978-3-319-20010-1>. Citado na pág. 15, 50, 51, 52, 53.

LAALA, W.; GUEDINI, S.; ZOUZOU, S. Novel approach for diagnosis and detection of broken bar in induction motor at low slip using fuzzy logic. In: *8th IEEE Symp. Diagnostics Electr. Mach. Power Electron. Drives*. IEEE, 2011. p. 511–516. ISBN 978-1-4244-9301-2. Disponível em: <http://ieeexplore.ieee.org/document/6063671>. Citado na pág. 27, 58, 61, 63, 74.

LEE, S. B. et al. Identification of false rotor fault indications produced by on-line MCSA for medium voltage induction machines. In: *2015 61st IEEE Pulp Pap. Ind. Conf.* IEEE, 2015. p. 1–9. ISBN 978-1-4799-7113-8. Disponível em: <http://ieeexplore.ieee.org/document/7165709>. Citado na pág. 58.

LIANG, X.; EDMONWANDERKHOE, K. Condition monitoring techniques for induction motors. In: *2017 IEEE Ind. Appl. Soc. Annu. Meet.* IEEE, 2017. p. 1–10. ISBN 978-1-5090-4894-6. Disponível em: <http://ieeexplore.ieee.org/document/8101860>. Citado na pág. 15, 22, 33, 34, 57, 61.

MARTIN-DIAZ, I. et al. An Experimental Comparative Evaluation of Machine Learning Techniques for Motor Fault Diagnosis Under Various Operating Conditions. *IEEE Trans. Ind. Appl.*, v. 54, n. 3, p. 2215–2224, may 2018. ISSN 0093-9994. Disponível em: http://www.ieee.org/publications_standards/publications/rights/index.html <https://ieeexplore.ieee.org/document/82810>. Citado na pág. 50, 62, 63, 76, 77.

MCKINNON, D. L. Using a six fault zone approach for predictive maintenance on motors. In: *2007 Electr. Insul. Conf. Electr. Manuf. Expo.* IEEE, 2007. p. 253–264. ISBN 978-1-4244-0446-9. Disponível em: <http://ieeexplore.ieee.org/document/4562630>. Citado na pág. 27.

MICERINO, F. J. *Impacto da utilização de motores de alto rendimento e sua aplicação no leilão de eficiência energética*. 122 p. Tese (Dissertação de Mestrado) — Universidade de São Paulo, São Paulo, may 2012. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/3/3143/tde-14062013-150047/>>. Citado na pág. 22.

Morais de Sousa, C. *UMA ABORDAGEM INTELIGENTE APLICADA NA DETECÇÃO DE BARRAS ROMPIDAS EM MOTORES DE INDUÇÃO OPERANDO EM BAIXO ESCORREGAMENTO*. 111 p. Tese (Dissertação de Mestrado) — Universidade Nove de Julho, 2017. Citado na pág. 15, 22, 33, 61, 63.

NAHA, A. et al. A Method for Detecting Half-Broken Rotor Bar in Lightly Loaded Induction Motors Using Current. *IEEE Trans. Instrum. Meas.*, v. 65, n. 7, p. 1614–1625, jul 2016. ISSN 0018-9456. Disponível em: <<http://ieeexplore.ieee.org/document/7450652/>>. Citado na pág. 58.

NAHA, A. et al. Low Complexity Motor Current Signature Analysis Using Sub-Nyquist Strategy With Reduced Data Length. *IEEE Trans. Instrum. Meas.*, v. 66, n. 12, p. 3249–3259, dec 2017. ISSN 0018-9456. Disponível em: <<http://ieeexplore.ieee.org/document/8019870/>>. Citado na pág. 59.

NG, A. *CS229 Lectures Notes*. 2012. Citado na pág. 15, 46, 47, 48.

PARK, S. H.; GOO, J. M.; JO, C.-H. Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean J. Radiol.*, v. 5, n. 1, p. 11, 2004. ISSN 1229-6929. Disponível em: <<https://doi.org/10.3348/kjr.2004.5.1.11> <https://synapse.koreamed.org/DOIX.php?id=10.3348/kjr.2004.5.1.11>>. Citado na pág. 56.

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. Evaluating Classifiers Using ROC Curves. *IEEE Lat. Am. Trans.*, v. 6, n. 2, p. 215–222, jun 2008. Disponível em: <<http://ieeexplore.ieee.org/document/4609920/>>. Citado na pág. 55.

PUCHE-PANADERO, R. et al. Improved Resolution of the MCSA Method Via Hilbert Transform, Enabling the Diagnosis of Rotor Asymmetries at Very Low Slip. *IEEE Trans. Energy Convers.*, v. 24, n. 1, p. 52–59, mar 2009. ISSN 0885-8969. Disponível em: <<http://ieeexplore.ieee.org/document/4749311/>>. Citado na pág. 27, 33, 43, 44, 58, 63, 74.

ROÇAS, A. A. C. *Análise de Qualidade e Processos de Bobinagem de Motores de Indução de Baixa Tensão*. 112 p. Tese (Dissertação de Mestrado) — Faculdade de Engenharia da Universidade Porto, Fundação Universidade Regional de Blumenau, 2012. Citado na pág. 31.

ROMERO-TRONCOSO, R. et al. Rotor unbalance and broken rotor bar detection in inverter-fed induction motors at start-up and steady-state regimes by high-resolution spectral analysis. *Electr. Power Syst. Res.*, Elsevier B.V., v. 133, p. 142–148, apr 2016. ISSN 03787796. Disponível em: <<http://dx.doi.org/10.1016/j.epsr.2015.12.009> <https://linkinghub.elsevier.com/retrieve/pii/S0378779615003867>>. Citado na pág. 28, 58.

ROMERO-TRONCOSO, R. J. et al. Broken rotor bar detection in VSD-fed induction motors at startup by high-resolution spectral analysis. In: *2014 Int. Conf. Electr. Mach.* IEEE, 2014. p. 1848–1854. ISBN 978-1-4799-4389-0. Disponível em: <<http://ieeexplore.ieee.org/document/6960435/>>. Citado na pág. 58, 63.

- SANTOS, C. d. J. *Avaliação do uso de classificadores para verificação de atendimento a critérios de seleção em programas sociais*. 88 p. Tese (Dissertação de Mestrado) — Universidade Federal de Juiz de Fora, 2017. Citado na pág. 54.
- SAUER, I. L. et al. A comparative assessment of Brazilian electric motors performance with minimum efficiency standards. *Renew. Sustain. Energy Rev.*, Elsevier, v. 41, p. 308–318, jan 2015. ISSN 13640321. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2014.08.053> <https://linkinghub.elsevier.com/retrieve/pii/S1364032114007291>>. Citado na pág. 22.
- SETTE, H. S. N. F. Transformada de Hilbert-Fundamentos. In: . [S.l.: s.n.], 2017. p. 31. Citado na pág. 43.
- SILVA, L. R. B. *Classificação de Falhas em Máquinas Elétricas usando Redes Neurais, Modelos Wavelet e Medidas de Informação*. Tese (Dissertação de Mestrado) — Universidade Tecnológica Federal do Paraná, 2014. Citado na pág. 22.
- SOUSA, K. M. et al. Broken Bar Fault Detection in Induction Motor by Using Optical Fiber Strain Sensors. *IEEE Sens. J.*, v. 17, n. 12, p. 3669–3676, jun 2017. ISSN 1530-437X. Disponível em: <<http://ieeexplore.ieee.org/document/7904691/>>. Citado na pág. 59.
- Steven W. Smith. *Digital signal processing*. Second edi. San Diego, California: [s.n.], 1999. v. 2. 688 p. ISBN 0-9660176-7-6. Disponível em: <<http://ieeexplore.ieee.org/document/826412/>>. Citado na pág. 15, 39, 40, 42.
- STILLWELL, J. *Mathematics and Its History*. [S.l.]: Springer Science & Business Media, 2010. 622 p. ISBN 144196052X, 9781441960528. Citado na pág. 42.
- THOMSON, W.; FENGER, M. Case histories of current signature analysis to detect faults in induction motor drives. In: *IEEE Int. Electr. Mach. Drives Conf. 2003. IEMDC'03*. IEEE, 2003. v. 3, p. 1459–1465. ISBN 0-7803-7817-2. Disponível em: <<http://ieeexplore.ieee.org/document/1210644/>>. Citado na pág. 27, 61.
- TRACHI, Y. et al. Induction Machines Fault Detection Based on Subspace Spectral Estimation. *IEEE Trans. Ind. Electron.*, v. 63, n. 9, p. 5641–5651, sep 2016. ISSN 0278-0046. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7473929> <http://ieeexplore.ieee.org/document/7473929/>>. Citado na pág. 58, 74.
- VALLES-NOVO, R. et al. Broken bar detection on squirrel cage induction motors with MCSA and EMD. In: *2014 IEEE Int. Instrum. Meas. Technol. Conf. Proc. IEEE*, 2014. p. 993–998. ISBN 978-1-4673-6386-0. ISSN 10915281. Disponível em: <<http://ieeexplore.ieee.org/document/6860892/>>. Citado na pág. 34, 59.
- WIDODO, A.; YANG, B.-S.; HAN, T. Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. *Expert Syst. Appl.*, v. 32, n. 2, p. 299–312, feb 2007. ISSN 09574174. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0957417405003398>>. Citado na pág. 62, 63.
- XU, B. et al. Improvement of the Hilbert Method via ESPRIT for Detecting Rotor Fault in Induction Motors at Low Slip. *IEEE Trans. Energy Convers.*, v. 28, n. 1, p. 225–233, mar 2013. ISSN 0885-8969. Disponível em: <<http://ieeexplore.ieee.org/document/6407981/>>. Citado na pág. 27, 43, 44, 58, 61, 63, 71.

ZHANG, P. et al. A Survey of Condition Monitoring and Protection Methods for Medium-Voltage Induction Motors. *IEEE Trans. Ind. Appl.*, v. 47, n. 1, p. 34–46, jan 2011. ISSN 0093-9994. Disponível em: <http://ieeexplore.ieee.org/document/5620974/>.

Citado na pág. 27.

ZHU, W.; ZENG, N.; WANG, N. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations. In: *NESUG 2010 Heal. Care Life Sci.* [S.l.: s.n.], 2010. p. 1–9. Citado na pág. 54, 56.

Apêndices

Neste apêndice são apresentados outros resultados oriundos da pesquisa, aplicando janelas amostrais e técnicas de pré-processamento que geraram resultados inferiores aqueles apresentados no capítulo de resultados.

A.1 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO A *FFT* EM UMA JANELA DE 10s

O processo de escolha dos melhores classificadores foi realizado analisando os parâmetros de avaliação descritos anteriormente na tabela 2.4. Neste sentido, a tabela A.1 apresenta os diferentes valores obtidos para os classificadores *KNN*, em função da variação do número dos vizinhos mais próximos (K). Neste caso os dados foram processados em uma janela amostral igual a 10s. Assim, é possível observar os resultados alcançados em termos das medidas de avaliação dos classificadores e assim determinar o mais adequado para cada caso.

Tabela A.1: *Escolha dos melhores parâmetros para o classificador KNN em uma janela de 10s*

K	S	E	Ef	CCM	TA
1	0,732	0,75	0,741	0,482	74,107
3	0,75	0,821	0,785	0,573	78,571
5	0,643	0,892	0,768	0,553	76,785
7	0,643	0,857	0,75	0,512	75
9	0,607	0,857	0,732	0,479	73,214
10	0,518	0,964	0,741	0,539	74,107

A figura A.1 mostra todas as curvas *ROC* geradas para os classificadores *KNN*, com as suas respectivas áreas para cada valor de vizinhos mais próximos.

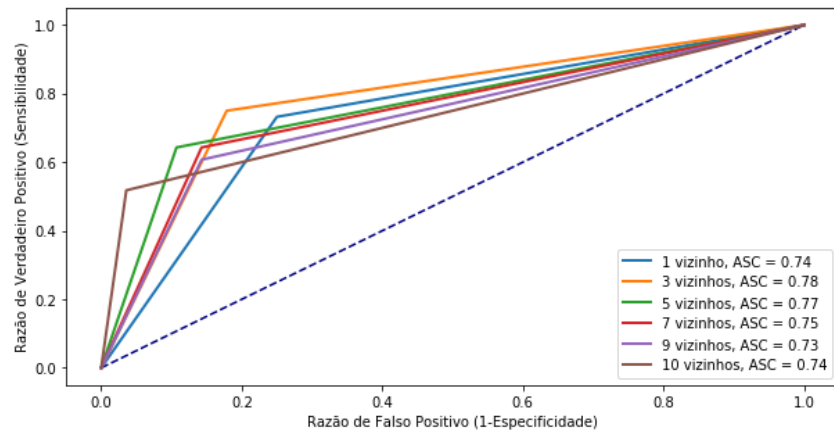


Figura A.1: *Curvas ROC do classificador KNN para diferentes valores de K*

De acordo com a tabela A.1, os melhores resultados foram alcançados com o classificador parametrizado em 3 vizinhos mais próximos. Por sua vez, a tabela A.2 mostra a matriz de confusão gerada para este classificador.

Tabela A.2: *Matriz de confusão gerada para o melhor classificador KNN utilizando a FFT em uma janela de 10s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	10	52
	Negativo	14	46	61
Total		56	56	112

A predição obtida para esse classificador foi fraca, segundo a tabela 2.5, com uma ASC igual a 0,78. O valor de CCM foi igual a 0,57 próximo a uma classificação aleatória e a taxa de acerto ficou em torno de 78,57%. O classificador obteve mais acertos negativos do que positivos com a especificidade em torno de 0,82 e a sensibilidade em 0,75.

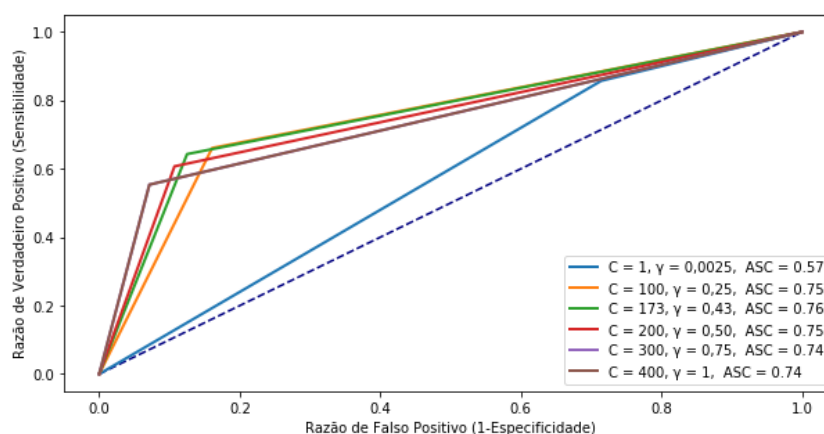
A.1.1 AVALIAÇÃO DO CLASSIFICADOR SVM USANDO A FFT EM UMA JANELA DE 10S

A tabela A.3 apresenta os resultados obtidos para a avaliação do classificador SVM , em função dos parâmetros de C e γ .

Tabela A.3: Escolha dos melhores parâmetros para o classificador SVM em uma janela de 10s

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,857	0,285	0,571	0,174	57,142
100 e 0,25	0,660	0,839	0,75	0,508	75
173 e 0,43	0,643	0,875	0,759	0,532	75,892
200 e 0,50	0,607	0,893	0,75	0,521749195	75
300 e 0,75	0,553	0,928	0,741	0,520	74,107
400 e 1	0,553	0,928	0,741	0,520	74,107

A figura A.2 mostra todas as curvas *ROC* geradas para o classificador *SVM*, em cada configuração dos parâmetros escolhidos.

**Figura A.2:** Curvas *ROC* do classificador SVM para diferentes parâmetros de C e γ

De acordo com a tabela A.3, os melhores resultados das métricas de avaliação foram usados com o classificador *SVM* parametrizado com $C = 173$ e $\gamma = 0,43$. A tabela A.4 mostra a matriz de confusão obtida para o melhor modelo *SVM*.

Tabela A.4: Matriz de confusão do classificador SVM utilizando o FFT em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	36	7	43
	Negativo	20	49	69
Total		56	56	112

A predição encontrada para esse classificador também foi fraca, conforme a tabela 2.5, com uma ASC igual a 0,76. O CCM foi igual a 0,53 próximo a uma classificação aleatória e a taxa de acerto ficou em torno de 75,892%. O classificador neste caso também obteve mais acertos negativos do que positivos, com a especificidade em torno de 0,87 e a sensibilidade em 0,64.

A.1.2 AVALIAÇÃO DO CLASSIFICADOR RL USANDO A FFT EM UMA JANELA DE 10s

A tabela A.5 mostra os resultados obtidos para a RL e neste caso não foi alterado qualquer parâmetro do classificador.

Tabela A.5: Métricas para a análise do classificador RL em uma janela de 10s

S	E	Ef	CCM	TA
0,75	0,75	0,75	0,50	75

A tabela A.6 mostra a matriz de confusão obtida para o classificador RL . A sua curva ROC é mostrada na figura A.3.

Tabela A.6: Matriz de confusão do classificador RL utilizando a FFT como ferramenta de extração das características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	14	56
	Negativo	14	42	56
Total		56	56	112

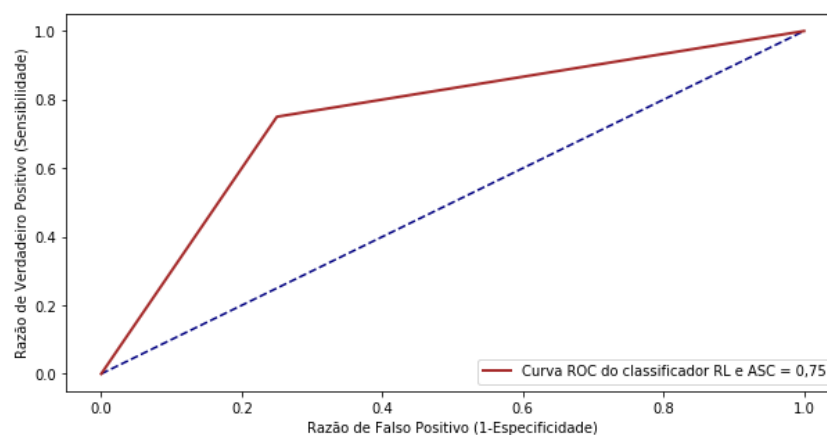


Figura A.3: Curva ROC gerada para a RL

A exemplo dos dois casos anteriores, a previsão para esse classificador também foi fraca com uma ASC igual a 0,75 e um valor de CCM igual a 0,5 próximo a uma classificação aleatória e a taxa de acerto foi de 75%. Neste caso, o classificador obteve o mesmo número de acertos negativos e positivos com a especificidade em torno de 0,75 e a sensibilidade em 0,75.

A.1.3 AVALIAÇÃO DO CLASSIFICADOR KNN USANDO FFT E CARACTERÍSTICAS *Estatísticas* EM UMA JANELA DE 10s

Neste caso foi avaliada a performance do classificador KNN para entradas obtidas nos domínios do tempo e da frequência, aplicando a técnica FFT e as grandezas estatísticas descritas no Capítulo 2. A tabela A.7 apresenta os resultados alcançados para o classificador KNN , para diferentes valores de vizinhos mais próximos, em uma janela amostral igual a 10s

Tabela A.7: Resultados do classificador KNN usando FFT e características *Estatísticas* como entradas em uma janela de 10s

K	S	E	Ef	CCM	TA
1	0,928	0,821	0,875	0,754	87,5
3	0,910	0,857	0,884	0,769	88,393
5	0,875	0,875	0,875	0,75	87,5
7	0,875	0,893	0,884	0,768	88,393
9	0,821	0,910	0,866	0,735	86,607
10	0,803	0,928	0,866	0,738	86,607

A figura A.4 mostra todas as curvas ROC geradas, com as suas respectivas áreas, para cada valor de vizinhos mais próximos.

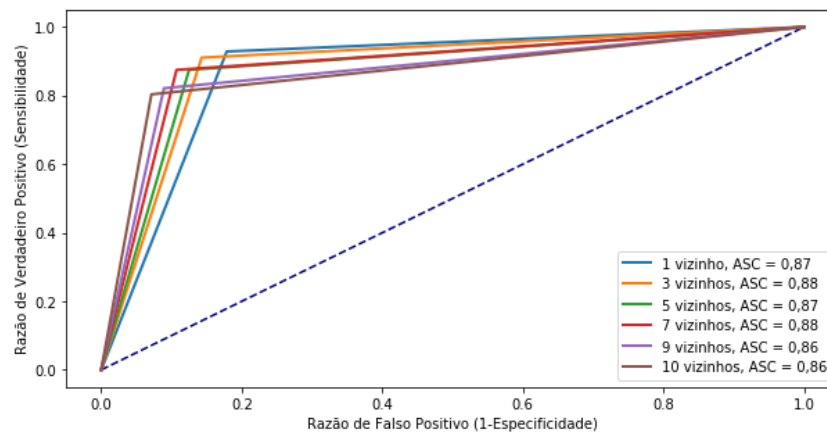


Figura A.4: Curvas ROC geradas para o classificador KNN usando FFT e Estatísticas como entradas em uma janela de 10s

De acordo com a tabela A.7, os melhores resultados foram obtidos para o classificador parametrizado com 3 vizinhos mais próximos.

Tabela A.8: Matriz de confusão do classificador KNN utilizando o FFT e EST como ferramenta de extração de características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	51	8	59
	Negativo	5	48	53
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,88, o CCM foi de 0,769 próximo a uma classificação perfeita e a taxa de acerto foi igual a 88,39%. O classificador obteve mais acertos positivos do que negativos com um sensibilidade em torno de 0,910 e uma especificidade em 0,857.

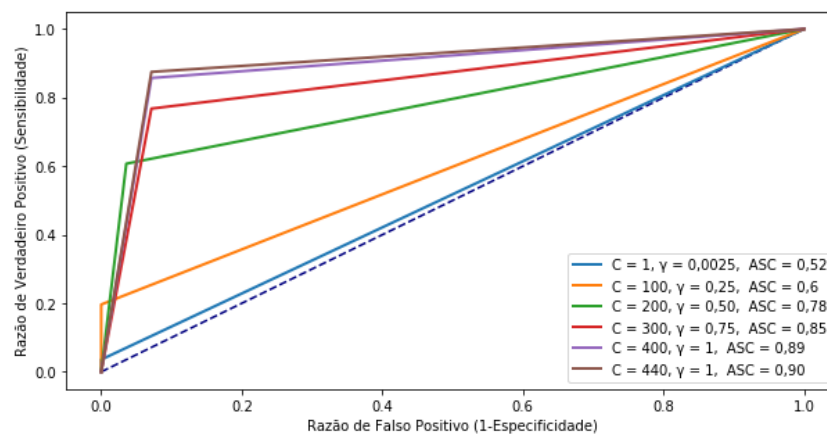
A.1.4 AVALIAÇÃO DO CLASSIFICADOR SVM COM CARACTERÍSTICAS EXTRAÍDAS USANDO A FFT E ESTATÍSTICAS EM UMA JANELA DE 10S

De maneira similar ao caso anterior, a tabela A.9 apresenta as principais métricas encontradas para o classificador SVM com os parâmetros de C e γ .

Tabela A.9: Escolha dos melhores parâmetros para o classificador SVM em uma janela de 10s

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,035	1	0,518	0,134	51,785
100 e 0,25	0,196	1	0,598	0,330	59,821
173 e 0,43	0,607	0,964	0,785	0,611	78,571
200 e 0,50	0,768	0,928	0,848	0,705	84,821
300 e 0,75	0,857	0,928	0,892	0,787	89,28
440 e 1,1	0,875	0,928	0,901	0,804	90,178

A figura A.5 mostra todas as curvas ROC geradas para o classificador SVM, em cada condição de parametrização.

**Figura A.5:** Curvas ROC do classificador SVM usando ambas as características FFT e Estatísticas em uma janela amostral de 10s

De acordo com a tabela A.10, os melhores resultados para o classificador SVM neste caso foram obtidos com $C = 440$ e $\gamma = 1,1$. A tabela A.10 mostra a matriz de confusão do melhor modelo SVM.

Tabela A.10: Matriz de confusão do classificador SVM utilizando FFT e Estatísticas como ferramenta de extração das características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	49	4	53
	Negativo	7	52	59
Total		56	56	112

A previsão para esse classificador foi excelente, segundo a tabela 2.5, com uma *ASC* em torno de 0,90. O valor de *CCM* ficou igual a 0,804727, próximo a uma classificação perfeita. A taxa de acerto foi de 90,17% e o classificador obteve mais acertos negativos do que positivos com um especificidade em torno de 0,928 e uma sensibilidade igual a 0,875.

A.1.5 AVALIAÇÃO DO CLASSIFICADOR *RL* COM CARACTERÍSTICAS EXTRAÍDAS USANDO A *FFT* E *Estatísticas* EM UMA JANELA DE 10s

A proposta aqui foi avaliar mais uma vez o desempenho de uma simples regressão logística com os dados extraídos com a técnica *FFT* e as grandezas *Estatísticas*. A tabela A.11 apresenta os resultados obtidos nesse caso.

Tabela A.11: *Métricas para a análise do classificador RL em uma janela de 10s*

S	E	Ef	CCM	TA
0,75	0,75	0,75	0,5	75

Como nos casos anteriores, foi construída a matriz de confusão para os dados de validação. A tabela A.12 mostra a referida matriz para o classificador *RL* e a sua respectiva curva *ROC* é mostrada na figura A.6.

Tabela A.12: *Matriz de confusão do classificador RL utilizando a FFT e Estatísticas como ferramenta de extração das características em uma janela de 10s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	14	56
	Negativo	14	42	56
Total		56	56	112

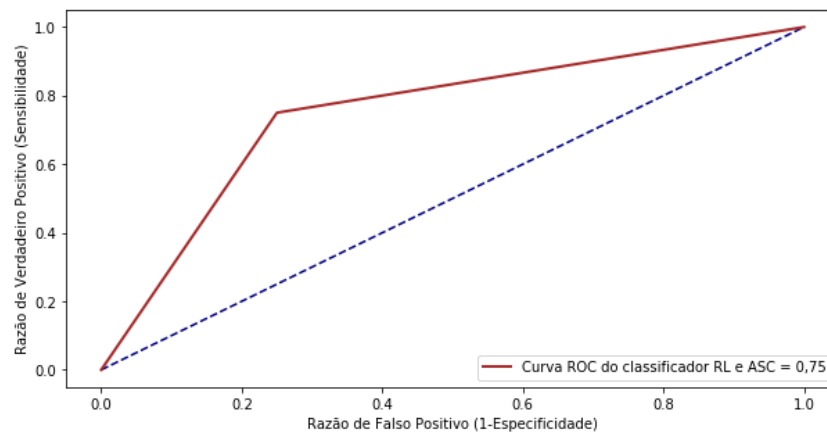


Figura A.6: *Curvas ROC do classificador RL utilizando FFT e Estatísticas*

Nota-se, neste caso, que a previsão para esse classificador foi fraca, como revela os dados da tabela 2.5. Neste sentido, a figura 2.13 mostra a *ASC* em torno de 0,75. O valor de *CCM* foi de 0,5, próximo de uma classificação aleatória e a taxa de acerto foi de 75%. O classificador obteve o mesmo número de acertos negativos e positivos com a especificidade em torno de 0,75 e a sensibilidade em 0,75.

A.1.6 AVALIAÇÃO DO CLASSIFICADOR *KNN* COM CARACTERÍSTICAS EXTRAÍDAS USANDO A *HT* E A *FFT* EM UMA JANELA DE 10S

Nessa fase do estudo foram avaliados os desempenhos dos classificadores para as características extraídas da corrente do motor usando as técnicas *HT* e *FFT*, ainda para uma janela amostral igual a 10s.

Como nos casos anteriores, foram encontrados os melhores parâmetros para cada tipo de classificador. Neste sentido, a tabela A.13 apresenta os resultados obtidos para o classificador *KNN*, em função do número de vizinhos mais próximos.

Tabela A.13: *Escolha dos melhores parâmetros para o classificador KNN em uma janela de 10s*

K	S	E	Ef	CCM	TA %
1	0,857	0,821	0,839	0,679	83,928
3	0,857	0,857	0,857	0,714	85,714
5	0,857	0,892	0,875	0,75	87,5
7	0,857	0,893	0,875	0,75	87,5
9	0,857	0,893	0,875	0,75	87,5
10	0,839	0,893	0,867	0,733	86,607

A figura A.7 mostra uma comparação entre todas as curvas *ROC* geradas e as suas respectivas áreas para cada valor de vizinhos mais próximos.

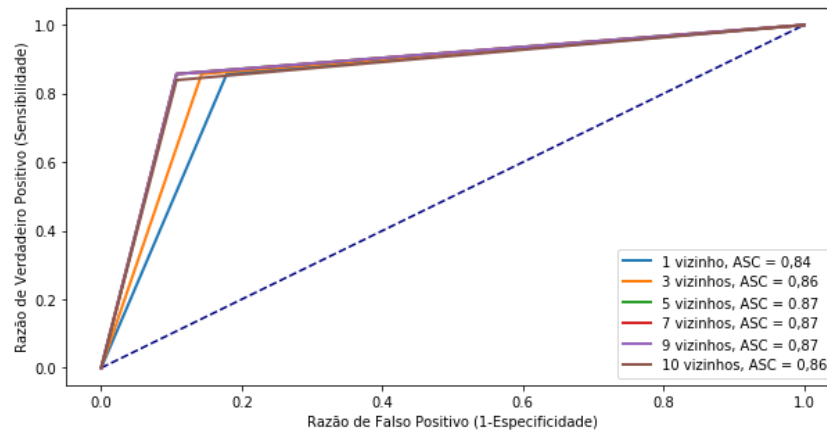


Figura A.7: Curvas *ROC* para o classificador *KNN* usando as técnicas *HT* e a *FFT*

De acordo com a tabela A.13, os melhores resultados das métricas de avaliação foram usados com o classificador parametrizado em 9 vizinhos mais próximos. A tabela A.14 apresenta a matriz de confusão para o melhor classificador.

Tabela A.14: Matriz de confusão do classificador *KNN* utilizando o *HT* e *FFT* como ferramenta de extração de características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	48	6	54
	Negativo	8	50	58
Total		56	56	112

A previsão para esse classificador foi boa, segundo os dados descritos na tabela 2.5, com uma *ASC* em torno de 0,87. Neste caso, o valor de *CCM* foi de 0,767, mais próximo, portanto, de uma classificação perfeita. A taxa de acerto foi de 87,5%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,982 e uma sensibilidade em 0,767.

A.1.7 AVALIAÇÃO DO CLASSIFICADOR *SVM* COM CARACTERÍSTICAS EXTRAÍDAS USANDO *HT* E *FFT* EM UMA JANELA DE 10s

Nesta fase foi novamente avaliado o desempenho de um classificador *SVM*, tendo como entradas as características extraídas com as técnicas combinadas *HT* e *FFT*. A tabela A.15

apresenta as principais métricas utilizadas para a avaliação do classificador *SVM*, com os parâmetros de ajuste como C e γ .

Tabela A.15: Escolha dos melhores parâmetros para o classificador *SVM* em uma janela de 10s

C e γ	S	E	Ef	CCM	TA
1 e 0,43	0,768	0,982	0,875	0,768	87,5
1 e 0,0025	0,964	0,321	0,643	0,373	64,285
100 e 0,25	0,839	0,910	0,875	0,752	87,5
200 e 0,50	0,785	0,965	0,875	0,762	87,5
300 e 0,75	0,785	0,964	0,875	0,762	87,5
400 e 1	0,768	0,981	0,875	0,767	87,5

A figura A.8 mostra todas as curvas *ROC* geradas para o classificador *SVM*, considerando os diferentes valores de C e γ .

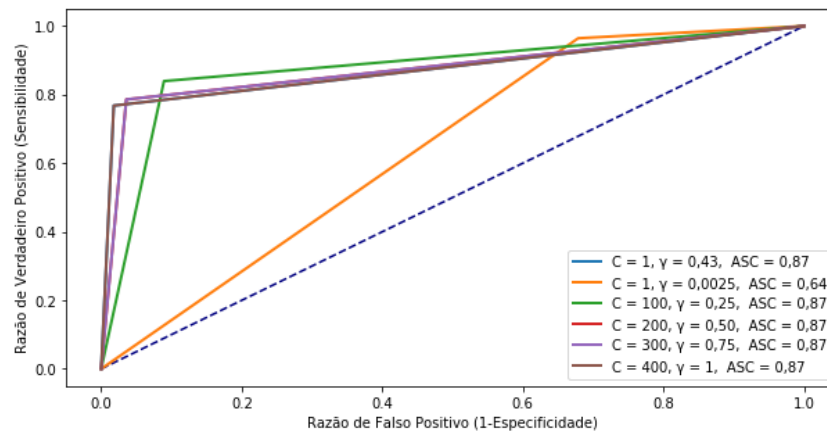


Figura A.8: Curvas *ROC* do classificador *SVM*

De acordo com a tabela A.15, os melhores resultados foram alcançados com o classificador *SVM* parametrizado com $C = 1$ e $\gamma = 0,43$. A tabela A.16 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.16: Matriz de confusão do classificador *KNN* utilizando *HT* e *FFT* como ferramenta de extração das características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	43	1	44
	Negativo	13	55	68
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,87; o CCM foi de 0,768, próximo a uma classificação perfeita. A taxa de acerto foi de 87,5%. O classificador obteve mais acertos negativos do que positivos nesse caso, com um especificidade em torno de 0,982 e uma sensibilidade igual a 0,767.

A.1.8 AVALIAÇÃO DO CLASSIFICADOR RL COM CARACTERÍSTICAS EXTRAÍDAS USANDO HT E FFT EM UMA JANELA DE 10s

Nesta fase foi novamente avaliado o desempenho da regressão logística, mas considerando como entradas as características da corrente extraídas com as técnicas HT e FFT . A tabela A.17 mostra as métricas de análise do classificador RL obtidas nesse caso.

Tabela A.17: Métricas para a análise do classificador RL em uma janela de 10s

S	E	Ef	CCM	TA
0,714	0,946	0,830	0,679	83,035

A tabela A.18 mostra a matriz de confusão do classificador RL , com a sua respectiva curva ROC como mostra a figura A.9.

Tabela A.18: Matriz de confusão do classificador RL utilizando HT e FFT como ferramenta de extração das características em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	40	3	43
	Negative	16	53	69
Total		56	56	112

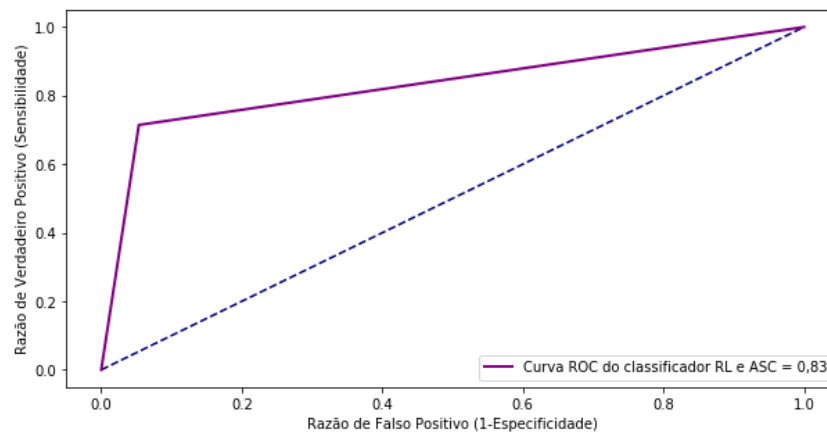


Figura A.9: *Curvas ROC do classificador RL utilizando HT e FFT*

A previsão para esse classificador foi boa segundo a tabela 2.5. A figura A.9 mostra uma *ASC* em torno de 0,83 e um valor de *CCM* igual a 0,679, entre uma classificação aleatória e perfeita, além de alcançar uma taxa de acerto igual a 83%. Neste caso, o classificador obteve mais acertos negativos do que acertos positivos, com a especificidade em torno de 0,94 e a sensibilidade em 0,71.

A.1.9 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO AS GRANDEZAS *Estatísticas* E COMBINADAS COM AQUELAS EXTRAÍDAS POR MEIO DA TÉCNICA *HT* E *FFT* EM UMA JANELA DE 10S

Nesta fase foram extraídas da corrente de fase do motor as grandezas estatísticas, conforme definidas no Capítulo 2, e aquelas oriundas da aplicação das técnicas conjuntas *HT* e *FFT*, em uma janela amostral igual a 10s. A tabela A.19 apresenta os valores encontrados para cada parâmetro de avaliação do classificador *KNN*, em função da variação no número de vizinhos mais próximos.

Tabela A.19: *Escolha dos melhores parâmetros para o classificador KNN em uma janela de 10s*

K	S	E	Ef	CCM	TA %
1	0,875	0,857	0,866	0,732	86,607
3	0,892	0,893	0,893	0,785	89,285
5	0,875	0,875	0,875	0,75	87,5
7	0,893	0,875	0,884	0,768	88,393
9	0,857	0,910	0,884	0,769	88,392
10	0,821	0,929	0,875	0,754	87,5

A figura A.10 mostra todas as curvas *ROC* geradas com as suas respectivas áreas.

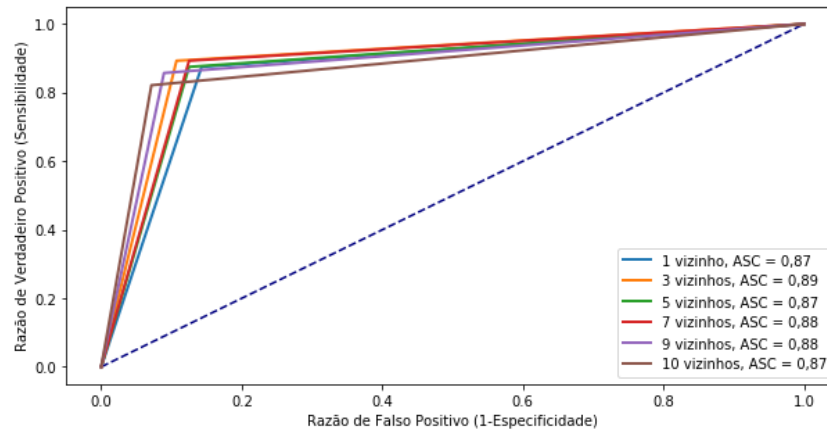


Figura A.10: *Curvas ROC do classificador KNN*

Mais uma vez, tomando por base a tabela A.19, os melhores resultados foram encontrados com o classificador parametrizado em 3 vizinhos mais próximos. A tabela A.20 mostra a matriz de confusão deste modelo *KNN*.

Tabela A.20: *Matriz de confusão do classificador KNN utilizando HT e FFT e Estatísticas como ferramenta de extração de características em uma janela de 10s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	50	6	56
	Negativo	6	50	56
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma *ASC* em torno de 0,89. O valor obtido para o *CCM* foi de 0,785, logo, mais perto de uma classificação perfeita e a taxa de acerto foi de 89,28%. O classificador obteve a mesma quantidade de acertos negativos e positivos com o valor em torno de 0,892.

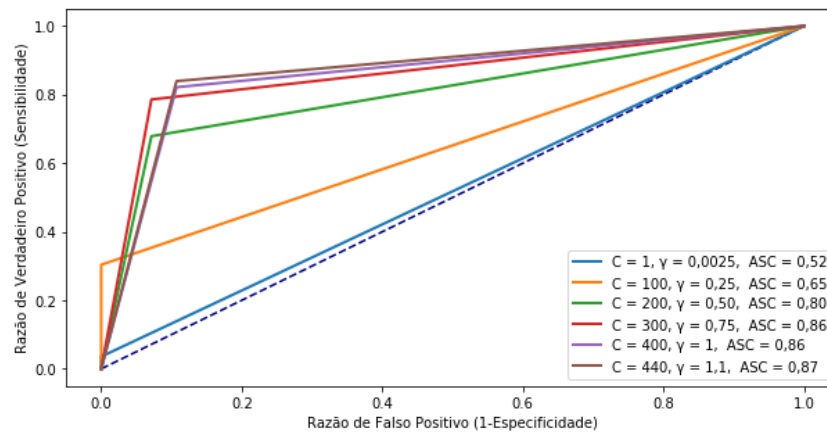
A.1.10 AVALIAÇÃO DO CLASSIFICADOR *SVM* USANDO AS GRANDEZAS *Estatísticas* E AQUELAS EXTRAÍDAS COM AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 10S

A exemplo do caso anterior, nesta fase foram avaliadas as mesmas características de entrada, mas considerando o uso de um classificador do tipo *SVM*. A tabela A.21 apresenta os resultados obtidos para cada combinação de parâmetros C e γ , a fim de encontrar o melhor classificador.

Tabela A.21: Escolha dos melhores parâmetros para o classificador SVM em uma janela de 10s

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,035	1	0,518	0,134	51,785
100 e 0,25	0,303	1	0,652	0,423	65,178
200 e 0,50	0,679	0,929	0,804	0,627	80,357
300 e 0,75	0,785	0,929	0,857	0,722	85,714
400 e 1	0,821	0,893	0,857	0,716	85,714
440 e 1,1	0,839	0,893	0,866	0,733	86,607

A figura A.11 ilustra todas as curvas ROC geradas para o classificador SVM, em cada caso de parametrização.

**Figura A.11:** Curvas ROC para o classificador SVM usando diferentes configurações de parâmetros

De acordo com a tabela A.22, os melhores resultados foram alcançados com o classificador SVM configurado com o valor de $C = 440$ e o valor de $\gamma = 1,1$. A tabela A.16 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.22: Matriz de confusão para o classificador SVM utilizando as entradas processadas com as grandezas Estatísticas e com as técnicas HT e FFT, em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	47	6	53
	Negativo	9	50	59
Total		56	56	112

A previsão alcançada para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,87. O valor de CCM foi de 0,733, próximo a uma classificação perfeita. A taxa de acerto foi de 86,6%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,893 e uma sensibilidade em 0,839. Vale destacar que esse classificador obteve um desempenho inferior ao anterior, ou seja, com o uso do KNN . Ademais, nota-se que a variação dos parâmetros C e γ provocou neste último caso uma alteração importante na performance do classificador, diferentemente do observado para o classificador KNN .

A.1.11 AVALIAÇÃO DO CLASSIFICADOR RL USANDO AS GRANDEZAS *Estatísticas* E AQUELAS EXTRAÍDAS COM AS TÉCNICAS HT E FFT EM UMA JANELA DE 10s

Nesta fase foi avaliado novamente o desempenho de uma regressão logística para os mesmos dados de entrada processados nos dois casos anteriores, ou seja, com o uso dos modelos KNN e SVM . A tabela A.17 mostra as métricas obtidas para o classificador RL .

Tabela A.23: Métricas para a análise do classificador RL em uma janela de 10s

S	E	Ef	CCM	TA
0,821	0,910	0,867	0,735	86,607

A tabela A.18 mostra a matriz de confusão gerada nesse caso e a figura A.12 ilustra a sua respectiva curva ROC .

Tabela A.24: Matriz de confusão para o classificador RL utilizando as entradas processadas com as grandezas *Estatísticas* e com as técnicas HT e FFT , em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	46	5	51
	Negativo	10	51	61
Total		56	56	112

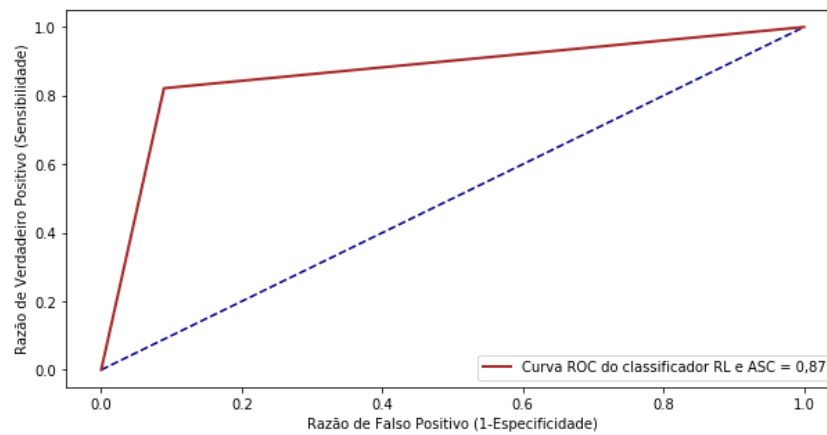


Figura A.12: Curva ROC obtida para o classificador *RL*, usando as entradas processadas com as grandezas Estatísticas e as técnicas HT e FFT

A previsão obtida para esse classificador foi boa, novamente segundo a tabela 2.5. A figura A.12 mostra a *ASC* em torno de 0,87, para um valor de *CCM* próximo a 0,735, logo, mais próxima de uma classificação perfeita. A taxa de acerto foi de 86,6%. O classificador obteve mais acertos negativos do que acertos positivos nesse caso, com a especificidade em torno de 0,91 e a sensibilidade em 0,82. Nota-se que o uso do classificador *RL* alcançou uma *ASC* em torno de 0,87, portanto, igual ao valor obtido para o classificador *SVM*, e ligeiramente inferior ao valor gerado para o classificador *KNN*, que obteve uma *ASC* igual a 0,89.

A.1.12 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO APENAS AS GRANDEZAS Estatísticas COMO ENTRADAS EM UMA JANELA DE 10s

Nos próximos três casos foram avaliados os desempenhos dos mesmos classificadores escolhidos para o presente trabalho, mas usando como entradas apenas as grandezas estatísticas descritas no Capítulo 2, ainda para uma janela amostral igual a 10s. Logo, foram avaliados nesses casos apenas os dados processados no domínio do tempo. A tabela A.25 apresenta os diferentes resultados obtidos com o classificador *KNN*, em função da variação do número de vizinhos mais próximos.

Tabela A.25: Resultados obtidos com o classificador KNN, em uma janela de 10s, para diferentes valores de K

K	S	E	Ef	CCM	TA %
1	0,893	0,839	0,857	0,716	85,714
3	0,857	0,893	0,848	0,696	84,82
5	0,839	0,803	0,821	0,643	82,143
7	0,857	0,803	0,83	0,66	83,035
9	0,839	0,857	0,848	0,696	84,821
10	0,785	0,893	0,839	0,682	83,928

A figura A.13 mostra todas as curvas *ROC* geradas, com as suas respectivas áreas para cada valor de vizinhos mais próximos.

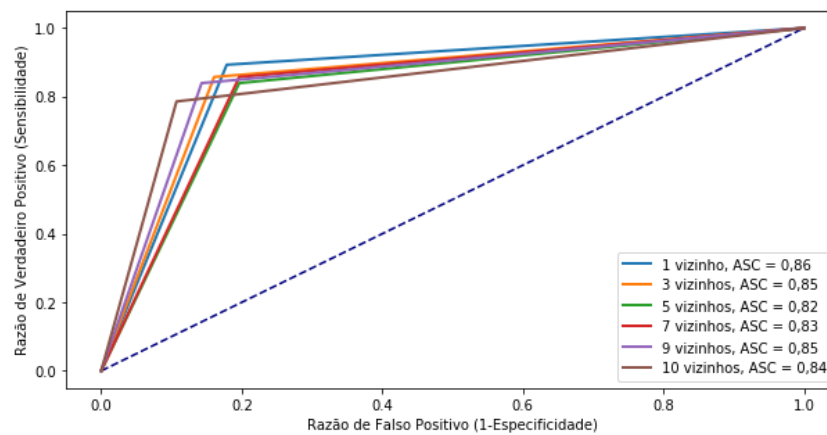


Figura A.13: Curvas *ROC* geradas com o classificador KNN usando apenas entradas no domínio do tempo

De acordo com a tabela A.25, os melhores resultados foram obtidos com o classificador configurado com apenas 1 vizinho mais próximo. A tabela A.26 mostra a matriz de confusão para este classificador.

Tabela A.26: Matriz de confusão do classificador KNN utilizando características Estatísticas como entradas em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	50	10	60
	Negativo	6	46	52
Total		56	56	112

A previsão obtida para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,86. O valor de CCM nesse caso foi de 0,716, assim mais perto de uma classificação perfeita. A taxa de acerto foi de 85,714%. O classificador obteve uma maior quantidade de acertos positivos do que negativos, com o valor de sensibilidade em torno de 0,893 e a especificidade em 0,821. As duas próximas seções apresentam os valores obtidos para os classificadores SVM e RL , nas mesmas condições de entrada e janela amostral.

A.1.13 AVALIAÇÃO DO CLASSIFICADOR SVM USANDO APENAS AS GRANDEZAS *Estatísticas* COMO ENTRADAS EM UMA JANELA DE 10s

O classificador SVM foi testado aqui também apenas para as entradas extraídas no domínio do tempo, ou seja, com o uso das grandezas *Estatísticas*, na janela amostral em 10s. A tabela A.27 apresenta os resultados alcançados com o uso do classificador SVM , dada a variação dos parâmetros C e γ .

Tabela A.27: Avaliação dos melhores parâmetros para o classificador SVM usando apenas as grandezas *Estatísticas* como entradas

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,035	1	0,518	0,134	51,785
100 e 0,25	0,339	0,982	0,660	0,419	66,0714
200 e 0,50	0,714	0,910	0,8125	0,637	81,25
300 e 0,75	0,803	0,892	0,848	0,699	84,821
400 e 1	0,875	0,839	0,857	0,714	85,714
440 e 1,1	0,857	0,821	0,839	0,679	83,928

A figura A.14 mostra todas as curvas ROC geradas para o classificador SVM em cada caso.

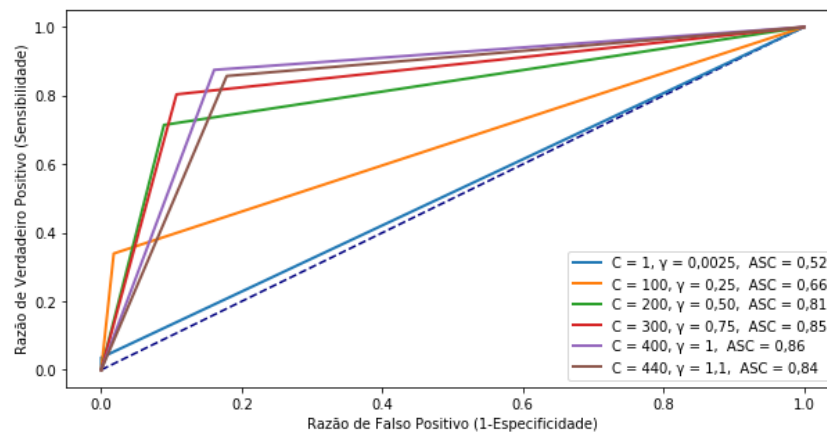


Figura A.14: Curvas ROC geradas para o classificador SVM usando apenas as grandezas Estatísticas como entradas

De acordo com a tabela A.28, os melhores resultados das métricas de avaliação foram obtidos com o classificador SVM parametrizado com $C = 400$ e $\gamma = 1$. A tabela A.28 mostra a matriz de confusão obtida com a classificação testada para o melhor modelo.

Tabela A.28: Matriz de confusão do classificador SVM com características Estatísticas em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	49	9	58
	Negativo	7	47	54
Total		56	56	112

A previsão para esse classificador foi boa, mais uma vez segundo a tabela 2.5 e com uma ASC em torno de 0,86 e o valor CCM foi de 0,733, logo, próximo a uma classificação perfeita. A taxa de acerto foi de 86,6%. O classificador obteve mais acertos positivos do que negativos nessa condição, com uma especificidade em torno de 0,839 e uma sensibilidade igual a 0,875.

A.1.14 AVALIAÇÃO DO CLASSIFICADOR RL USANDO APENAS AS GRANDEZAS Estatísticas COMO ENTRADAS EM UMA JANELA DE 10S

A regressão logística foi novamente testada, considerando como entradas apenas as grandezas extraídas da corrente do motor no domínio do tempo, em uma janela de 10s. A tabela A.29 apresenta os resultados obtidos para avaliação do classificador RL.

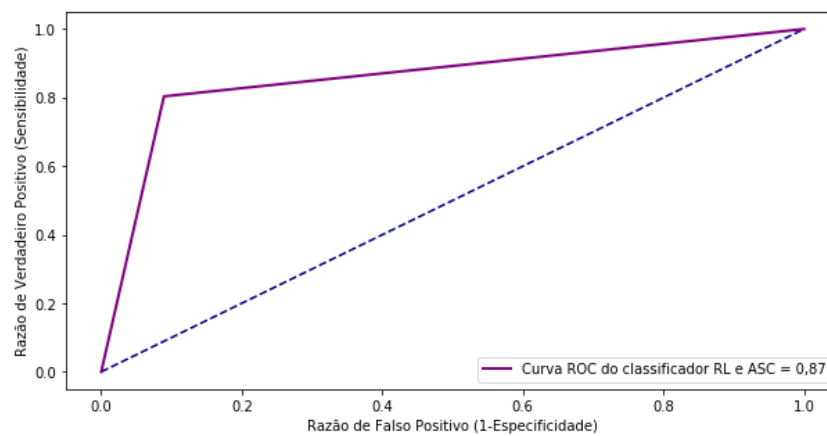
Tabela A.29: Métricas para a análise do classificador *RL* em uma janela de 10s

S	E	Ef	CCM	TA
0,803	0,910	0,857	0,718	85,714

A tabela A.30 mostra a matriz de confusão obtida durante os testes do classificador *RL* e a figura A.15 ilustra a sua respectiva curva *ROC*.

Tabela A.30: Matriz de confusão do classificador *RL* com características Estatísticas em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	49	9	58
	Negative	7	47	54
Total		56	56	112

**Figura A.15:** Curva *ROC* do classificador *RL* utilizando entradas no domínio do tempo

A previsão para esse classificador foi boa, segundo a tabela 2.5. A figura A.15 mostra a *ASC* em torno de 0,87, para um valor de *CCM* em torno de 0,714, portanto, mais próximo de uma classificação perfeita. A taxa de acerto foi de 85,71%. O classificador obteve mais acertos positivos do que acertos negativos nesse caso, com a especificidade em torno de 0,839 e a sensibilidade igual a 0,875. Nota-se, portanto, que o classificador *RL* obteve um desempenho um pouco superior ($ASC = 0,87$) aquele observado para os outros dois modelos ($ASC = 0,86$), tomando por base as entradas extraídas apenas no domínio do tempo, ou seja, sem o uso das técnicas *HT* e/ou *FFT*.

A.1.15 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO A *FFT* EM UMA JANELA DE 5s

Todos os casos avaliados anteriormente, isto é, para diferentes entradas e com os três classificadores testados, foram definidos para uma janela amostral da corrente igual a 10s. Como a proposta deste trabalho é comparar ainda o desempenho das diferentes configurações em um volume diferente de amostras discretizadas, a seguir são apresentados e discutidos os resultados obtidos para uma janela amostral igual a 5s, logo, com um número de amostras reduzido à metade em relação aos casos avaliados anteriormente. Assim, a tabela A.31 apresenta os resultados obtidos com o uso do classificador *KNN*, para diferentes valores de vizinhos mais próximos, tomando como entradas as características extraídas apenas com a técnica *FFT*, em uma janela amostral igual a 5s. A figura A.16 mostra todas as curvas *ROC* geradas nesse caso, com as suas respectivas áreas para cada valor de vizinhos mais próximos.

Tabela A.31: Escolha dos melhores parâmetros para o classificador *KNN*, usando apenas entradas extraídas com a *FFT* em uma janela de 5s

K	S	E	Ef	CCM	TA
1	0,714	0,643	0,678	0,358	67,857
3	0,607	0,768	0,687	0,379	68,75
5	0,571	0,839	0,705	0,426	70,535
7	0,625	0,821	0,723	0,455	72,321
9	0,589	0,839	0,714	0,442	71,428
10	0,518	0,928	0,723	0,489	72,321

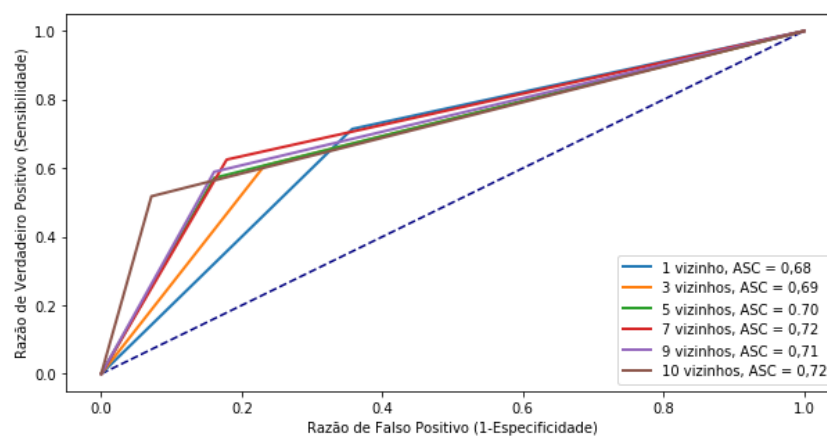


Figura A.16: Curvas *ROC* geradas com o classificador *KNN*, em uma janela de 5s, usando apenas entradas extraídas com a *FFT*

De acordo com a tabela A.31, os melhores resultados das métricas de avaliação foram alcançados com o classificador parametrizado com 10 vizinhos mais próximos. A tabela A.32 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.32: *Matriz de confusão do classificador KNN utilizando o FFT como ferramenta de extração das características em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	29	4	33
	Negativo	27	52	79
Total		56	56	112

A previsão para esse classificador foi fraca, segundo o que determina a tabela 2.5, com uma *ASC* de 0,72. O valor de *CCM* foi de 0,48, portanto está mais próximo de uma classificação aleatória. A taxa de acerto foi de 72,32%. O classificador obteve mais acertos negativos do que positivos, com a especificidade em torno de 0,928 e a sensibilidade em 0,518.

A.1.16 AVALIAÇÃO DO CLASSIFICADOR *SVM* USANDO A *FFT* EM UMA JANELA DE 5S

Neste caso foi avaliado o desempenho do classificador *SVM*, para as mesmas entradas do caso anterior e também na mesma janela amostral. A tabela A.33 apresenta os diferentes valores encontrados para o classificador *SVM*, em função da variação dos parâmetros *C* e γ .

Tabela A.33: *Escolha dos melhores parâmetros para o classificador SVM em uma janela de 5s*

C e γ	S	E	Ef	CCM	TA
1 e 0,43	0,518	0,892	0,705	0,443	70,535
1 e 0,0025	0,7143	0,518	0,616	0,236	61,607
100 e 0,25	0,589	0,928	0,759	0,550	75,892
200 e 0,50	0,589	0,839	0,714	0,443	71,428
300 e 0,75	0,607	0,803	0,705	0,419	70,535
400 e 1	0,625	0,839	0,732	0,475	73,214

A figura A.17 mostra as curvas *ROC* geradas com o classificador *SVM* para cada configuração.

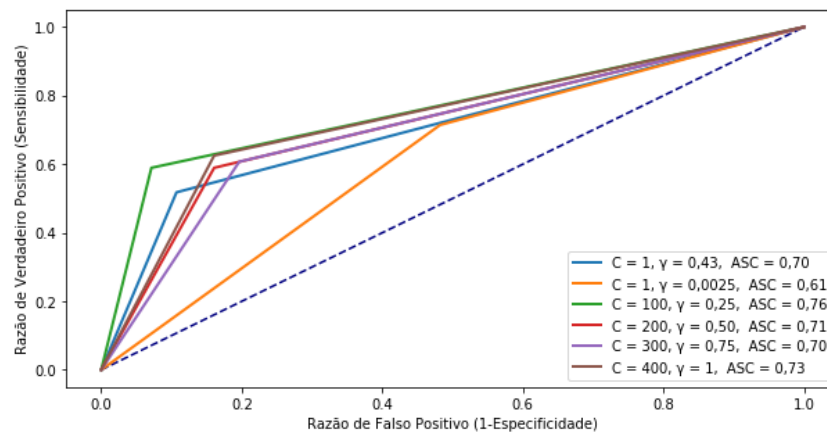


Figura A.17: Curvas ROC do classificador SVM utilizando as entradas extraídas com a FFT, em uma janela de 5s

De acordo com a tabela A.33, os melhores resultados foram observados com o classificador SVM parametrizado com $C = 100$ e $\gamma = 0,25$. A tabela A.34 apresenta a matriz de confusão para o melhor modelo escolhido.

Tabela A.34: Matriz de confusão do classificador SVM utilizando o FFT como ferramenta de extração das características em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	33	4	37
	Negativo	23	52	75
Total		56	56	112

A previsão para esse classificador foi fraca, segundo a tabela 2.5, com uma ASC igual a 0,76. O valor de CCM foi de 0,55, portanto, entre uma classificação aleatória e perfeita. A taxa de acerto foi de 75,893%. O classificador obteve mais acertos negativos do que positivos, com a especificidade em torno de 0,928 e a sensibilidade igual a 0,589.

A.1.17 AVALIAÇÃO DO CLASSIFICADOR RL USANDO A FFT EM UMA JANELA DE 5S

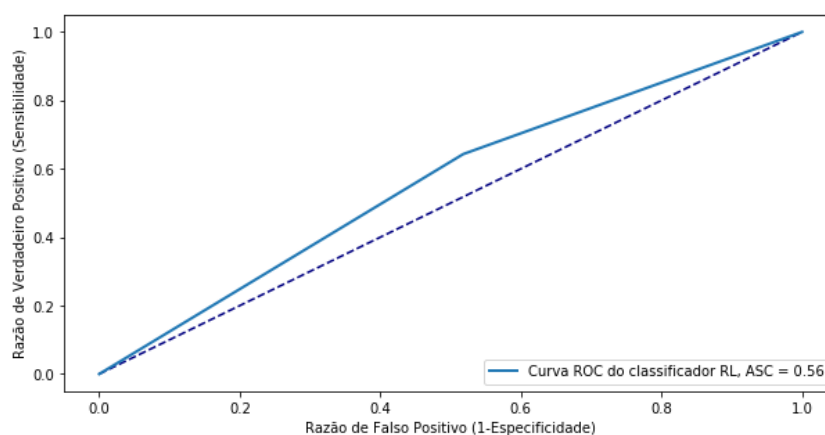
A regressão logística também foi testada para as mesmas condições avaliadas nos dois casos anteriores. A seguir são apresentadas as métricas de avaliação do classificador RL (tabela A.35), bem como a sua matriz de confusão (tabela A.36) e a curva ROC gerada (figura A.18).

Tabela A.35: Métricas para a análise do classificador RL em uma janela de 5s

S	E	Ef	CCM	TA
0,643	0,482	0,562	0,127	56,25%

Tabela A.36: Matriz de confusão do classificador RL utilizando a FFT em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	36	29	65
	Negativo	20	27	47
Total		56	56	112

**Figura A.18:** Curva ROC do classificador RL utilizando FFT em uma janela de 5s

Os resultados obtidos para esse classificador, ou seja, a sua previsão, foi aleatória segundo a tabela 2.5. O classificador obteve uma *ASC* igual a 0,56 e um valor de *CCM* de 0,127, logo, próxima a uma classificação aleatória. A taxa de acerto foi de 56,2%. O classificador obteve mais acertos positivos do que negativos, com a sensibilidade em torno de 0,643 e a especificidade em 0,482. Dos três classificadores testados, para as mesmas entradas e na mesma janela amostral, é possível notar que o classificador *SVM* alcançou a melhor performance, tomando por base o valor obtido para a *ASC*, no caso em torno de 0,76 na melhor configuração.

A.1.18 AVALIAÇÃO DO CLASSIFICADOR *KNN* COM CARACTERÍSTICAS EXTRAÍDAS USANDO A *FFT* E *Estatísticas* EM UMA JANELA DE 5s

A exemplo dos testes realizados com uma janela de 10s, foi avaliado também o desempenho dos classificadores com a combinação de entradas nos domínios do tempo e da frequência, para uma janela amostral igual a 5s. Neste sentido, a seguir são apresentados os resultados obtidos para as características extraídas usando a técnica *FFT* e as grandezas estatísticas descritas no Capítulo 2, em uma janela de tempo de 5s. A tabela A.37 apresenta os resultados da classificação obtidos com os modelos *KNN*, em função da variação no número de vizinhos mais próximos.

Tabela A.37: Escolha dos melhores parâmetros para o classificador *KNN* em uma janela de 5s

K	S	E	Ef	CCM	TA
1	0,803	0,857	0,830	0,661	83,03
3	0,785	0,928	0,857	0,722	85,714
5	0,75	0,928	0,839	0,689	83,928
7	0,7143	0,946	0,830	0,679	83,035
9	0,7143	0,964	0,839	0,70	83,928
10	0,696	0,982	0,839	0,70	83,928

A figura A.19 mostra todas as curvas *ROC* geradas, com as suas respectivas áreas, para cada valor de vizinhos mais próximos.

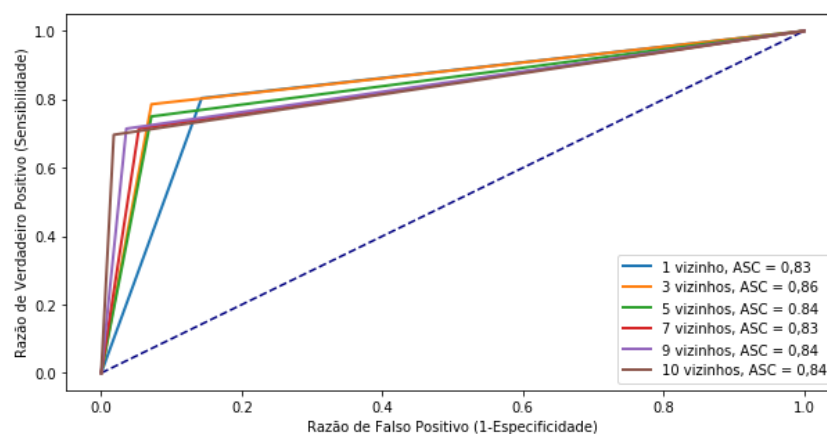


Figura A.19: Curvas *ROC* obtidas para o classificador *KNN*, em uma janela de 5s, usando as técnicas *FFT* e *Estatísticas*

Tomando por base os resultados da tabela A.37, os melhores resultados foram obtidos com o classificador configurado com 3 vizinhos mais próximos. A tabela A.38 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.38: *Matriz de confusão do classificador KNN utilizando o FFT e Estatísticas como ferramenta de extração de características em uma janela de 10s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	44	4	48
	Negativo	12	52	64
Total		56	56	112

Conforme a tabela 2.5, a previsão obtida para esse classificador foi boa, com uma ASC em torno de 0,86. O valor de CCM foi de 0,722, mais perto, portanto, de uma classificação perfeita do que aleatória. A taxa de acerto foi de 85,7142%. O classificador obteve mais acertos negativos do que positivos nesse caso, com uma sensibilidade em torno de 0,928 e uma especificidade em 0,785.

A.1.19 AVALIAÇÃO DO CLASSIFICADOR SVM COM CARACTERÍSTICAS EXTRAÍDAS USANDO A FFT E $Estatísticas$ EM UMA JANELA DE 5s

Neste caso o classificador SVM foi testado para as mesmas condições anteriores avaliadas com o uso do KNN . A tabela A.39 apresenta os principais resultados avaliados para o classificador SVM , considerando a variação dos parâmetros C e γ .

Tabela A.39: *Escolha dos melhores parâmetros para o classificador SVM em uma janela de 5s*

C e γ	S	E	Ef	CCM	TA
1 e 4,6	0,696	1	0,848	0,730	84,821
1 e 0,0025	0,035	1	0,518	0,134	51,785
100 e 0,25	0,125	0,964	0,544	0,164	54,464
200 e 0,50	0,714	0,910	0,8125	0,637	81,25
300 e 0,75	0,732	0,893	0,8125	0,633	81,25
400 e 1,1	0,803	0,893	0,848	0,699	84,821

A figura A.20 mostra as curvas ROC geradas do classificador SVM para cada conjunto de parametrização.

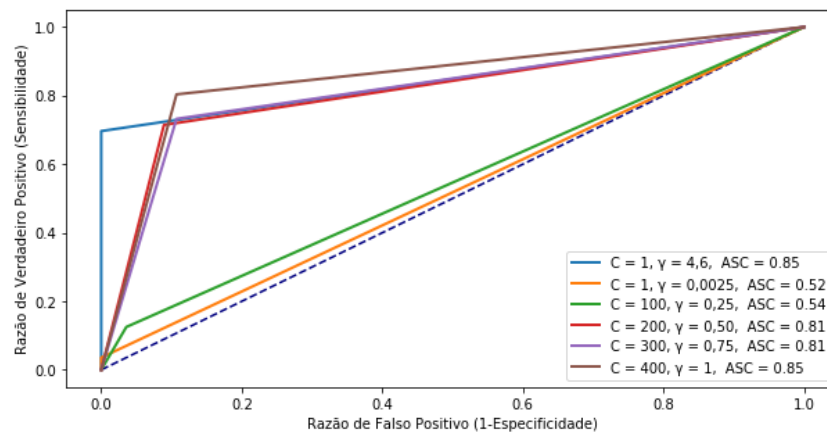


Figura A.20: Curvas ROC do classificador SVM em uma janela de 5s, usando entradas extraídas com FFT e Estatísticas

De acordo com a tabela A.39, pode-se notar que os melhores resultados foram obtidos com o classificador SVM configurado com os valores $C = 1$ e $\gamma = 4,6$. A tabela A.40 mostra a matriz de confusão gerada para o melhor classificador.

Tabela A.40: Matriz de confusão do classificador SVM utilizando o FFT e Estatísticas como ferramenta de extração das características em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	0	39
	Negativo	17	56	73
Total		56	56	112

A previsão para esse classificador foi excelente, segundo a tabela 2.5, com uma ASC em torno de 0,85, com um valor de CCM igual a 0,73, próximo a uma classificação perfeita. A taxa de acerto foi de 84,821%. O classificador acertou todos os valores da classe negativa, com uma especificidade igual a 1 e uma sensibilidade de 0,696.

A.1.20 AVALIAÇÃO DO CLASSIFICADOR RL COM CARACTERÍSTICAS EXTRAÍDAS USANDO A FFT E Estatísticas EM UMA JANELA DE 5S

Nesta fase foi testada a regressão logística, com as mesmas entradas e a mesma janela amostral dos casos avaliados com os classificadores KNN e SVM. A tabela A.41 mostra as métricas testadas para análise de desempenho do classificador RL, com características extraídas utilizando as entradas FFT e EST.

Tabela A.41: Resultados do classificador *RL* em uma janela de 5s, usando entradas geradas a partir da *FFT* e *EST*

S	E	Ef	CCM	TA
0,697	0,928	0,812	0,642	81,25

A tabela A.42 mostra a matriz de confusão do classificador *RL* e a figura A.21 ilustra a sua respectiva curva *ROC*.

Tabela A.42: Matriz de confusão do classificador *RL* utilizando a *FFT* e as grandezas estatísticas como entradas em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	4	43
	Negativo	17	52	69
Total		56	56	112

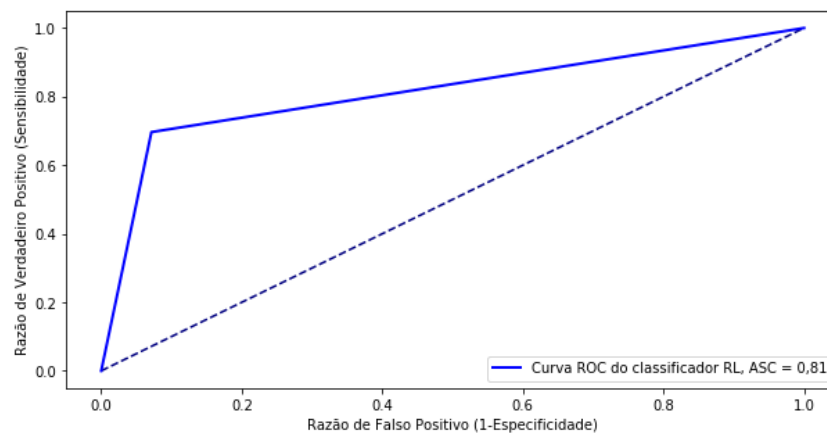


Figura A.21: Curva *ROC* do classificador *RL* utilizando *FFT* e Estatísticas

A previsão para esse classificador foi boa, ainda segundo a tabela 2.5. A figura A.21 mostra a *ASC* obtida, no caso em torno de 0,81, para um valor de *CCM* em torno de 0,642, assim, entre uma classificação aleatória e perfeita. A taxa de acerto foi de 81,25%. O classificador obteve mais acertos negativos do que acertos positivos, com uma especificidade em torno de 0,928 e uma sensibilidade em 0,696. Vale destacar que, para as mesmas condições de entrada, o classificador *KNN* obteve o melhor desempenho usando como referência o valor de *ASC* (*ASC* igual a 0,86), na sua melhor configuração com 3 vizinhos mais próximos.

A.1.21 AVALIAÇÃO DO CLASSIFICADOR *KNN* COM CARACTERÍSTICAS EXTRAÍDAS USANDO *HT* E *FFT* EM UMA JANELA DE 5s

Aqui os mesmos classificadores foram testados, para a mesma janela amostral, mas considerando como entradas as características da corrente do motor extraídas com as técnicas combinadas *HT* e *FFT*. A tabela A.43 apresenta os resultados obtidos para a avaliação do classificador *KNN*, com os diferentes valores de vizinhos mais próximos.

Tabela A.43: Resultados da classificação gerada com o *KNN*, em uma janela de 5s, para diferentes valores de *K*

K	S	E	Ef	CCM	TA %
1	0,785	0,875	0,830	0,663	83,035
3	0,75	0,9107	0,830	0,669	83,035
5	0,732142857	0,928	0,830	0,674	83,035
7	0,75	0,892	0,821	0,649	82,143
9	0,75	0,946	0,848	0,710	84,821
10	0,732142857	0,964	0,848	0,716	84,821

A figura A.22 mostra todas as curvas *ROC* geradas e as suas respectivas áreas, para cada valor de vizinhos mais próximos.

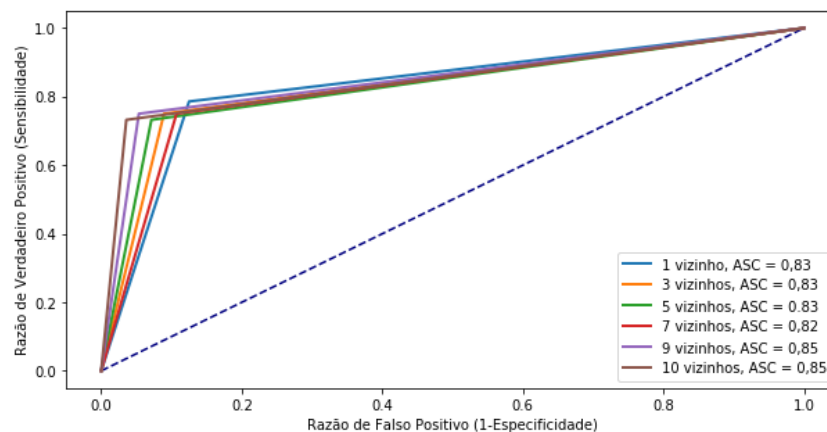


Figura A.22: Curvas *ROC* do classificador *KNN*

É possível observar na tabela A.43 que os melhores resultados foram obtidos com o classificador configurado com 10 vizinhos mais próximos. A tabela A.44 mostra a matriz de confusão gerada para o melhor modelo *KNN*.

Tabela A.44: *Matriz de confusão do classificador KNN utilizando HT e FFT como ferramenta de extração de características em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	41	2	43
	Negativo	15	54	69
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma *ASC* em torno de 0,85 e um valor de *CCM* de 0,715, mais perto de uma classificação perfeita. A taxa de acerto foi de 84,821%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,964 e uma sensibilidade em 0,732.

A.1.22 AVALIAÇÃO DO CLASSIFICADOR *SVM* COM CARACTERÍSTICAS EXTRAÍDAS USANDO *HT* E *FFT* EM UMA JANELA DE 5S

A seguir são apresentados os resultados obtidos com o classificador *SVM*, para as mesmas entradas do caso anterior. A tabela A.45 mostra os resultados alcançados com os diversos ajustes de parâmetros para *C* e γ .

Tabela A.45: *Escolha dos melhores parâmetros para o classificador SVM em uma janela de 5s*

C e γ	S	E	Ef	CCM	TA
1 e 0,43	0,75	1	0,875	0,774	87,5
1 e 0,0025	0,964	0,339	0,652	0,389	65,179
100 e 0,25	0,75	0,982	0,867	0,752	86,607
200 e 0,50	0,75	0,982	0,867	0,752	86,607
300 e 0,75	0,75	1	0,875	0,774	87,5
400 e 1	0,75	0,982	0,867	0,753	86,607

A figura A.23 mostra as curvas *ROC* geradas nas diversas configurações do classificador *SVM*.

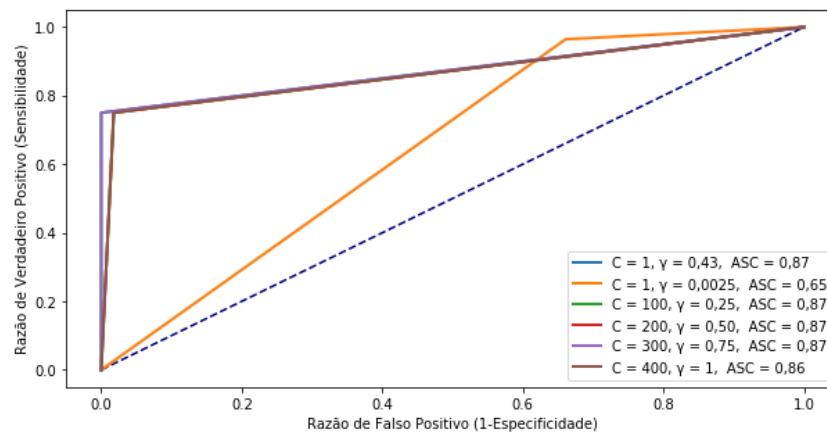


Figura A.23: Curvas ROC do classificador SVM, em uma janela de 5s, usando as técnicas HT e FFT

De acordo com a tabela A.45, os melhores resultados foram identificados com o classificador SVM ajustado com valores de $C = 300$ e $\gamma = 0,75$. Por sua vez, a tabela A.46 mostra a matriz de confusão gerada para o melhor caso.

Tabela A.46: Matriz de confusão do classificador SVM utilizando o HT e FFT como ferramenta de extração das características em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	0	42
	Negativo	14	56	70
Total		56	56	112

Logo, a previsão para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,87. O valor de CCM foi de 0,775, próximo a uma classificação perfeita. A taxa de acerto foi de 87,5%. O classificador acertou todos a classe negativas. A especificidade ficou em 1 e a sensibilidade igual a 0,75.

A.1.23 AVALIAÇÃO DO CLASSIFICADOR RL COM CARACTERÍSTICAS EXTRAÍDAS USANDO HT E FFT EM UMA JANELA DE 5S

A regressão logística foi testada nas mesmas condições dos dois casos anteriores, ou seja, usando as entradas extraídas com a combinação de ambas as técnicas no domínio da frequência. Neste sentido, a tabela A.47 apresenta os resultados gerados com o classificador RL, com características extraídas usando as técnicas HT e FFT, para uma janela de 5s.

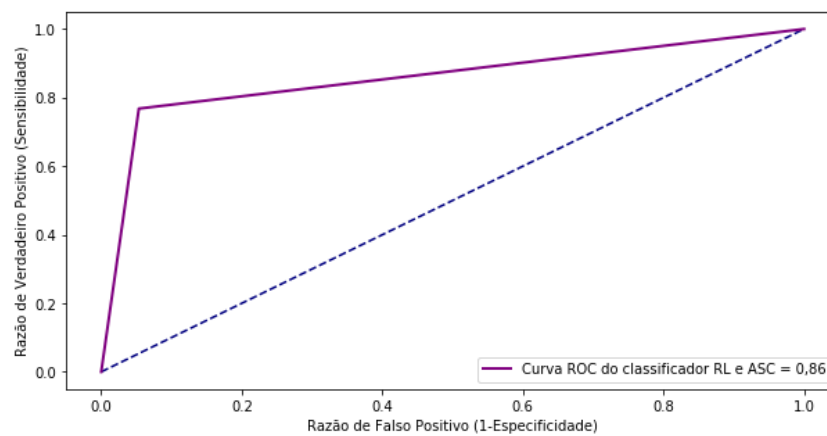
Tabela A.47: Métricas para a análise do classificador *RL* em uma janela de 5s

S	E	Ef	CCM	TA
0,768	0,946	0,857	0,725	85,714

A tabela A.48 mostra a matriz de confusão do classificador *RL* e a figura A.24 ilustra a sua respectiva curva *ROC*.

Tabela A.48: Matriz de confusão do classificador *RL* utilizando a *FFT* como ferramenta de extração das características em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	43	3	46
	Negativo	13	53	66
Total		56	56	112

**Figura A.24:** Curva *ROC* do classificador *RL* utilizando *HT* e *FFT*

A previsão para esse classificador foi boa, segundo a tabela 2.5. A figura A.9 mostra a *ASC* obtida, no caso em torno de 0,86, para o *CCM* em torno de 0,725, mais próxima de uma classificação perfeita, com a taxa de acerto em torno de 85,7143%. O classificador obteve mais acertos negativos do que acertos positivos neste último caso, com a especificidade em torno de 0,946 e a sensibilidade em 0,767. Verifica-se, portanto, que a regressão logística obteve um desempenho muito próximo ao verificado para o melhor classificador, ou seja, para o *SVM*, considerando os valores obtidos para *ASC* em cada caso.

A.1.24 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO COMO ENTRADAS AS CARACTERÍSTICAS EXTRAÍDAS COM AS GRANDEZAS *Estatísticas* E AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 5s

Nessa fase do estudo foi avaliada a performance dos três classificadores, com a combinação das características extraídas nos domínios do tempo e da frequência, ou seja, com as entradas obtidas das grandezas estatísticas e aquelas geradas pelas técnicas *HT* e *FFT* em uma janela de 5s. A tabela A.49 mostra os resultados obtidos com o classificador *KNN*, para os diferentes valores de vizinhos mais próximos.

Tabela A.49: Resultados do classificador *KNN*, em uma janela de 5s, para diferentes valores de *K*

K	S	E	Ef	CCM	TA%
1	0,803	0,839	0,821	0,643	82,142
3	0,75	0,910	0,830	0,669	83,035
5	0,714	0,946	0,830	0,679	83,035
7	0,714	0,946	0,830	0,679	83,036
9	0,732	0,946	0,839	0,694	83,928
10	0,732	0,946	0,839	0,694	83,928

A figura A.25 ilustra todas as curvas *ROC* geradas com as suas respectivas áreas, para cada valor de vizinhos mais próximos.

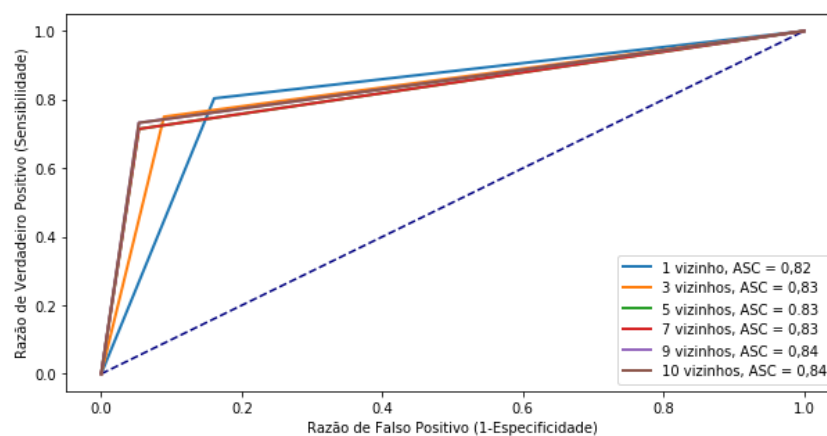


Figura A.25: Curvas *ROC* geradas para o classificador *KNN* em uma janela de 5s, usando características de entrada nos domínios do tempo e da frequência

De acordo com a tabela A.49, os melhores resultados das métricas de avaliação foram usados com o classificador parametrizado em 10 vizinhos mais próximos. A tabela A.50 mostra a matriz de confusão do melhor modelo encontrado.

Tabela A.50: *Matriz de confusão do classificador KNN utilizando as técnicas HT e FFT, com as grandezas Estatísticas, em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	41	3	44
	Negativo	15	53	68
Total		56	56	112

Nota-se que a previsão para esse classificador foi boa, segundo a tabela 2.5, com uma *ASC* em torno de 0,84. O valor de *CCM* ficou em torno de 0,695, entre uma classificação aleatória e uma classificação perfeita. A taxa de acerto foi de 83,928%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,946 e uma sensibilidade em 0,732.

A.1.25 AVALIAÇÃO DO CLASSIFICADOR *SVM* USANDO COMO ENTRADAS AS CARACTERÍSTICAS EXTRAÍDAS COM AS GRANDEZAS *Estatísticas* E AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 5S

No presente caso, o classificador *SVM* foi testado para as mesmas entradas anteriores e na mesma janela amostral. A tabela A.51 apresenta as principais métricas utilizadas para a avaliação do classificador *SVM*, com os diferentes ajustes de *C* e γ .

Tabela A.51: *Avaliação dos parâmetros do classificador SVM em uma janela de 5s*

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,0357	1	0,517	0,134	51,785
100 e 0,25	0,089	0,982	0,535	0,158	53,571
200 e 0,50	0,571	0,946	0,758	0,558	75,892
300 e 0,75	0,732	0,910	0,821	0,653	82,142
400 e 1	0,768	0,892	0,830	0,665	83,035
440 e 1,1	0,696	1	0,848	0,730	84,821

A figura A.26 mostra todas as curvas *ROC* do classificador *SVM* para cada parametrização.

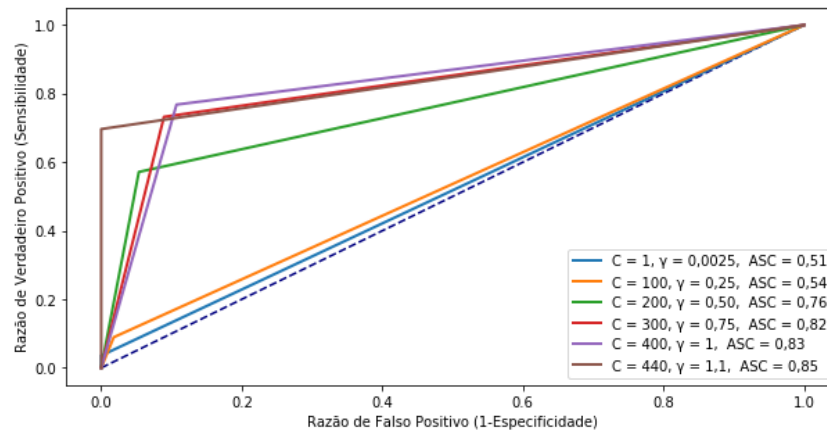


Figura A.26: *Curvas ROC geradas do classificador SVM em uma janela de 5s, usando entradas nos domínios do tempo e da frequência*

De acordo com a tabela A.51, os melhores resultados foram alcançados com o classificador SVM configurado com $C = 440$ e $\gamma = 1,1$. A tabela A.52 mostra a matriz de confusão do melhor modelo encontrado.

Tabela A.52: *Matriz de confusão do classificador SVM utilizando HT e FFT e Estatísticas como ferramenta de extração das características em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	0	39
	Negativo	17	56	73
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,85, o valor de CCM em 0,731, logo, próximo a uma classificação perfeita. A taxa de acerto foi de 84,821%. O classificador encontrou todos os acertos negativos. A especificidade ficou igual a 1 e a sensibilidade igual a 0,696.

A.1.26 AVALIAÇÃO DO CLASSIFICADOR RL COM CARACTERÍSTICAS EXTRAÍDAS USANDO HT E FFT E $Estatísticas$ EM UMA JANELA DE 5S

Finalmente, as mesmas entradas foram testadas para a regressão logística, na janela amostral de 5s, como será apresentado a seguir. A tabela A.53 mostra os valores obtidos para o classificador RL , com as características de entrada extraídas usando as técnicas HT e FFT , com as grandezas $Estatísticas$.

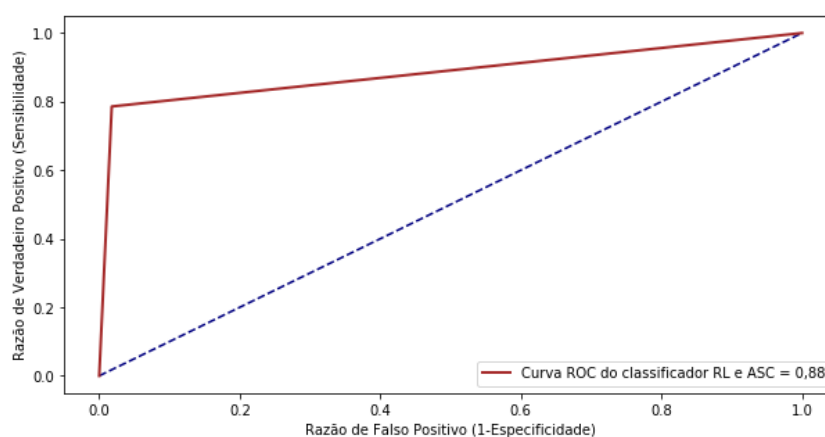
Tabela A.53: Métricas para a análise do classificador *RL* em uma janela de 5s

S	E	Ef	CCM	TA
0,785	0,982	0,884	0,783	88,392

A tabela A.54 mostra a matriz de confusão do classificador *RL* e a figura A.27 ilustra a sua respectiva curva *ROC*.

Tabela A.54: Matriz de confusão do classificador *RL* utilizando as técnicas *HT* e *FFT*, com as Estatísticas, em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	44	1	45
	Negativo	12	55	67
Total		56	56	112

**Figura A.27:** Curva *ROC* do classificador *RL* utilizando *HT* *FFT* e Estatísticas

Verifica-se que a previsão obtida para esse classificador foi boa, segundo a tabela 2.5. A figura A.27 mostra a *ASC* alcançada em torno de 0,88, para um valor de *CCM* em torno de 0,783, logo, mais próxima de uma classificação perfeita. A taxa de acerto foi de 88,393%. O classificador obteve mais acertos negativos do que acertos positivos, com a especificidade em torno de 0,982 e a sensibilidade em 0,785. Portanto, a regressão logística apresentou para essas entradas combinadas um desempenho superior aos outros dois classificadores, tomando por base os valores da *ASC* nas curvas *ROC* geradas nos melhores casos.

A.1.27 AVALIAÇÃO DO CLASSIFICADOR *KNN* COM CARACTERÍSTICAS *Estatísticas* EM UMA JANELA DE 5s

Como última análise na janela amostral de 5s, foram testados os mesmos classificadores usando como entradas apenas informações extraídas da corrente do motor no domínio do tempo, ou seja, com as grandezas estatísticas calculadas como apresentado no Capítulo 2. A tabela A.55 mostra os resultados obtidos para o classificador *KNN*, com os diferentes valores de vizinhos mais próximos.

Tabela A.55: Escolha dos melhores parâmetros para o classificador *KNN* em uma janela de 5s

K	S	E	Ef	CCM	TA %
1	0,821	0,821	0,821	0,643	82,142
3	0,803	0,875	0,839	0,680	83,928
5	0,768	0,928	0,848	0,705	84,821
7	0,732	0,911	0,821	0,653	82,143
9	0,714	0,911	0,8125	0,637	81,25
10	0,714	0,964	0,839	0,70	83,928

A figura A.28 mostra as curvas *ROC*, com as suas respectivas áreas para cada valor de vizinhos mais próximos.

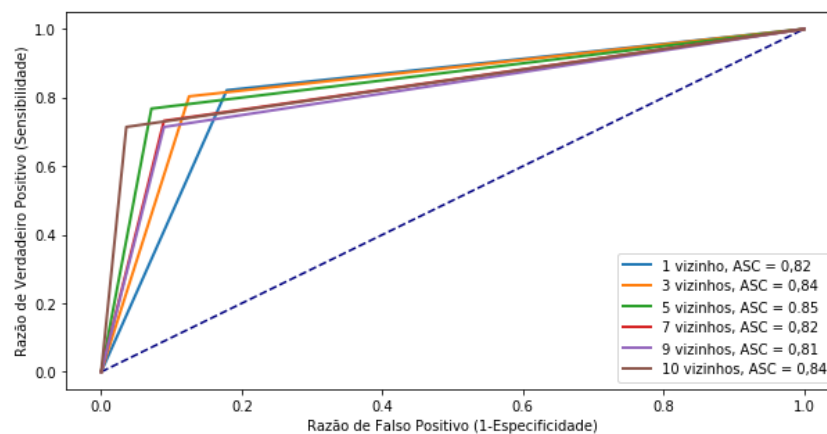


Figura A.28: Curvas *ROC* do classificador *KNN*, em uma janela de 5s, usando apenas entradas geradas no domínio do tempo

Como pode ser visto na tabela A.55, os melhores resultados foram obtidos para o classificador ajustado com 5 vizinhos mais próximos. A tabela A.56 mostra a matriz de confusão gerada para o melhor classificador.

Tabela A.56: *Matriz de confusão do classificador KNN utilizando características Estatísticas em uma janela de 5s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	43	4	47
	Negativo	13	52	65
Total		56	56	112

Conforme a tabela 2.5, nota-se que a previsão obtida para esse classificador foi boa, com uma *ASC* em torno de 0,85. Um valor de *CCM* ficou em torno de 0,706, assim, mais perto de uma classificação perfeita. A taxa de acerto foi de 84,821%. O classificador obteve uma maior quantidade de acertos negativos do que positivos, com o valor de sensibilidade em torno de 0,768 e a especificidade em 0,928.

A.1.28 AVALIAÇÃO DO CLASSIFICADOR *SVM* USANDO COMO ENTRADAS APENAS AS CARACTERÍSTICAS ESTATÍSTICAS EM UMA JANELA DE 5s

O classificador *SVM* também foi testado nas mesmas condições anteriores, isto é, usando apenas entradas processadas no domínio do tempo e uma janela amostral igual a 5s. A tabela A.57 apresenta os resultados do classificador *SVM*, para os diferentes ajustes de C e γ .

Tabela A.57: *Resultados do classificador SVM, em uma janela de 5s, para diferentes valores dos parâmetros*

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,035	1	0,518	0,135	51,785
100 e 0,25	0,196	0,928	0,562	0,183	56,25
200 e 0,50	0,643	0,892	0,768	0,553	76,785
300 e 0,75	0,768	0,875	0,821	0,646	82,142
400 e 1	0,839	0,875	0,857	0,714	85,714
1 e 4,2	0,696	1	0,848	0,730	84,821

A figura A.29 mostra as curvas *ROC* do classificador *SVM*, para cada caso de parametrização.

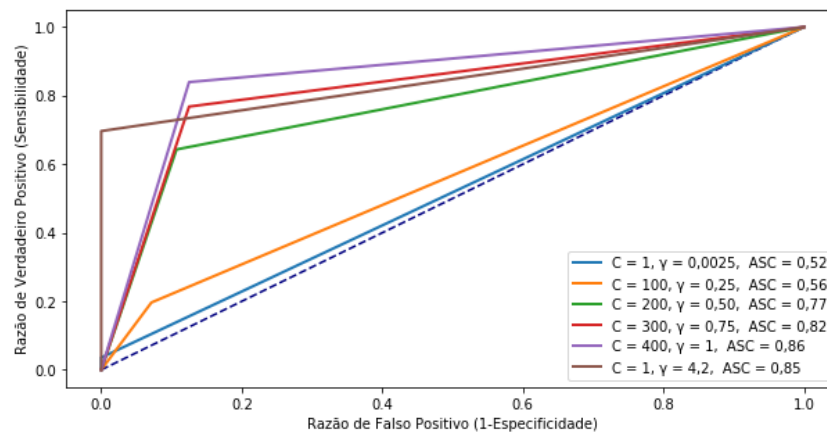


Figura A.29: Curvas ROC do classificador SVM em uma janela de 5s, usando apenas entradas no domínio do tempo

De acordo com a tabela A.57, os melhores resultados foram alcançados para este classificador com os parâmetros configurados como $C = 400$ e $\gamma = 1$. A tabela A.58 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.58: Matriz de confusão do classificador KNN utilizando o FFT como ferramenta de extração das características em uma janela de 10s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	47	7	54
	Negativo	9	49	58
Total		56	56	112

Neste caso a previsão foi boa, segundo a tabela 2.5, com uma ASC em torno de 0,86. O valor de CCM foi em torno de 0,714, próximo a uma classificação perfeita. A taxa de acerto foi de 85,71%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,875 e uma sensibilidade em 0,839.

A.1.29 AVALIAÇÃO DO CLASSIFICADOR RL USANDO COMO ENTRADAS APENAS AS CARACTERÍSTICAS ESTATÍSTICAS EM UMA JANELA DE 5S

Como análise final para essa janela amostral, foi avaliado o desempenho da regressão logística usando como entradas apenas as características estatísticas calculadas conforme descrito no Capítulo 2. A tabela A.59 mostra os resultados obtidos com o uso classificador RL .

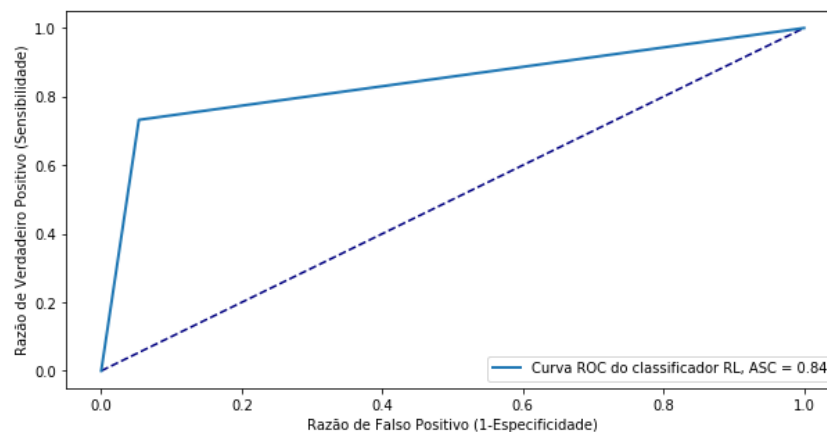
Tabela A.59: Métricas para a análise do classificador *RL* em uma janela de 5s

S	E	Ef	CCM	TA
0,732	0,946	0,839	0,694	83,928

A tabela A.60 mostra a matriz de confusão do classificador *RL* e a sua respectiva curva *ROC* é mostrada na figura A.30.

Tabela A.60: Matriz de confusão do classificador *RL* utilizando características Estatísticas em uma janela de 5s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	41	3	44
	Negativo	15	53	68
Total		56	56	112

**Figura A.30:** Curva *ROC* do classificador *RL* utilizando Estatísticas

A previsão para esse classificador também foi boa, conforme os dados previstos e mostrados na tabela 2.5. A figura A.30 mostra a *ASC* igual a 0,84, para um valor de *CCM* em torno de 0,69, logo, entre uma classificação aleatória e uma classificação perfeita. A taxa de acerto foi de 83,928%. O classificador obteve mais acertos positivos do que acertos negativos, com a especificidade em torno de 0,946 e a sensibilidade em 0,732. Na comparação, nota-se que o classificador *SVM* apresentou um desempenho ligeiramente superior aos encontrados para os outros dois modelos, considerando a área calculada pela curva *ROC*. De todo o modo, empregando apenas entradas no domínio do tempo, e para uma janela amostral igual a 5s, verifica-se que todos os classificadores obtiveram uma performance muito próxima, com um *ASC* médio igual a 0,85 ($ASC(KNN) = 0,85$, $ASC(SVM) = 0,86$ e $ASC(RL) = 0,84$).

A.1.30 AVALIAÇÃO DO CLASSIFICADOR *KNN* COM CARACTERÍSTICAS EXTRAÍDAS USANDO *HT* E *FFT* EM UMA JANELA DE 3s

A seguir serão apresentados os resultados obtidos para as diferentes entradas, mas tomando por base um conjunto de dados discretizados em uma janela de 3s, portanto, inferior aos dois casos anteriores (janelas de 10s e 5s). Inicialmente, foi avaliado o desempenho dos classificadores usando as características extraídas com as técnicas combinadas *HT* e *FFT*. A tabela A.61 apresenta os resultados testados para o classificador *KNN*, com os diferentes valores de vizinhos mais próximos.

Tabela A.61: Resultados do classificador *KNN* em uma janela de 3s

K	S	E	Ef	CCM	TA %
1	0,714	0,821	0,768	0,539	76,786
3	0,696	0,857	0,776	0,561	77,678
5	0,696	0,875	0,786	0,581	78,571
7	0,660	0,875	0,767	0,548	76,785
9	0,643	0,875	0,759	0,532	75,893
10	0,625	0,910	0,768	0,559	76,786

A figura A.31 mostra as curvas *ROC* geradas com as suas respectivas áreas, para cada valor de vizinhos mais próximos.

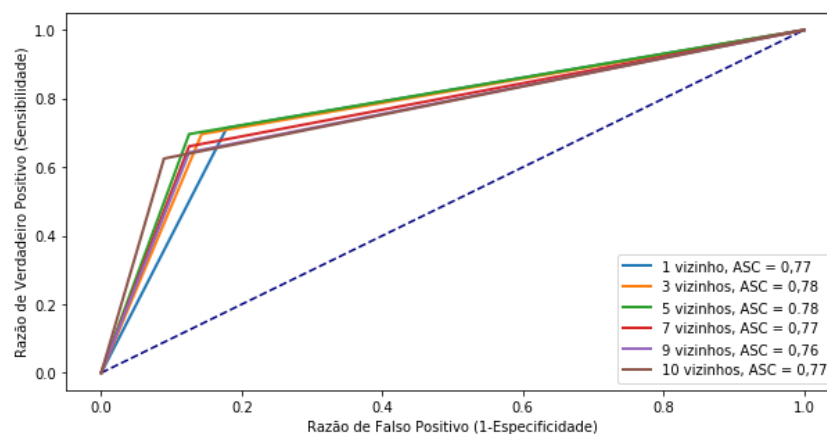


Figura A.31: Curvas *ROC* geradas com o classificador *KNN* em uma janela de 3s

Como pode ser observado na tabela A.61, os melhores resultados foram obtidos para o classificador configurado com 5 vizinhos mais próximos. A tabela A.62 mostra a matriz de confusão gerada com o melhor modelo.

Tabela A.62: *Matriz de confusão do classificador KNN utilizando HT e FFT como ferramenta de extração de características em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	7	46
	Negativo	17	49	66
Total		56	56	112

A previsão neste caso foi fraca com uma ASC em torno de 0,78. O valor de CCM foi de 0,58, entre uma classificação aleatória e perfeita. A taxa de acerto foi de 78,57%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,875 e uma sensibilidade em 0,696.

A.1.31 AVALIAÇÃO DO CLASSIFICADOR SVM COM CARACTERÍSTICAS EXTRAÍDAS USANDO HT E FFT EM UMA JANELA DE 3S

O classificador SVM também foi testado para as mesmas entradas do caso anterior. A tabela A.63 apresenta as principais métricas utilizadas para a avaliação do classificador SVM , com os diferentes parâmetros de C e γ .

Tabela A.63: *Avaliação dos parâmetros encontrados para o classificador SVM em uma janela de 3s*

C e γ	S	E	Ef	CCM	TA
1 e 1,7	0,643	0,964	0,803	0,641	80,357
1 e 0,0025	0,928	00,303	0,616	0,297	61,607
100 e 0,25	0,678	0,857	0,867	0,544	76,785
200 e 0,50	0,696	0,928	0,867	0,642	81,25
300 e 0,75	0,660	0,928	0,875	0,611	79,464
400 e 1	0,660	0,964	0,867	0,655	81,25

A figura A.32 mostra as curvas ROC do classificador SVM para cada caso de parametrização.

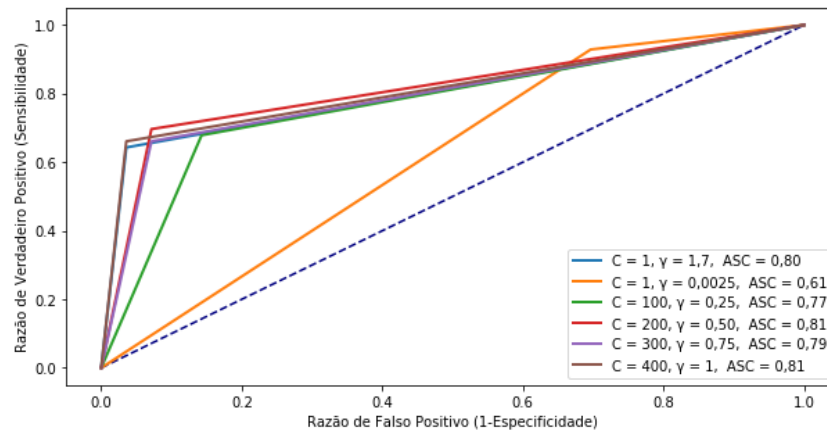


Figura A.32: *Curvas ROC do classificador SVM*

De acordo com a tabela A.63, os melhores resultados das métricas de avaliação foram usados com o classificador *SVM* parametrizado com $C = 400$ e $\gamma = 1$. A tabela A.64 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.64: *Matriz de confusão do classificador SVM utilizando o HT e FFT como ferramenta de extração das características em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	37	2	39
	Negativo	19	54	73
Total		56	56	112

Neste caso a previsão foi boa, segundo a tabela 2.5, com uma *ASC* em torno de 0,81 e um *CCM* de 0,66, entre uma classificação aleatória e uma classificação perfeita. A taxa de acerto foi de 81,25%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,96 e uma sensibilidade em 0,66.

A.1.32 AVALIAÇÃO DO CLASSIFICADOR *RL* COM CARACTERÍSTICAS EXTRAÍDAS USANDO *HT* E *FFT* EM UMA JANELA DE 3S

Por fim, a regressão logística também foi testada com as mesmas entradas dos dois casos anteriores e na janela de 3s. A tabela A.65 mostra as métricas de análise do classificador *RL*, com características extraídas utilizando as técnicas *HT* e *FFT*.

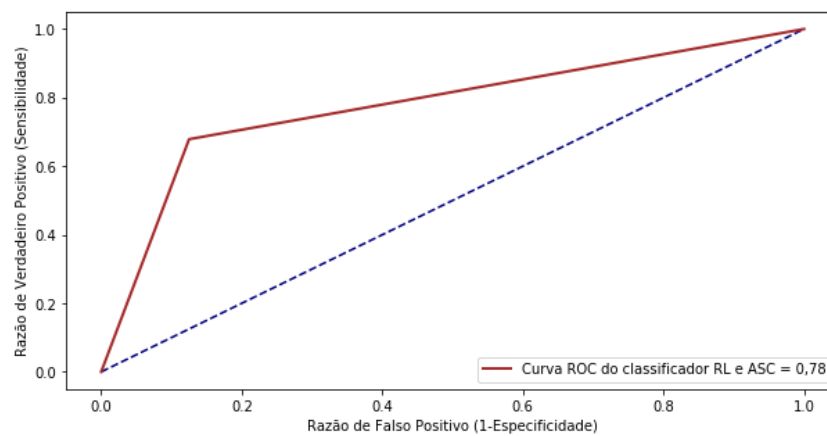
Tabela A.65: Métricas para a análise do classificador *RL* em uma janela de 3s

S	E	Ef	CCM	TA
0,678	0,875	0,777	0,453	77,678

A tabela A.66 mostra a matriz de confusão do classificador *RL* e a sua respectiva curva *ROC* está ilustrada na figura A.33.

Tabela A.66: Matriz de confusão do classificador *RL* utilizando a *FFT* como ferramenta de extração das características em uma janela de 3s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	38	7	45
	Negativo	18	49	67
Total		56	56	112

**Figura A.33:** Curva *ROC* do classificador *RL* utilizando *HT* e *FFT*

A classificação nesse caso foi fraca, segundo a tabela 2.5. A figura A.33 mostra a *ASC* em torno de 0,78 e o valor de *CCM* em 0,453, mais próximo de uma classificação aleatória. A taxa de acerto foi de 77,678%. O classificador obteve mais acertos negativos do que acertos positivos, com a especificidade em torno de 0,875 e a sensibilidade em 0,679. Em comparação com os três classificadores, nota-se que o melhor *SVM* apresentou um desempenho superior aquele encontrado para os modelos *KNN* e *RL*, usando novamente a *ASC* como referência.

A.1.33 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO COMO ENTRADAS AS CARACTERÍSTICAS *Estatísticas* E AQUELAS EXTRAÍDAS COM AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 3S

Nesta fase do trabalho foi avaliado novamente o desempenho dos classificadores para entradas combinadas nos domínios do tempo e da frequência, mas para uma janela amostral de 3s. A tabela A.67 apresenta os resultados obtidos com o classificador *KNN*, para os diversos ajustes dos vizinhos mais próximos.

Tabela A.67: Escolha dos melhores parâmetros para o classificador *KNN* em uma janela de 3s

K	S	E	Ef	CCM	TA%
1	0,768	0,768	0,768	0,535	76,786
3	0,768	0,785	0,777	0,553	77,678
5	0,75	0,839	0,794	0,591	79,464
7	0,732	0,875	0,803	0,613	80,357
9	0,732	0,946	0,839	0,695	83,928
10	0,7143	0,982	0,848	0,723	84,821

A figura A.25 mostra todas as curvas *ROC* com as suas respectivas áreas para cada valor de vizinhos mais próximos.

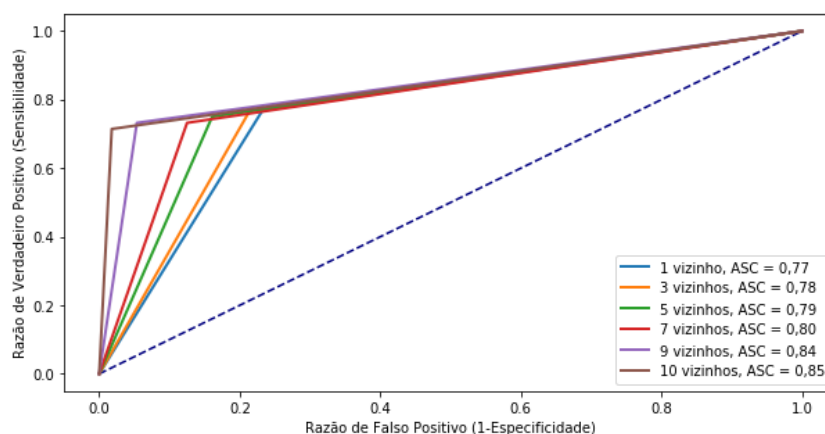


Figura A.34: Curvas *ROC* do classificador *KNN* na janela de 3s e entradas nos domínios do tempo e da frequência

De acordo com a tabela A.67, os melhores resultados das métricas de avaliação foram encontrados com o classificador configurado em 10 vizinhos mais próximos. A tabela A.68 mostra a matriz de confusão para o melhor modelo escolhido.

Tabela A.68: *Matriz de confusão do classificador KNN utilizando HT e FFT e Estatísticas como ferramenta de extração de características em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	40	1	41
	Negativo	16	55	71
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5, com uma *ASC* em torno de 0,85 e o *CCM* igual a 0,723, mais perto de uma classificação perfeita. A taxa de acerto foi de 84,82%. O classificador obteve mais acertos negativos do que positivos, com uma especificidade em torno de 0,982 e uma sensibilidade em 0,714.

A.1.34 AVALIAÇÃO DO CLASSIFICADOR *SVM* USANDO COMO ENTRADAS AS CARACTERÍSTICAS *Estatísticas* E AQUELAS EXTRAÍDAS COM AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 3S

Aqui o classificador *SVM* também foi testado para as mesmas entradas anteriores. A tabela A.69 apresenta as principais métricas utilizadas para a avaliação do classificador *SVM* com os parâmetros de C e γ .

Tabela A.69: *Escolha dos melhores parâmetros para o classificador SVM em uma janela de 3s*

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,428	0,535	0,482	-0,036	48,214
100 e 0,25	0,696	1	0,848	0,731	84,821
200 e 0,50	0,696	1	0,848	0,731	84,821
300 e 0,75	0,696	1	0,848	0,731	84,821
400 e 1	0,696	1	0,848	0,731	84,821
440 e 1,1	0,696	1	0,848	0,731	84,821

A figura A.35 mostra as curvas *ROC* do classificador *SVM* para cada caso de parametrização.

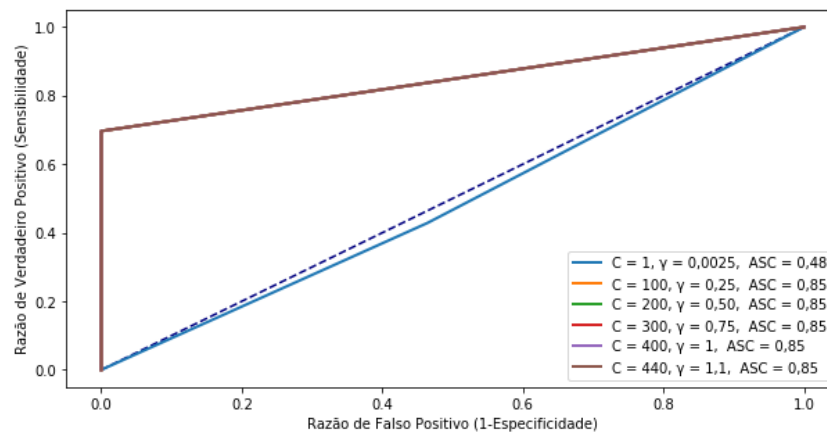


Figura A.35: *Curvas ROC do classificador SVM*

De acordo com a tabela A.69, os melhores resultados foram obtidos com o classificador *SVM* configurado com $C = 100$ e $\gamma = 0,25$. A tabela A.70 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.70: *Matriz de confusão do classificador SVM utilizando HT e FFT e Estatísticas como ferramenta de extração das características em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	39	0	39
	Negativo	17	56	73
Total		56	56	112

A previsão para esse classificador foi boa, segundo a tabela 2.5 com uma *ASC* em torno de 0,85, o *CCM* foi de 0,73, próximo a uma classificação perfeita. A taxa de acerto foi de 84,82%. O classificador obteve todos os acertos negativos. A especificidade ficou igual a 1 e a sensibilidade igual a 0,696. O modelo acertou todas as classificações para a classe negativa.

A.1.35 AVALIAÇÃO DO CLASSIFICADOR *RL* USANDO COMO ENTRADAS AS CARACTERÍSTICAS *Estatísticas* E AQUELAS EXTRAÍDAS COM AS TÉCNICAS *HT* E *FFT* EM UMA JANELA DE 3S

Como última avaliação para esse conjunto de entradas, foi testado o desempenho da regressão logística. A tabela A.71 mostra os parâmetros de análise do classificador *RL*, com características extraídas utilizando a *HT FFT*, além das grandezas *Estatísticas*.

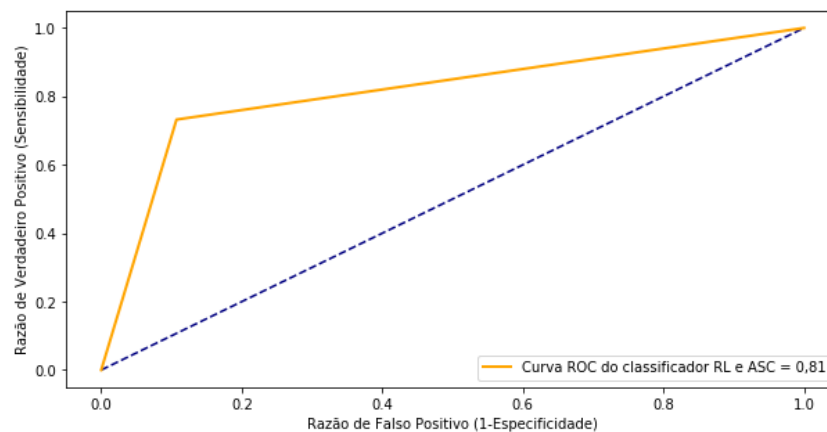
Tabela A.71: Métricas para a análise do classificador *RL* em uma janela de 3s

S	E	Ef	CCM	TA
0,732	0,893	0,812	0,497	81,25

A tabela A.72 mostra a matriz de confusão do classificador *RL*, com a sua respectiva curva *ROC*, como mostra a figura A.36.

Tabela A.72: Matriz de confusão do classificador *RL* utilizando *HT FFT* e *Estatísticas* como ferramenta de extração das características em uma janela de 3s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	41	6	47
	Negativo	15	50	65
Total		56	56	112

**Figura A.36:** Curva *ROC* do classificador *RL* utilizando *HT FFT* e *Estatísticas*

A previsão para esse classificador foi boa segundo, a tabela 2.5. A figura A.27 mostra a *ASC* em torno de 0,81, o *CCM* foi de 0,496, mais próxima de uma classificação aleatória. A taxa de acerto foi de 81,25%. O classificador obteve mais acertos negativos do que acertos positivos com a especificidade em torno de 0,893 e a sensibilidade em 0,732. Na comparação dos três classificadores, nota-se neste caso que os modelos *KNN* e *SVM* obtiveram o mesmo desempenho, usando novamente como referência o valor da *ASC*, sendo esta superior ao observado para a regressão logística.

A.1.36 AVALIAÇÃO DO CLASSIFICADOR *KNN* USANDO COMO ENTRADAS AS CARACTERÍSTICAS *Estatísticas* EM UMA JANELA DE 3s

Finalmente, para a janela amostral igual a 3s foi feita uma comparação de desempenho dos classificadores usando apenas as entradas extraídas no domínio do tempo, ou seja, com o uso das grandezas estatísticas mostradas no Capítulo 2. A tabela A.73 apresenta as principais métricas utilizadas para a avaliação do classificador *KNN* com os diferentes valores de vizinhos mais próximos.

Tabela A.73: Avaliação dos melhores parâmetros para o classificador *KNN* em uma janela de 3s

V	S	E	Ef	CCM	TA %
1	0,803	0,821	0,8125	0,625	81,25
3	0,732	0,821	0,776	0,556	77,678
5	0,732	0,821	0,777	0,556	77,678
7	0,75	0,875	0,812	0,63	81,25
9	0,75	0,9286	0,839	0,689	83,928
10	0,732	0,964	0,848	0,716	84,821

A figura A.37 mostra uma comparação entre todas as curvas *ROC* geradas, com as suas respectivas áreas para cada valor de vizinhos mais próximos.

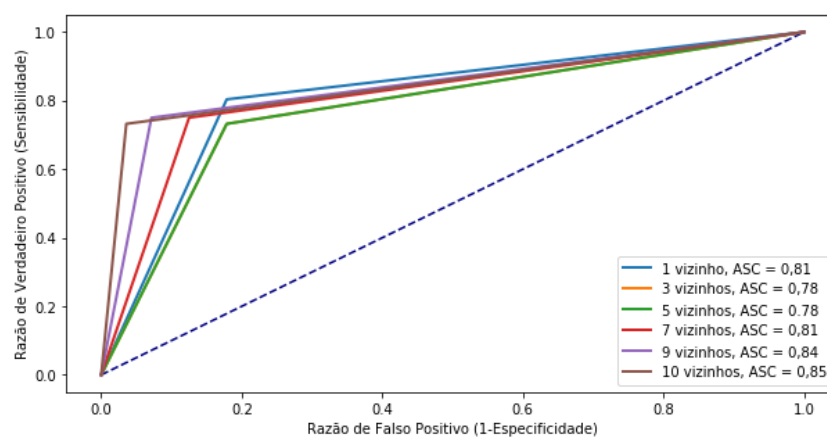


Figura A.37: Curvas *ROC* do classificador *KNN* em uma janela de 3s, usando apenas entradas no domínio do tempo

De acordo com a tabela A.73, os melhores resultados das métricas de avaliação foram encontrados com o classificador parametrizado em 10 vizinho mais próximos. A tabela A.74 mostra a matriz de confusão para o melhor modelo testado.

Tabela A.74: *Matriz de confusão do classificador KNN utilizando características Estatísticas em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	41	2	43
	Negativo	15	54	69
Total		56	56	112

A previsão para esse classificador foi boa, conforme a tabela 2.5, indicando uma ASC em torno de 0,85 e um valor de CCM próximo a 0,716, logo, mais perto de uma classificação perfeita. A taxa de acerto foi de 84,821%. O classificador obteve uma maior quantidade de acertos negativos do que positivos com o valores de sensibilidade em torno de 0,732 e a especificidade em 0,964.

A.1.37 AVALIAÇÃO DO CLASSIFICADOR SVM USANDO COMO ENTRADAS AS CARACTERÍSTICAS *Estatísticas* EM UMA JANELA DE 3s

A seguir são apresentados os resultados obtidos com o uso do classificador SVM , nas mesmas condições do caso anterior. A tabela A.75 apresenta os resultados deste classificador, para os diferentes valores de C e γ .

Tabela A.75: *Escolha dos melhores parâmetros para o classificador SVM em uma janela de 3s*

C e γ	S	E	Ef	CCM	TA
1 e 0,0025	0,035	1	0,518	0,135	51,785
100 e 0,25	0,214	1	0,607	0,346	60,714
200 e 0,50	0,643	0,893	0,768	0,553	76,786
300 e 0,75	0,714	0,839	0,777	0,558	77,678
400 e 1	0,75	0,839	0,794	0,591	79,464
1 e 17	0,75	0,839	0,794	0,591	79,464

A figura A.38 mostra as curvas ROC do classificador SVM , para cada conjunto de parâmetros.

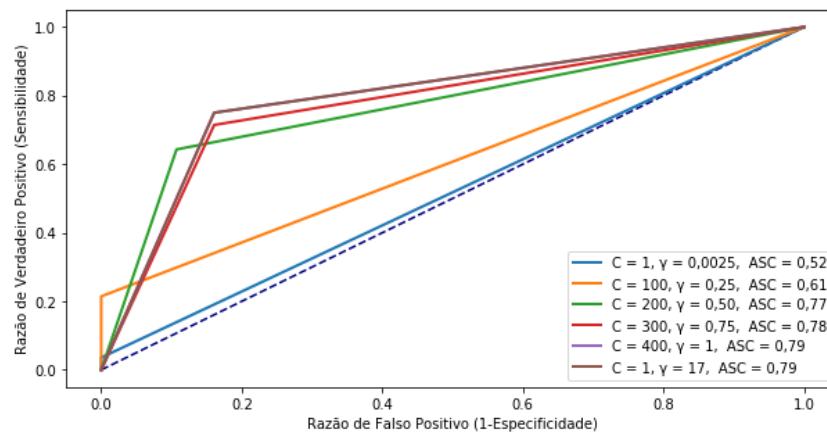


Figura A.38: *Curvas ROC do classificador SVM*

De acordo com a tabela A.75, os melhores resultados foram observados para o classificador ajustado com $C = 1$ e $\gamma = 17$. A tabela A.76 mostra a matriz de confusão do melhor modelo escolhido.

Tabela A.76: *Matriz de confusão do classificador KNN usando como entradas as características Estatísticas em uma janela de 3s*

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	9	51
	Negativo	14	47	61
Total		56	56	112

A previsão para esse classificador foi fraca, segundo a tabela 2.5, com uma ASC em torno de 0,79, o CCM foi de 0,59, portanto, entre uma classificação aleatória e uma classificação perfeita. A taxa de acerto foi de 79,464%. O classificador obteve mais acertos negativos do que positivos com um especificidade em torno de 0,875 e uma sensibilidade em 0,839.

A.1.38 AVALIAÇÃO DO CLASSIFICADOR RL COM CARACTERÍSTICAS *Estatísticas* EM UMA JANELA DE 3S

Como última análise para esse conjunto de entradas, foi testado novamente o desempenho da regressão logística. Deste modo, a tabela A.77 mostra os resultados das métricas de avaliação do classificador RL .

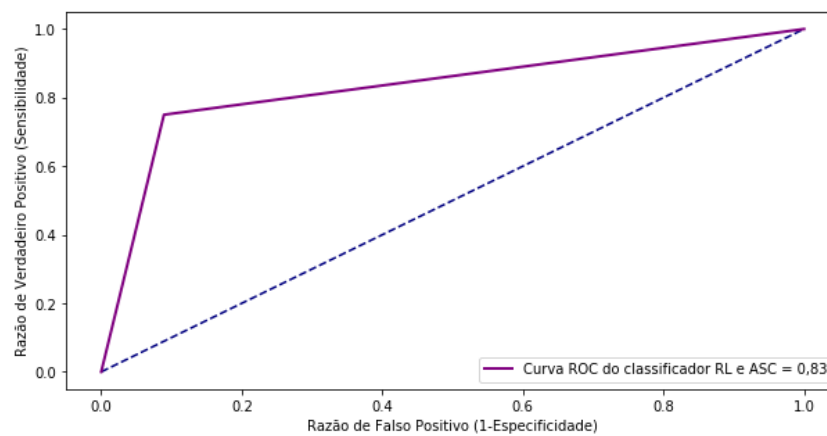
Tabela A.77: Métricas para a análise do classificador *RL* em uma janela de 3s

S	E	Ef	CCM	TA
0,75	0,91	0,830	0,519	83,035

A tabela A.78 mostra a matriz de confusão do classificador *RL*, com a sua respectiva curva *ROC* ilustrada na figura A.39.

Tabela A.78: Matriz de confusão do classificador *RL* utilizando características Estatísticas em uma janela de 3s

		Diagnóstico verdadeiro		Total
		Positivo	Negativo	
Teste	Positivo	42	5	47
	Negativo	14	51	65
Total		56	56	112

**Figura A.39:** Curva *ROC* do classificador *RL* utilizando Estatísticas

Nota-se que a previsão obtida para a regressão foi boa, segundo a tabela 2.5. A figura A.39 mostra a *ASC* em torno de 0,83, para um *CCM* de 0,519, entre uma classificação aleatória e uma classificação perfeita. A taxa de acerto foi de 83,035%. O classificador obteve mais acertos negativos do que acertos positivos, com a especificidade em torno de 0,91 e a sensibilidade em 0,75. Neste último caso, o classificador *KNN* obteve o melhor desempenho, com uma *ASC* = 0,85. A regressão logística obteve a segunda melhor performance com uma *ASC* = 0,83 e o classificador *SVM* obteve o pior desempenho com *ASC* = 0,79.

B.1 RESUMO DAS CONFIGURAÇÕES EMPREGADAS EM TODOS OS MODELOS

No presente trabalho foram avaliadas diferentes variações para os três classificadores testados, em todas as janelas amostrais. A tabela B.1 mostra todos os modelos utilizados nas janelas estudadas.

Tabela B.1: *Total das variações avaliadas no trabalho*

Janelas	Modelos
10s	30 para o <i>KNN</i> , 30 para o <i>SVM</i> e 5 para a <i>RL</i>
5s	30 para o <i>KNN</i> , 30 para o <i>SVM</i> e 5 para a <i>RL</i>
3s	18 para o <i>KNN</i> , 18 para o <i>SVM</i> e 3 para a <i>RL</i>
1s	6 para o <i>KNN</i> , 6 para o <i>SVM</i> e 1 para a <i>RL</i>
0,5	6 para o <i>KNN</i> , 6 para o <i>SVM</i> e 1 para a <i>RL</i>
0,25s	6 para o <i>KNN</i> , 6 para o <i>SVM</i> e 1 para a <i>RL</i>

Como apresentado, foram selecionados os melhores modelos com as suas respectivas configurações (parâmetros) e cenários. Para as janelas iguais a 10s e 5s, foram escolhidas as 15 melhores abordagens, três para cada um dos cinco cenários, logo, foram selecionados o melhor *KNN*, o melhor *SVM* e a melhor *RL* para cada cenário. Na janela de 3s, foram avaliados os 9 melhores modelos, três para cada um dos três cenários. Para as janelas de 1s, 0,5s e 0,25s foram escolhidas as três melhores abordagens para apenas um cenário. A tabela B.2 apresenta os melhores modelos.

Tabela B.2: *Quantidade dos melhores modelos selecionados*

Janelas	Melhores modelos
10s	5 <i>KNN</i> , 5 <i>SVM</i> e 5 <i>RL</i> (Total = 15 modelos)
5s	5 <i>KNN</i> , 5 <i>SVM</i> e 5 <i>RL</i> (Total = 15 modelos)
3s	3 <i>KNN</i> , 3 <i>SVM</i> e 3 <i>RL</i> (Total = 9 modelos)
1s	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> (Total = 3 modelos)
0,5s	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> (Total = 3 modelos)
0,25s	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> (Total = 3 modelos)

Vale mencionar que, a tabela B.2 mostra os melhores modelos, contudo, na primeira linha, por exemplo, as cinco variações do modelo *KNN* foram escolhidas individualmente para cada cenário, e este procedimento foi repetido para o *SVM* e *RL*. Para cada cenário há um modelo do *KNN*, *SVM* e *RL*, como mostram as tabelas B.3, B.4 e B.5.

Tabela B.3: *Quantidade de modelos para cada cenário para 10s e 5s*

Cenários	Melhores modelos
FFT	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelos
FFT + EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo
HT + FFT	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo
FFT + FFT + EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo
EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo

Tabela B.4: *Quantidade de modelos para cada cenário para 3s*

Cenários	Melhores modelos
HT + FFT	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo
FFT + FFT + EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo
EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo

Tabela B.5: *Quantidade de modelos para cada cenário para 1s, 0,5s e 0,25s*

Cenários	Melhores modelos
EST	1 <i>KNN</i> , 1 <i>SVM</i> e 1 <i>RL</i> = 3 modelo