

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SANTA  
CATARINA – CÂMPUS FLORIANÓPOLIS  
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA  
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**RAFAEL NOBORO TOMINAGA**

**IDENTIFICAÇÃO DO TIPO E DA SEVERIDADE DO  
DESBALANCEAMENTO DE ROTORES POR MEIO DA ANÁLISE DA  
ASSINATURA DE CORRENTE COM ALGORITMOS DE  
APRENDIZADO DE MÁQUINA MULTICLASSES**

**FLORIANÓPOLIS, 2021.**

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SANTA  
CATARINA – CÂMPUS FLORIANÓPOLIS  
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA  
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**RAFAEL NOBORO TOMINAGA**

**IDENTIFICAÇÃO DO TIPO E DA SEVERIDADE DO  
DESBALANCEAMENTO DE ROTORES POR MEIO DA ANÁLISE DA  
ASSINATURA DE CORRENTE COM ALGORITMOS DE  
APRENDIZADO DE MÁQUINA MULTICLASSES**

Trabalho de Conclusão de Curso  
submetido ao Instituto Federal de  
Educação, Ciência e Tecnologia de Santa  
Catarina como parte dos requisitos para  
obtenção do título de Engenheiro(a)  
Eletricista.

Orientador:  
Prof. Sergio Luciano Avila, Dr.

Coorientador:  
Prof. Cesar Alberto Penz, Dr.

**FLORIANÓPOLIS, 2021.**

Ficha de identificação da obra elaborada pelo autor.

Tominaga, Rafael Noboro  
Identificação do tipo e da severidade do desbalanceamento de rotores por meio da análise da assinatura de corrente com algoritmo de aprendizado de máquina multiclases / Rafael Noboro Tominaga ; orientação de Sergio Luciano Avila; coorientação de Cesar Alberto Penz. - Florianópolis, SC, 2021.  
Trabalho de Conclusão de Curso (TCC) - Instituto Federal de Santa Catarina, Câmpus Florianópolis. Bacharelado em Engenharia Elétrica. Departamento Acadêmico de Eletrotécnica.  
Inclui Referências.

1. Máquinas elétricas rotativas. 2. Análise de sinais de corrente. 3. Identificação de desbalanceamento. 4. Aprendizado de máquina. I. Avila, Sergio Luciano. II. Penz, Cesar Alberto. III. Instituto Federal de Santa Catarina. Departamento Acadêmico de Eletrotécnica. IV. Título.

# IDENTIFICAÇÃO DO TIPO E DA SEVERIDADE DO DESBALANCEAMENTO DE ROTORES POR MEIO DA ANÁLISE DA ASSINATURA DE CORRENTE COM ALGORITMOS DE APRENDIZADO DE MÁQUINA MULTICLASSES

RAFAEL NOBORO TOMINAGA

Este trabalho foi julgado adequado para obtenção do título de Engenheiro(a) Eletricista e aprovado na sua forma final pela banca examinadora do Curso de Graduação em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina.

Florianópolis, 12 de março, 2021.

Banca Examinadora:

 SERGIO LUCIANO  
AVILA:02074392907  
2021.03.28 18:51:56  
-03'00'

---

Prof. Sérgio L. Avila, Dr.

Orientador

 Assinado de forma digital por  
CESAR ALBERTO  
PENZ:02172115932  
Dados: 2021.03.29 14:22:53  
-03'00'

---

Prof. Cesar A. Penz, Dr.

Coorientador

 EDSON SORATO:  
55654657020  
2021.03.29 19:50:  
49-03'00'

---

Prof. Edson Sorato, Me.



---

Prof. Fernando Lila, Me.

Dedico este trabalho ao meu pai Osvaldo Noboro Tominaga e a minha mãe Ana Paula Tominaga por serem mutuamente responsáveis por todas as conquistas que eu consegui durante a minha trajetória no curso de engenharia elétrica.

## **AGRADECIMENTOS**

Primeiramente, gostaria agradecer aos meus pais Osvaldo Noboro Tominaga e Ana Paula Tominaga, na qual eu dedico o trabalho, pelo todo apoio profissional e pessoal que eles me deram ao longo de todos os anos da minha vida.

Posteriormente, eu agradeço Osvaldo Tominaga e Eduardo Noboro Tominaga pelo todo apoio que eles me deram como irmão e ex-aluno de engenharia.

Agradeço meu orientador Prof. Sergio Luciano Avila por ter me convidado a participar do projeto de pesquisa que me deu a oportunidade para começar a desenvolver este trabalho. Também, agradeço pelo todo auxílio dado durante toda esta trajetória.

Gostaria de agradecer meu coorientador Cesar Alberto Penz pelo todo apoio técnico fornecido durante as minhas atividades na bancada em estudo e pelo aprendizado que me proporcionou ao longo deste período.

Também gostaria de agradecer Jonas Fortunato Honorato pelo apoio inicial que me foi fornecido e por ser autor de outro trabalho de conclusão de curso na qual o trabalho atual visa dar continuidade.

Agradeço também Felipe Teixeira, Heloiza Schaberle, Lucas Imamura, Rayan Steinbach, Suzi Youssef, assim como todas as outras pessoas do Grupo de Pesquisa em Computação Científica para Engenharia (PECCE) que me acompanharam paralelamente o desenvolvimento deste trabalho e também tornaram a minha experiência no projeto de pesquisa muito melhor.

Por fim, agradeço a todos os meus amigos da engenharia elétrica que estiveram me acompanhando e motivando o meu desenvolvimento acadêmico durante todos este tempo.

We're not on our journey to save the world but to save ourselves.  
But in doing that you save the world.  
The influence of a vital person vitalizes (Joseph Campbell).

## RESUMO

A saúde de uma máquina elétrica rotativa pode ser avaliada através do monitoramento de grandezas elétricas e mecânicas associadas à sua operação. Quanto mais informação estiver disponível, mais fácil o seu estado operacional pode ser diagnosticado. Para se estudar este tema, o Grupo de Pesquisas em Computação Científica para a Engenharia (PECCE) construiu uma bancada de teste que permite inserir desbalanceamentos propositais sobre o eixo de um motor de indução. No presente trabalho, fica-se restrito a diagnósticos a partir da análise dos harmônicos de corrente do motor. A partir dessa assinatura, e entendendo-se que se trata de um problema com características não-lineares e multiclasse, fez-se o uso de ferramentas computacionais para proceder a identificação (tipo) e a classificação (severidade) dos desbalanceamentos. Os algoritmos estudados aqui foram Máquina de Vetores de Suporte (SVM), Árvore de Decisão (DTC) e *One-vs-One* (OvO). Os resultados obtidos atestam a eficiência da metodologia proposta. A comparação das estratégias de aprendizado de máquina, bem como os cuidados necessários quanto ao processamento de sinais, também são contribuições deste trabalho para a área da engenharia de manutenção industrial.

**Palavras-chave:** Máquinas elétricas rotativas. Análise de sinais de corrente. Identificação de desbalanceamento. Aprendizado de máquina.

## ABSTRACT

The rotating electrical machine health can be evaluated by monitoring electrical and mechanical parameters associated with its operation. The more information is available, the easier the diagnosis of the machine operational condition can be. In order to study this subject, a test bench has been built. It allows inserting purposeful unbalances on the axis of an induction motor. The present work is restricted to diagnostics using the motor current harmonics analysis. Based on this signature, and understanding that is a problem with nonlinear and multiclass characteristics, computational tools were used to identify (type) and to classify (severity) the unbalances. The algorithms studied here were Support Vector Machine (SVM), Decision Tree (DTC) and One-vs-One (OvO). The results prove that the proposed methodology is efficiency. The comparison of machine learning strategies, as well as the necessary care regarding signal processing, is also contributions of this work to industrial maintenance engineering domain.

**Keywords:** Rotating electrical machines. Current signature analysis. Unbalance identification. Machine learning.

## LISTA DE FIGURAS

Figura 1 – Tipos de desbalanceamento em rotores .....	17
Figura 2 – Parte mecânica da bancada – desbalanceamento acoplado .....	18
Figura 3 – Esquemático elétrico da placa do transdutor de corrente .....	21
Figura 4 – Diagrama de ligação dos sinais elétricos .....	23
Figura 5 – Diagrama de ligação dos sinais mecânicos .....	24
Figura 6 – Categorias de ML .....	26
Figura 7 – Diagrama para escolha de algoritmos de ML .....	29
Figura 8 – Etapas para a validação de um classificador no ML .....	31
Figura 9 – Gráfico da desordem em função da relação P/T .....	35
Figura 10 – Definição do tronco da DTC parte A.....	36
Figura 11 – Definição do tronco da DTC parte B.....	36
Figura 12 – Definição do ramo de clima ensolarado da DTC parte A .....	37
Figura 13 – Definição do ramo de clima ensolarado da DTC parte B .....	38
Figura 14 – Definição do ramo de clima chuvoso da DTC parte A.....	38
Figura 15 – Definição do ramo de clima chuvoso da DTC parte B.....	39
Figura 16 – Árvore de decisão final obtido da DTC.....	39
Figura 17 – Pseudocódigo da DTC .....	40
Figura 18 – Condição inicial para um problema a ser solucionado por SVM .....	42
Figura 19 – Possíveis limites de decisão para o problema no SVM.....	43
Figura 20 – Apresentação do hiperplano do problema no SVM.....	44
Figura 21 – Definição de vetores no espaço vetorial do problema no SVM .....	45
Figura 22 – Definição de vetores para determinar o hiperplano no SVM .....	48
Figura 23 – Parafusos utilizados para o desbalanceamento.....	54
Figura 24 – Parafuso de desbalanceamento acoplado no disco de inércia .....	55
Figura 25 – Seleção de componentes no domínio da frequência.....	58
Figura 26 – Visão geral da componente fundamental na fase T .....	62
Figura 27 – Histograma simplificado da componente fundamental .....	63
Figura 28 – Histograma detalhado da componente 0,5x fundamental .....	63
Figura 29 – Distribuição de probabilidade da componente fundamental.....	64
Figura 30 – Distribuição cumulativa da componente 5,0x fundamental .....	65
Figura 31 – <i>Boxplot</i> simplificado da componente fundamental .....	66
Figura 32 – <i>Boxplot</i> detalhado da componente fundamental .....	66
Figura 33 – Desempenho na detecção de tipo e severidade (ponto).....	70

Figura 34 – Desempenho na detecção de tipo e severidade (elipse).....	71
Figura 35 – Exemplo de critério de avaliação para o melhor algoritmo.....	71

## LISTA DE TABELAS

Tabela 1 – Lista dos componentes utilizados para a instrumentação .....	22
Tabela 2 – Exemplo de algoritmos de ML e aplicações .....	28
Tabela 3 – Matriz de confusão .....	31
Tabela 4 – Conjunto de dados para explicar DTC.....	33
Tabela 5 – Massa de parafusos, roscas e arruelas, em gramas .....	55
Tabela 6 – Massa dos níveis, em gramas, e comparação .....	56
Tabela 7 – Matriz de confusão do algoritmo DTC com critério gini em 5,0 Hz.....	74

## LISTA DE ABREVIATURAS E SIGLAS

DAE – Departamento Acadêmico de Eletrotécnica

DTC – Árvore de Decisão (*Decision Tree Classifier*)

ENAM – Encontro Nacional de Máquinas Rotativas

IFSC – Instituto Federal de Santa Catarina

ISO – Organização Internacional de Normalização (*International Organization for Standardization*)

MCSA – Análise de Assinatura de Corrente do Motor (*Motor Current Signature Analysis*)

ML – Aprendizado de Máquina (*Machine Learning*)

NI – *National Instruments*

OvO – *One-versus-One*

OvR – *One-versus-Rest*

PECCE – Grupo de Pesquisa de Computação Científica para Engenharia

SVM – Máquina de Vetores de Suporte (*Support Vector Machine*)

USB – *Universal Serial Bus*

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>12</b>
<b>1.1</b>	<b>Justificativa</b> .....	<b>12</b>
<b>1.2</b>	<b>Definição do Problema</b> .....	<b>13</b>
<b>1.3</b>	<b>Objetivos</b> .....	<b>14</b>
1.3.1	Objetivo Geral .....	14
1.3.2	Objetivos Específicos .....	14
<b>1.4</b>	<b>Estrutura do trabalho</b> .....	<b>14</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>16</b>
<b>2.1</b>	<b>Desbalanceamentos em rotores</b> .....	<b>16</b>
<b>2.2</b>	<b>Assinatura do transitório de corrente</b> .....	<b>18</b>
<b>2.3</b>	<b>Aprendizado de máquina</b> .....	<b>25</b>
2.3.1	Árvore de decisão (DTC) .....	32
2.3.2	Máquina de vetores de suporte (SVM).....	40
2.3.3	One-versus-One (OvO).....	51
<b>3</b>	<b>DESCRIÇÃO DA METODOLOGIA</b> .....	<b>53</b>
<b>3.1</b>	<b>Aquisição dos sinais</b> .....	<b>53</b>
<b>3.2</b>	<b>Processamento de dados e sua interpretação</b> .....	<b>56</b>
<b>4</b>	<b>RESULTADOS OBTIDOS</b> .....	<b>60</b>
<b>4.1</b>	<b>Sem auxílio de ML</b> .....	<b>60</b>
<b>4.2</b>	<b>Com auxílio de ML</b> .....	<b>67</b>
<b>5</b>	<b>CONCLUSÃO</b> .....	<b>75</b>
	<b>REFERÊNCIAS</b> .....	<b>76</b>
	<b>APÊNDICES</b> .....	<b>80</b>
	APÊNDICE A – Artigo completo .....	81
	APÊNDICE B – Comprovante de apresentação.....	92
	APÊNDICE C – Slides apresentados no congresso.....	93
	<b>ANEXOS</b> .....	<b>104</b>
	ANEXO A – Folder de chamada do congresso .....	105

## 1 INTRODUÇÃO

O diagnóstico eficaz de falhas, por meio da instrumentação de monitoramento em tempo real, pode melhorar a segurança e a confiabilidade na operação de máquinas rotativas (MESHGIN-KELK *et al.*, 2017). Ele também pode fornecer uma base sólida nos fundamentos da causa e efeito da falha. A maioria das técnicas existentes para monitoramento de condições e diagnóstico de falhas é baseada em medições de corrente (THOMSON; CULBERT, 2017), vibração (FINLEY; SAUER; LOUTFI, 2015), campo eletromagnético externo (SANTOS *et al.*, 2017) e/ou temperatura (HEY *et al.*, 2016).

O Grupo de Pesquisa de Computação Científica para Engenharia (PECCE) do Instituto Federal de Santa Catarina (IFSC), Câmpus Florianópolis, desenvolve há alguns anos estudos sobre o diagnóstico preditivo em máquinas elétricas.

O presente trabalho é uma continuação do trabalho de conclusão de curso do aluno Honorato (2017). O mesmo fez uso da análise de assinatura de corrente do motor (*Motor Current Signature Analysis – MCSA*) por meio de ferramentas de processamento de sinais clássicos, a saber, as transformadas de Fourier e Hilbert.

Aqui, a interpretação do MCSA foi feita por meio de algoritmos de aprendizado de máquina (*Machine Learning algorithms – ML*) (RUSSELL, 2018).

### 1.1 Justificativa

A MCSA tem sido usada para monitoramento de condições em aplicações industriais (THOMSON; CULBERT, 2017). Uma vantagem do MCSA em relação a outras técnicas é a simplicidade de sua instrumentação. Não há requisitos para ajustes mecânicos complexos, como os necessários para sensores de vibração mecânica; ou mesmo atenção para a complexidade do ambiente, como a presença de campos eletromagnéticos externos. Entretanto, cabe ressaltar que a interpretação da MCSA não é evidente, devido principalmente aos ruídos inerentes ao sinal em

ambiente industrial. Assim, MCSA necessita de algoritmos complementares que em geral podem ser custosos computacionalmente.

O principal benefício do emprego de algoritmos de ML é a menor dependência da experiência do engenheiro para se obter informações sobre o problema. A aplicação de técnicas de ML para o diagnóstico de comportamento de máquinas rotativas se tornou alvo de muitas pesquisas nos últimos anos (DINEVA *et al.*, 2019).

É consenso que, quanto mais informações estiverem disponíveis, mais fácil é o gerenciamento da integridade de um sistema.

## 1.2 Definição do Problema

O processo de desalinhamento do rotor de um motor elétrico é um problema relevante e bastante comum na indústria (MATSUSHITA, 2017). Com uma instrumentação de baixo custo e fácil de instalar (MCSA), desenvolveu-se uma metodologia que usa ML para identificar o tipo e a severidade de falhas de desbalanceamento.

De acordo com a ISO 1940-1 (ISO, 2003), existem quatro tipos de falhas de desbalanceamento (estático, dinâmico, acoplado e quase-estático). Identificar três ou mais tipos de ocorrência pode ser chamado de classificação multiclasse.

Qualquer máquina deve ser parada para manutenção caso seja desejável estender a sua vida útil. Quando o equipamento retorna ao estado natural de operação, o transitório da corrente indicará a condição saudável da máquina. Conseqüentemente, recomenda-se um procedimento de ML dito supervisionado: a comparação direta entre dados ditos saudáveis e dados ainda não classificados (MOHAMMED; KHAN; BASHIER, 2017). Entre as técnicas mais relevantes, destaca-se Máquina de Vetores de Suporte (SVM), Árvore de Decisão (DTC) e *One-versus-One* (OvO). Elas se destacam, pois resolvem com eficiência uma classificação multiclasse e não linear. Não linear devido à disparidade da operação das condições do motor, por exemplo, variações de carga e fonte de alimentação.

Este estudo apresenta a metodologia de diagnóstico desenvolvida (MCSA + ML) e realiza um comparativo entre as três técnicas de ML citadas.

## **1.3 Objetivos**

### **1.3.1 Objetivo Geral**

Desenvolver uma metodologia para a identificação do desbalanceamento em rotores de máquinas elétricas rotativa, classificando o tipo e a severidade do problema, através da análise de assinatura de corrente do equipamento com algoritmos de aprendizado de máquina supervisionados multiclases.

### **1.3.2 Objetivos Específicos**

As alíneas abaixo mostram os objetivos específicos que o estudo visa buscar ao longo de seu desenvolvimento.

- a) Adquirir os sinais de corrente e criar um conjunto de dados que bem represente o desbalanceamento na bancada experimental;
- b) Estudar os algoritmos de ML e escolher três que lidem bem com as características do problema apresentado;
- c) Preparar o conjunto de dados através da filtragem dos parâmetros não relevantes para a utilização dos algoritmos de ML escolhidos;
- d) Aplicar os algoritmos de ML escolhidos nos dados filtrados;
- e) Avaliar o desempenho apresentado pelos algoritmos e a contribuição do estudo proposto para a engenharia de manutenção.

## **1.4 Estrutura do trabalho**

Este texto foi dividido em quatro capítulos. O primeiro capítulo é a introdução do estudo, visando apresentar as motivações e os objetivos para o trabalho. O segundo capítulo apresenta a fundamentação teórica necessária para o entendimento do conteúdo. Em essência, este capítulo apresenta os conceitos básicos a respeito de desbalanceamento em máquinas elétricas rotativas e, posteriormente, ele apresenta os conceitos relacionados com o aprendizado de máquina. O terceiro capítulo mostra as metodologias utilizadas para atender os

objetivos. O quarto capítulo apresenta os resultados obtidos e reflexões sobre o problema e as MLs. Por fim, apresentam-se as conclusões finais, propostas de continuação dos estudos e também informações a respeito da participação deste estudo, na forma de artigo científico, no IX Encontro Nacional de Máquinas Rotativas (ENAM).

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste tópico são apresentados os principais conceitos que são necessários para o entendimento do estudo proposto.

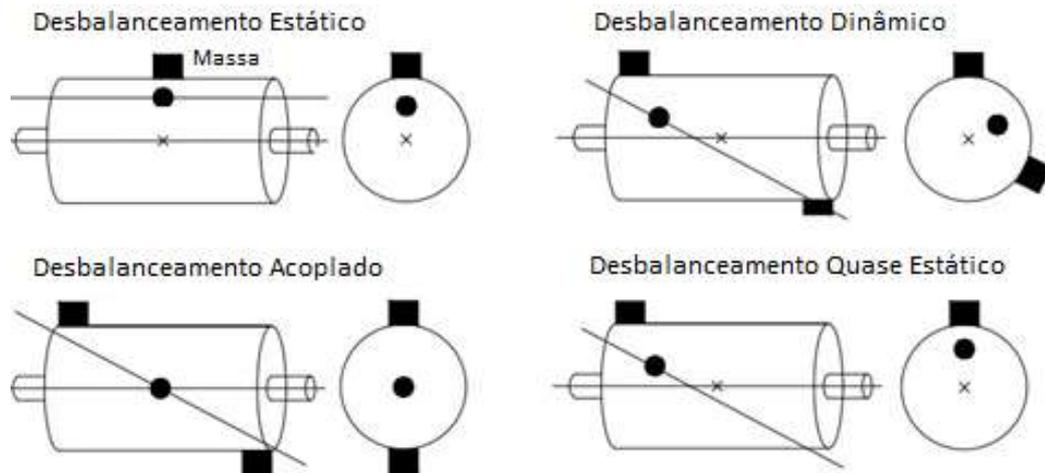
### 2.1 Desbalanceamentos em rotores

Define-se desbalanceamento como o desequilíbrio de massa existente com relação ao eixo de rotação do motor. Alternativamente, o desbalanceamento é a distância existente entre o centro de massa com o eixo de rotação do motor. Assimetrias, tolerâncias dos desvios de forma e imperfeições da montagem são responsáveis pela existência do desbalanceamento em motores (ISO, 2003).

Existem normas que delimitam limites de desbalanceamento, o número necessário de planos de correção e métodos para mensurar essa grandeza. A Organização Internacional de Normalização (ISO) possui uma norma cujo objetivo é a mensuração da vibração mecânica decorrida de desbalanceamento, a ISO 1940-1:2003. Esta norma faz recomendações sobre os requisitos de qualidade de balanço para rotores em um estado constante (rígido), de acordo com o tipo de maquinaria e a velocidade máxima de serviço.

Existem quatro tipos de desbalanceamento quando é levada em consideração a posição relativa entre os eixos de rotação e de massa (MATSUSHITA *et al.*, 2017): estático, quase estático, acoplado e dinâmico. A Figura 1 representa esses tipos de desbalanceamentos.

**Figura 1 – Tipos de desbalanceamento em rotores**



Fonte: Honorato (2017).

No desbalanceamento estático, existe um deslocamento paralelo do eixo de massa com relação ao eixo de rotação.

Esta denominação se deve ao fato de que quando a máquina é desligada, o eixo tende a girar até que o ponto de maior massa esteja para baixo, tornando-o facilmente detectável. Este desbalanceamento pode ser facilmente corrigido com a adição de uma massa de balanceamento diametralmente oposta ao centro de gravidade em um plano perpendicular ao eixo de rotação (HONORATO, 2017, p. 35).

O desbalanceamento acoplado é caracterizado por fazer um deslocamento não paralelo do eixo de massa com relação ao eixo de rotação. Todavia, para receber essa classificação, os eixos devem interceptar no centro de gravidade do rotor.

Este desbalanceamento é equivalente a duas massas em dois planos diferentes, deslocadas de  $180^\circ$  entre si. A amplitude desse desbalanceamento é determinada pelas massas responsáveis pelo desbalanceamento, pelo raio existente entre o eixo de rotação e as massas e pela distância entre elas. Este desbalanceamento pode ser corrigido com a adição de duas massas de balanceamento também em planos opostos e deslocadas de  $180^\circ$  entre si, de tal forma que as massas fiquem diametralmente opostas (HONORATO, 2017, p. 36).

Para ser classificado como um desbalanceamento dinâmico, o eixo de massa não pode interceptar o eixo de rotação no centro de gravidade e nem ser paralelo ao mesmo.

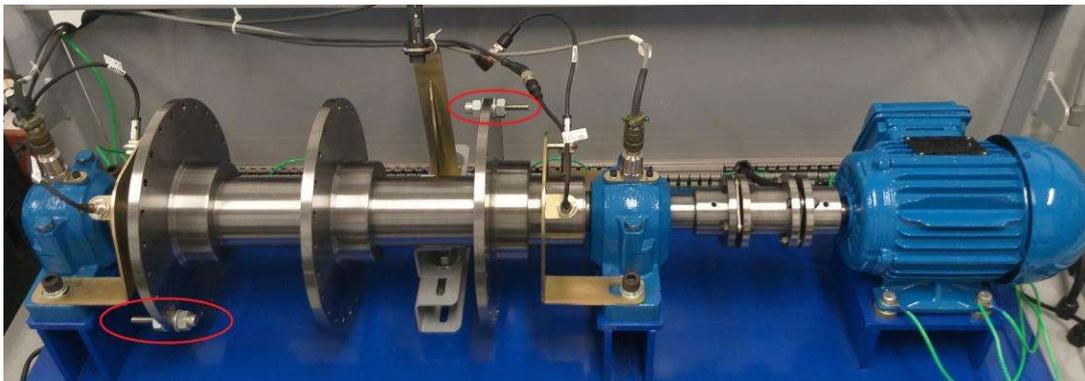
Este é o desbalanceamento mais comum, e pode ser corrigido com a adição de pelo menos duas massas em planos perpendiculares ao eixo de rotação. Em um rotor real, é comum que ocorra uma série de pequenos desbalanceamentos que podem ser definidos por um desbalanceamento dinâmico equivalente. Este desbalanceamento pode ser representado pela

combinação de um desbalanceamento estático e um acoplado (HONORATO, 2017, p. 36).

O desbalanceamento quase estático é uma variante do desbalanceamento dinâmico em que o eixo de massa é interceptado pelo eixo de rotação em um ponto diferente do centro de gravidade (HONORATO, 2017).

O PECCE possui uma bancada que permite o estudo controlado de desbalanceamentos em máquinas elétricas rotativas. As especificações mecânicas de bancada podem ser encontradas na dissertação de mestrado do aluno IFSC (MATSUO, 2017). A Figura 2 apresenta a parte mecânica da bancada. Os círculos em vermelho destacam as massas que emulam o comportamento do desbalanceamento do tipo acoplado.

**Figura 2 – Parte mecânica da bancada – desbalanceamento acoplado**



Fonte: Elaboração Própria (2020).

## 2.2 Assinatura do transitório de corrente

Aqui se apresenta toda a instrumentação que foi utilizada para o desenvolvimento deste estudo. Devido a sua natureza, este tópico foi dividido em duas partes: instrumentação mecânica, voltada para medição de grandezas de natureza mecânica, tais como distância e vibração; e instrumentação elétrica, cuja função é mensurar as grandezas de natureza elétrica, tais como as correntes nos enrolamentos do motor de indução.

Foi optado por utilizar o dispositivo de modelo NI USB-6212 do fabricante *National Instruments* (NI) (NATIONAL INSTRUMENTS, 2009). Este equipamento é um módulo de aquisição de dados multifuncional, alimentado via *Universal Serial Bus* (USB), otimizado para oferecer exatidão com taxas de amostragem altas. O

módulo oferece 16 entradas analógicas com 4 faixas de entrada programáveis ( $\pm 0,2$  V a  $\pm 10$  V) com resolução de 16 bits, taxa de amostragem de 400 kS/s, 2 saídas analógicas e 32 linhas de E/S digital.

O uso deste módulo permite a utilização das ferramentas de instrumentação virtual do *software LabVIEW*, o qual apresenta integração simplificada com o módulo, facilitando os processos de aquisição de dados (NATIONAL INSTRUMENTS, 2020b). Este é um *software* de engenharia de sistemas criado para aplicações envolvendo teste, medição e controle, com rápida integração ao hardware.

Para a medição de grandezas de natureza mecânica foram utilizados:

- a) Acelerômetro: Foram utilizados dois destes pertencentes ao modelo 601A01 do fabricante IMI Sensors (IMI SENSORS, 2020). Cada um destes foi instalado diretamente sobre os mancais da bancada. Este sensor é capaz de detectar acelerações de até  $490 \text{ m/s}^2$ , que é um valor muito superior em relação aos valores esperados. A sensibilidade de  $10,2 \text{ mV}/(\text{m/s}^2)$  indica que o sinal de saída deste sensor pode não ser interpretado apropriadamente pelo dispositivo de aquisição de dados, conforme mostrado anteriormente. Sendo assim, o sinal gerado de cada um destes sensores é condicionado por um condicionador de sinais do modelo MACXMCR-UI-UI-UP do fabricante Phoenix Contact (PHOENIX CONTACT, 2020). O mesmo foi configurado através de suas chaves DIP de tal forma a converter o sinal para uma faixa de medição de 0 a 5 V;
- b) Sensor indutivo: Foram utilizados dois destes sensores pertencentes ao modelo M12MG2-ICC20B-BP00,2-GS06 do fabricante Balluff para cada mancal presente na bancada, totalizando quatro sensores, de forma que estes ficassem ortogonais entre si (BALLUFF, 2020a). A faixa de medição é expressa pelo intervalo de (0,5 a 2) mm. Como os dispositivos estão próximos do eixo, essa faixa de medição limitada não é vista como um problema para a bancada. O sinal de saída varia de (4 a 20) mA, ou seja, será utilizado um resistor de tal forma a

permitir a leitura dessa grandeza por parte do dispositivo de aquisição de dados;

- c) Sensor óptico: Foi utilizado um único sensor óptico pertencente ao modelo BOS 18KF-PA-1LQP-S4-C do fabricante Balluff (BALLUFF, 2020b). Esse sensor foi instalado em um suporte que direciona a parte sensora do dispositivo para a parte girante da bancada. Uma fita reflexiva foi inserida na parte girante de tal forma que o sensor consiga identificar a conclusão de uma revolução. Esse sensor faz o controle de uma chave, cuja frequência de comutação é de até 1500 Hz. Como o motor da bancada visa operar em baixa rotação, não existe nenhum problema com relação à frequência de comutação;
- d) Fonte de alimentação: Foi utilizada uma fonte de alimentação de modelo MPL-1303M do fabricante Minipa (MINIPA, 2020). Esse dispositivo foi ajustado para operar como fonte de tensão de 24 V, uma vez que esse valor permite o funcionamento de todos os componentes pertencentes da parte mecânica.

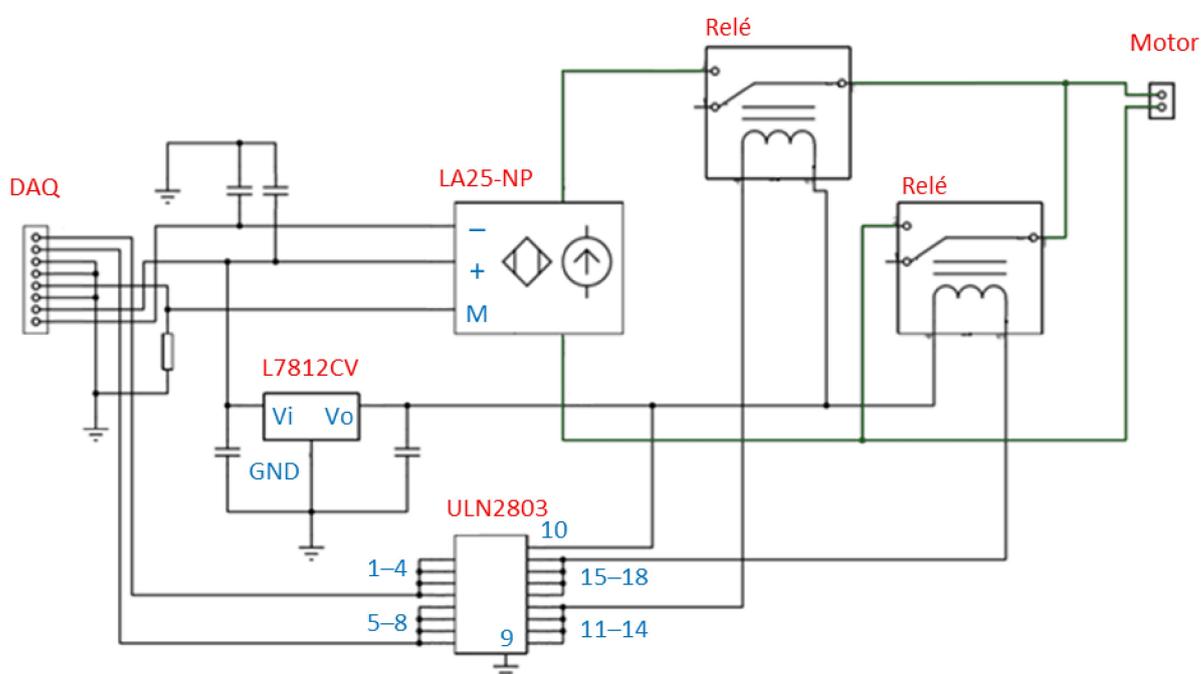
Para a medição de grandezas de natureza elétrica foram utilizados:

- a) Transdutor de corrente: Foi utilizado um transdutor de corrente para cada fase do motor de indução trifásico de modelo LA25-NP do fabricante LEM (LEM, 2020a). Apresenta erro máximo de  $\pm 0,5\%$  na temperatura de 25 °C, tempo de resposta para 90% da corrente nominal  $< 1\ \mu\text{s}$  e possui faixa de medição variável conforme a ligação de seus terminais, variando de (5 a 25) amperes. Vale ressaltar que esse transdutor está inserido em uma placa eletrônica composta por outros dispositivos eletrônicos que garantem a qualidade do sinal e a proteção do sistema. O circuito desenvolvido para sua instalação consiste de dois relés que podem isolar o transdutor, ainda permitindo a passagem da corrente, ou colocá-lo em série com a alimentação conforme desejado. Ambos os relés são acionados pelo circuito integrado ULN2803, que é constituído por 8 ligações Darlington de transistores NPN, sendo divididos 4 transistores para cada relé

(MOTOROLA, 2020). O esquemático do circuito da placa eletrônica pode ser observado na Figura 3;

- b) Fonte de alimentação: Foi utilizada uma fonte de alimentação de modelo RT-65C do fabricante Mean Well (MEAN WELL, 2020). Esse dispositivo foi ajustado para operar como fonte de tensão simétrica de 15 V, uma vez que esse valor permite o funcionamento de todos os componentes pertencentes da parte elétrica. Foi optado em utilizar uma fonte de tensão diferente em relação à fonte da parte mecânica para tentar isolar os sistemas elétricos e mecânicos da bancada.

**Figura 3 – Esquemático elétrico da placa do transdutor de corrente**



Fonte: Elaboração Própria (2020).

O resumo das informações mostradas até o momento pode ser observado na Tabela 1, que visa mostrar cada um dos componentes com o seu respectivo código de identificação.

Tabela 1 – Lista dos componentes utilizados para a instrumentação

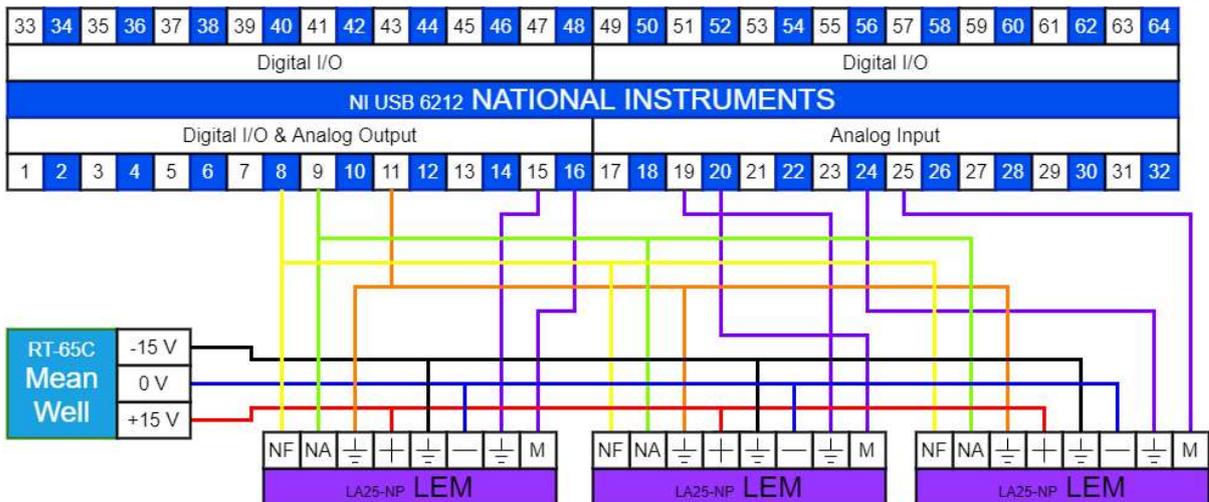
<b>Componente</b>	<b>Código</b>	<b>Fabricante</b>	<b>Modelo</b>
<b>Acelerômetro</b>	M1a45	IMI Sensors	601A01
<b>Acelerômetro</b>	M2a45	IMI Sensors	601A01
<b>Condicionador de Sinal</b>	CS1	Phoenix Contact	MACXMCR-UI-UI-UP
<b>Condicionador de Sinal</b>	CS2	Phoenix Contact	MACXMCR-UI-UI-UP
<b>Dispositivo de Aquisição de Sinal</b>	NI1	National Instruments	NI USB-6212
<b>Dispositivo de Aquisição de Sinal</b>	NI2	National Instruments	NI USB-6212
<b>Fonte de Alimentação</b>	FA1	Minipa	MPL-1303M
<b>Fonte de Alimentação</b>	FA2	Mean Well	RT-65C
<b>Sensor Indutivo</b>	M1d0	Balluff	M12MG2-ICC20B-BP00,2-GS04
<b>Sensor Indutivo</b>	M1d90	Balluff	M12MG2-ICC20B-BP00,2-GS05
<b>Sensor Indutivo</b>	M2d0	Balluff	M12MG2-ICC20B-BP00,2-GS06
<b>Sensor Indutivo</b>	M2d90	Balluff	M12MG2-ICC20B-BP00,2-GS07
<b>Sensor Óptico</b>	Fase	Balluff	BOS 18KF-PA-1LQP-S4-C
<b>Transdutor de Corrente</b>	TC1	LEM	LA25-NP
<b>Transdutor de Corrente</b>	TC2	LEM	LA25-NP
<b>Transdutor de Corrente</b>	TC3	LEM	LA25-NP

Fonte: Elaboração Própria (2020).

A Figura 4 mostra a atual configuração para os transdutores de grandezas elétricas. Foi utilizado um critério de cores na fiação para facilitar o entendimento do diagrama. As cores em vermelho, azul e preto representam, respectivamente, +15 V, GND e -15 V fornecidos por uma fonte de tensão simétrica. A cor roxa representa o

sinal dos transdutores, a qual a placa de aquisição de dados deve interpretar essas informações. As cores em amarela e verde se tratam de um sinal binário, gerado pela placa de aquisição de dados, para comandar os relés presentes nos transdutores de corrente. A função desses relés é isolar o transdutor de corrente do circuito de potência caso o operador desejar.

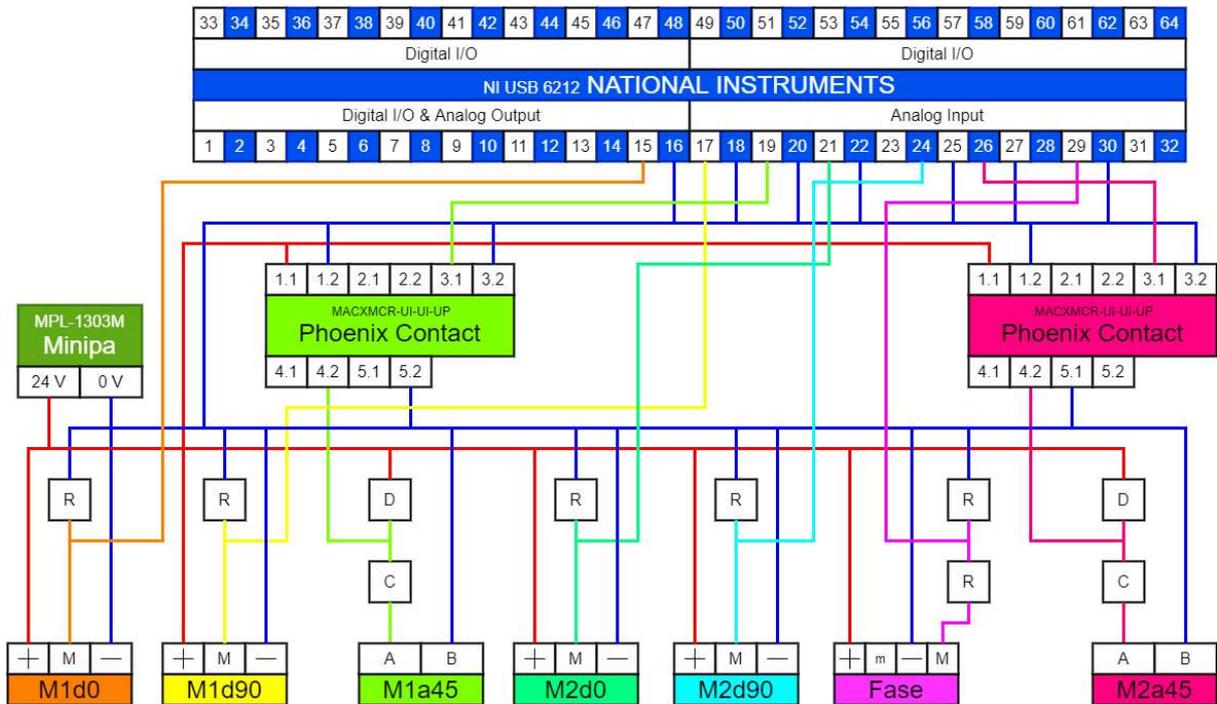
**Figura 4 – Diagrama de ligação dos sinais elétricos**



Fonte: Elaboração Própria (2020).

A Figura 5 mostra a atual configuração para os transdutores de grandezas mecânicas. Assim como no diagrama para grandezas elétricas, aqui também foi utilizado um critério de cores na fiação para facilitar o entendimento do diagrama. As cores em vermelho e azul representam, respectivamente, +24 V e GND fornecidos pela fonte de tensão, que é diferente ao que foi mostrada anteriormente. As demais cores representam os sinais que são captados pelos sensores de grandezas mecânicas. Vale ressaltar que: para os sensores indutivos, é utilizado um único resistor para converter o sinal de corrente gerado pelo mesmo para uma queda de tensão proporcional; para o sensor óptico, é utilizado um divisor de tensão para que o sinal possa ser interpretado pela placa da *National Instruments* (NI); para os acelerômetros, é utilizado um condicionador de sinal para cada um deles para converter o sinal gerado pelos sensores para uma grandeza que pode ser interpretada pela placa de aquisição de dados.

**Figura 5 – Diagrama de ligação dos sinais mecânicos**



Fonte: Elaboração Própria (2020).

Uma vez apresentadas as atuais condições e especificações técnicas da instrumentação presente na bancada, o estudo ainda permite fazer uma breve comparação com o trabalho executado pelo Honorato (2017).

Honorato (2017) desenvolveu uma metodologia para a detecção de desbalanceamentos em rotores de máquinas rotativas com base na utilização de apenas um dos transdutores de corrente. Sendo assim, a análise foi efetuada em apenas uma das fases do motor trifásico. Ele focou em estudos a respeito de falhas típicas em motores de indução trifásicos, em normas que delimitam os tipos de desbalanceamentos existentes com seus níveis aceitáveis e nos ensaios para a aquisição de dados do transdutor de corrente.

O estudo proposto aproveita todas as informações obtidas por Honorato (2017), todavia, aqui, também visa apresentar inovações no quesito da aquisição de dados através da instrumentação utilizada e também na apresentação de metodologias possíveis para a análise dos dados coletados.

Com base nas características do problema apresentado, foi feita uma análise para identificar os algoritmos de aprendizado de máquina que podem trazer

resultados satisfatórios no quesito de apresentar um diagnóstico de desbalanceamento em rotores.

### **2.3 Aprendizado de máquina**

O aprendizado de máquina pode ser visto como uma categoria de algoritmos dentro do que pode ser tido como inteligência computacional, a qual consiste no desenvolvimento de um algoritmo computacional capaz de executar uma tarefa sem a utilização de instruções explícitas. Esses algoritmos executam as tarefas com bases em padrões e inferências de um conjunto de dados (SHWARTZ; DAVID, 2014).

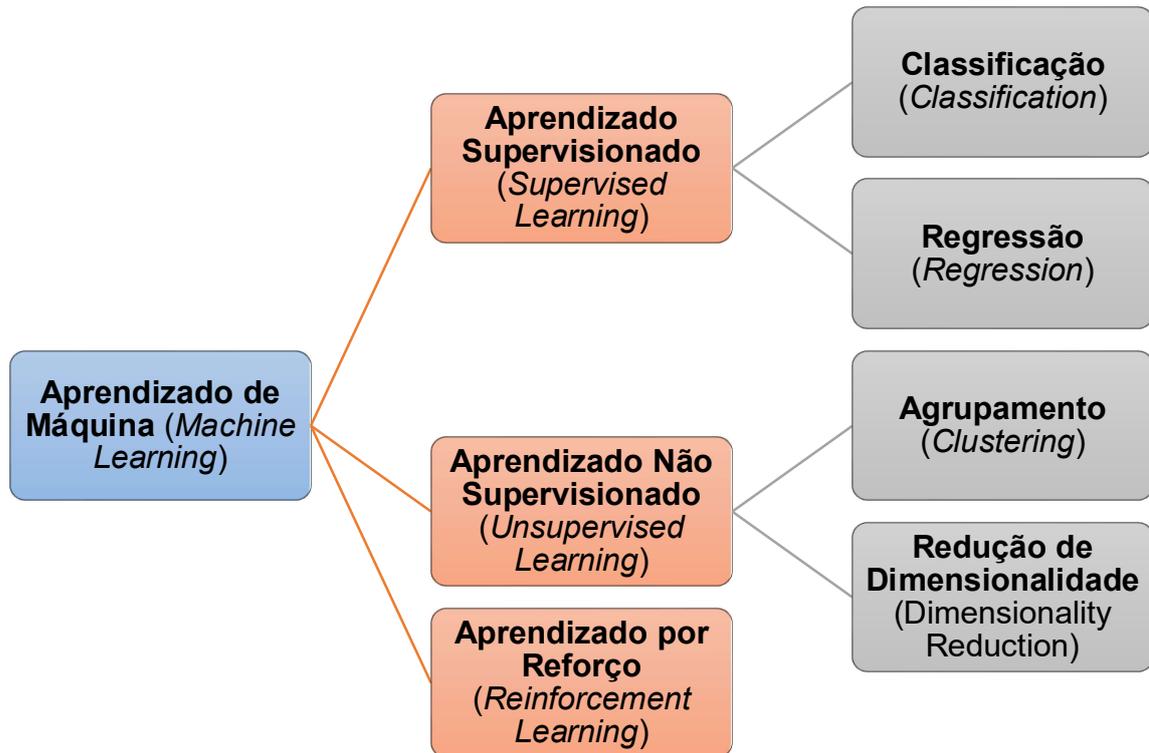
O ML possui sua aplicabilidade quando se encontram problemas cujas tarefas são difíceis de serem interpretadas diretamente. Nos casos que o problema consiste em apenas algumas tarefas simples de programar manualmente, a utilização do ML pode não ser a melhor solução para o problema.

O ML pode ser classificado em três tipos distintos em termos de abordagem: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço (MOHAMMED; KHAN; BASHIER, 2017). Outros autores defendem que existem mais tipos de ML, todavia, os três tipos mencionados são os que são mais abordados pelos mesmos.

A Figura 6 mostra as categorias possíveis para o ML. O termo em inglês é apresentado juntamente com o nome de cada categoria, pois esses são amplamente utilizados dentro da área em estudo. Também, são apresentadas algumas categorias para o aprendizado supervisionado e não supervisionado por essas também serem amplamente comentadas pelos especialistas que estudam o assunto.

A proposta para esse tópico é apresentar superficialmente todas as categorias presentes na Figura 6. Com isso, é possível obter uma visão geral a respeito do tema e ainda permite justificar a escolha dos algoritmos que são utilizados posteriormente pelo estudo.

Figura 6 – Categorias de ML



Fonte: Elaboração Própria (2020).

No aprendizado supervisionado, o conjunto de dados para o treinamento do algoritmo é rotulado. Entende-se dado rotulado quando o valor da saída para uma determinada entrada é conhecida, ou seja, em outras palavras, a classificação para uma determinada amostra é conhecida. Sendo assim, o algoritmo dito como supervisionado visa a desenvolver padrões e inferências com base em exemplos apresentados por um supervisor. O supervisor, que pode ser uma pessoa ou outra máquina, tem como objetivo ensinar a máquina com esse conjunto de dados rotulados (RUSSELL, 2018).

No aprendizado não supervisionado, não existe um supervisor nem dados de treinamento. O conjunto de dados que é utilizado pelo algoritmo é não rotulado, ou seja, não existe uma relação entre entrada e saída. Sendo assim, esse algoritmo visa encontrar padrões ocultos dentro desse conjunto de dados. A ausência dos rótulos para esse aprendizado pode ser causada devido inviabilidade financeira para a rotulagem ou da própria forma de que o conjunto de dados é gerado, que eventualmente impossibilita a rotulagem (RUSSELL, 2018).

O aprendizado por reforço é caracterizado pela ausência de um conjunto de dados. Nesse caso, o algoritmo visa aprender, ou seja, desenvolver regras ou inferências, com base em observações coletadas do ambiente de forma iterativa para executar ações que maximizem a recompensa ou minimizem os riscos. Na iteração, o algoritmo observa o problema e toma uma decisão. Se a decisão tomada atuou de forma positiva ao processo o algoritmo recebe uma recompensa. Caso a decisão tenha atuado de forma negativa ao processo o algoritmo recebe um reforço a fim de melhorar as suas tomadas de decisão (MOHAMMED; KHAN; BASHIER, 2017).

De fato, não existe um algoritmo melhor, que supera todos os outros em termos de desempenho e acurácia, mas sim, existe um algoritmo que melhor se adequa ao problema. Sendo assim, argumentos tais como aprendizado supervisionado é melhor do que o não supervisionado é inválido, a princípio. O mesmo argumento pode ser corrigido caso seja apresentado um contexto que apresente as características de um problema que possibilite tal afirmação.

Visando a existência desse fato, a Tabela 2 apresenta algumas informações extras a respeito das categorias existentes de ML (SMITH, 2020). Nas informações extras incluem alguns dos algoritmos que são pertencentes a uma categoria e também algumas aplicações associadas com a utilização desses algoritmos. Observa-se que é ressaltada a ideia de que todos os algoritmos de ML possuem a sua contribuição na sociedade.

Tabela 2 – Exemplo de algoritmos de ML e aplicações

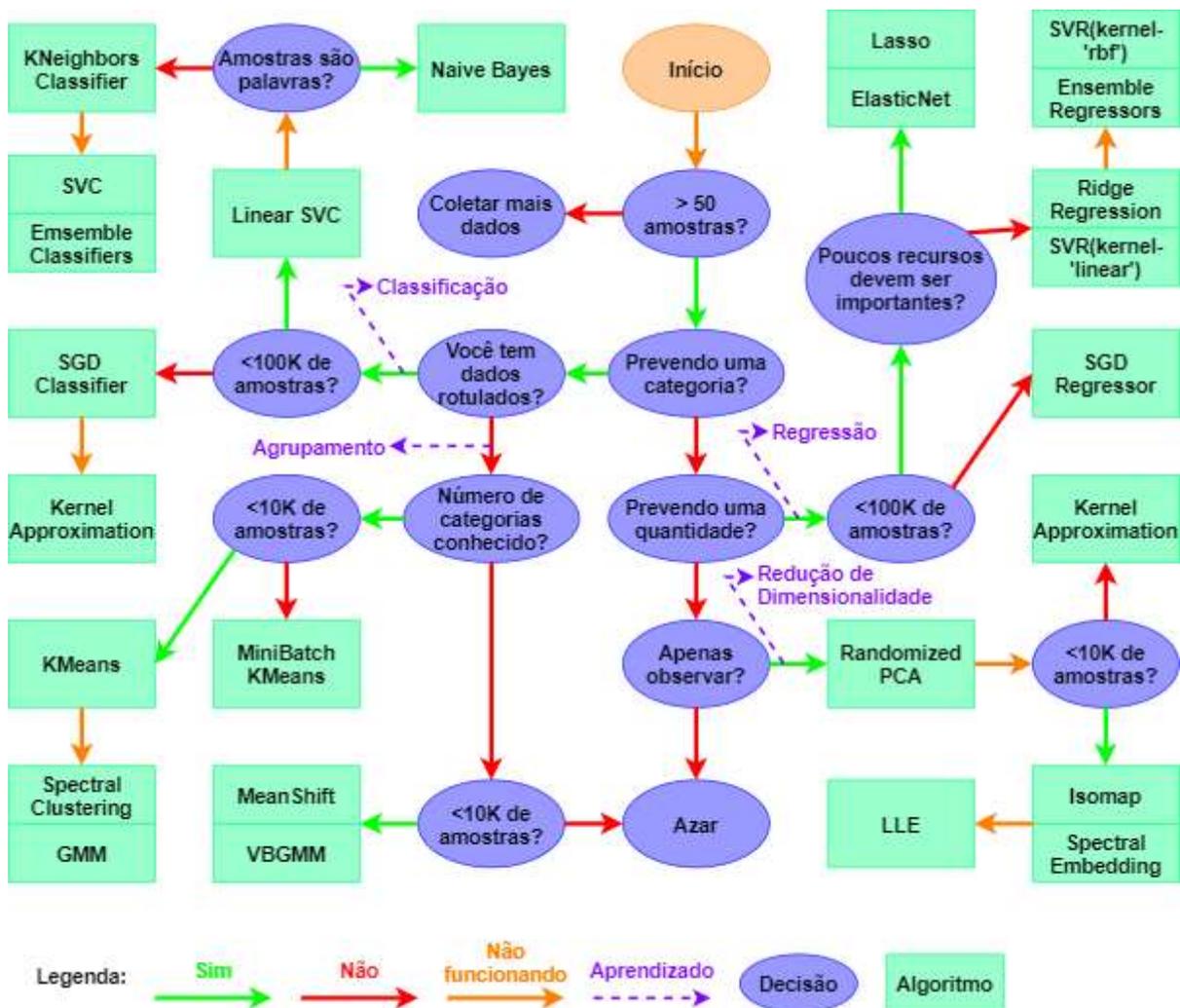
<b>Tipo de Aprendizado</b>	<b>Exemplos de algoritmos</b>	<b>Aplicação</b>
Aprendizado Supervisionado (Classificação)	<ul style="list-style-type: none"> <li>• <i>K-Nearest Neighbor (KNN)</i></li> <li>• <i>Naive Bayes</i></li> <li>• <i>Support Vector Machines (SVM)</i></li> <li>• <i>Decision Tree Classifier (DTC)</i></li> </ul>	<ul style="list-style-type: none"> <li>• Detecção de fraude</li> <li>• Classificação de imagem</li> <li>• Diagnóstico</li> </ul>
Aprendizado Supervisionado (Regressão)	<ul style="list-style-type: none"> <li>• <i>Linear Regression</i></li> <li>• <i>Decision Tree Regression (DTR)</i></li> </ul>	<ul style="list-style-type: none"> <li>• Previsões</li> <li>• Predições</li> <li>• Otimização</li> </ul>
Aprendizado Não Supervisionado (Agrupamento)	<ul style="list-style-type: none"> <li>• <i>K-Means</i></li> <li>• <i>K-Medoids</i></li> </ul>	<ul style="list-style-type: none"> <li>• Sistemas de recomendação</li> <li>• Classificação de clientes</li> <li>• Marketing direcionado</li> </ul>
Aprendizado Não Supervisionado (Redução de Dimensionalidade)	<ul style="list-style-type: none"> <li>• <i>Principal Component Analysis (PCA)</i></li> <li>• <i>Linear Discriminant Analysis (LDA)</i></li> </ul>	<ul style="list-style-type: none"> <li>• Visualização de grandes conjuntos de dados</li> <li>• Compressão de dados</li> <li>• Descobrir estruturas em conjunto de dados</li> </ul>
Aprendizado por Reforço	<ul style="list-style-type: none"> <li>• Q-learning</li> </ul>	<ul style="list-style-type: none"> <li>• Decisões em tempo real</li> <li>• Tarefas de aprendizado</li> <li>• IA para vídeo games</li> </ul>

Fonte: Adaptado de Smith (2020).

Visto a importância existente para todos os algoritmos de ML, é necessário identificar aquele que melhor se adequa para o problema proposto. Vale ressaltar que escolher aleatoriamente algoritmos para fazer testes e concluir que o mesmo não é próprio para a aplicação pode ser uma tarefa custosa. O tempo investido para entender o algoritmo e modelá-lo ao problema proposto pelo estudo pode ser muito grande. É por essa razão que se deve buscar as experiências obtidas em outros estudos relacionadas para a escolha do melhor algoritmo para a aplicação.

Sendo assim, para a escolha dos algoritmos a serem utilizados pelo estudo proposto, além das experiências obtidas em outros estudos que foram feitos no PECCE, também foram utilizados diagramas semelhantes ao apresentado na Figura 7, incluindo o mesmo. Este diagrama está disponível no site da *Scikit learn* (COURNAPEAU, 2020). O mesmo serve como referência para ajudar a escolher o melhor algoritmo para o estudo proposto. Para a escolha do algoritmo, deve-se começar pelo *start*. A partir desse ponto, perguntas objetivas de sim (*yes*) ou não (*no*) devem ser respondidas até chegar uma das quatro categorias existentes no diagrama. As quatro categorias são classificação (*classification*), regressão (*regression*), *clustering* e redução de dimensão (*dimensionality reduction*).

Figura 7 – Diagrama para escolha de algoritmos de ML



Fonte: Adaptado de Courneau (2020).

A escolha do diagrama da Figura 7 foi dada visto que o *Scikit learn* é um módulo programado na linguagem Python e o diagrama serve como uma orientação de como utilizar as funções desse módulo (COURNAPEAU, 2020). O estudo proposto utiliza esse módulo para chegar aos resultados.

Após a análise de outros estudos também realizados no PECCE e também com a utilização desses diagramas, foi observado que o problema proposto deve ser resolvido com aprendizado de máquina supervisionado através da classificação. É por essa razão que esse método recebe mais ênfase nesse estudo com relação ao método de regressão e até mesmo do que o método de agrupamento e redução de dimensionalidade do aprendizado não supervisionado.

O método de classificação, do aprendizado de máquina supervisionado, atribui rótulos para um determinado grupo de amostras não rotuladas com base em um conjunto de dados de treinamento rotulados. O mesmo ainda pode ser classificado quanto a número de classes possíveis. O classificador pode ser binário, quando existem apenas duas classes, ou pode ser multiclasse, quando existem três ou mais classes possíveis.

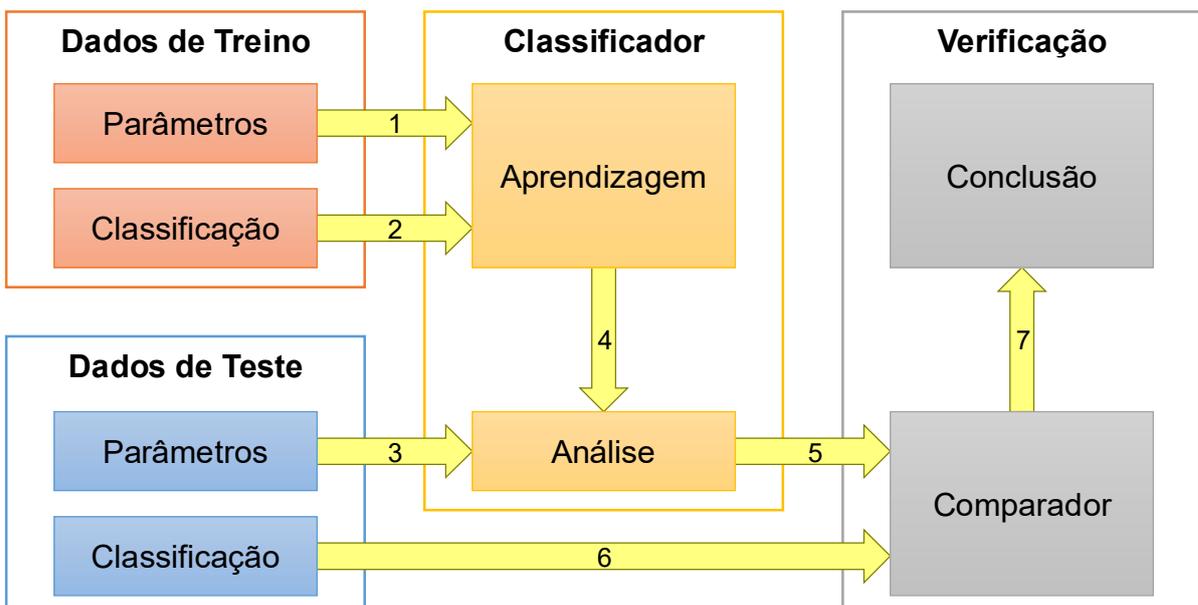
O procedimento que é usualmente utilizado para modelar, treinar e testar um algoritmo pertencente a esse grupo está transcrito na Figura 8. Primeiramente, um conjunto de dados suficientemente grande deve ser coletado para poder utilizar o algoritmo. Este conjunto de dados deve ser dividido em duas partes, os dados de treino e os dados de teste. Recomenda-se fazer uma divisão de tal forma que o conjunto de dados gerado pelos dados de treino seja maior do que o conjunto de dados gerado pelos dados de teste. Uma proporção usual é para cada quatro amostras, três sejam consideradas como um dado de treino e apenas uma como dado de teste (COURNAPEAU, 2020).

Posteriormente, o conjunto de dados de treinamento deve ser utilizado para o processo de aprendizagem do classificador. Observe que tanto os parâmetros (representado pela seta 1) quanto a classificação correta da amostra (seta 2) são utilizados nesse processo. O conjunto de dados de teste deve ser utilizado, posteriormente, para a análise do classificador. Observe que a classificação correta das amostras (seta 6) não é utilizada pelo classificador, apenas é utilizado os parâmetros (seta 3). Internamente ao classificador, o mesmo visa a gerar uma

classificação (seta 5) para cada uma das amostras com base na experiência obtida no processo de aprendizagem (seta 4).

Posteriormente, deve-se fazer uma comparação entre as classificações preditas (seta 5) com o resultado verdadeiro (seta 6) através de um comparador. Um comparador usual é a matriz de confusão, apresentada na Tabela 3. Essa matriz permite mensurar quantos erros ocorreram através de um cruzamento de informações entre o conjunto de classificações preditas com as classificações verdadeiras (RUSSELL, 2018). Define-se falso positivo quando um classificador define uma amostra como positivo, mas na verdade a mesma é negativa. Define-se falso negativo quando um classificador define uma amostra como negativa, mas na verdade a mesma é positiva. A partir dos resultados obtidos do comparador (seta 7), é possível tirar conclusões a respeito da eficácia do algoritmo de classificação para o problema proposto.

**Figura 8 – Etapas para a validação de um classificador no ML**



Fonte: Elaboração Própria (2020).

**Tabela 3 – Matriz de confusão**

Matriz de confusão	Classificação predita	
	Classificação verdadeira	Verdadeiro Positivo
	Falso Positivo	Verdadeiro Negativo

Fonte: Adaptado de Sammut e Webb (2011).

Visto que existem diversos algoritmos dentro do ML e que nem todos são pertinentes ao estudo, são comentados apenas aqueles que são utilizados.

### 2.3.1 Árvore de decisão (DTC)

DTC é conhecida no âmbito do ML como um algoritmo de classificação que permite observar com clareza todas as etapas de decisão utilizadas para distinguir os grupos de amostras. Isso se deve ao fato de que essas etapas podem ser facilmente representadas através de algum diagrama. Esse diagrama é frequentemente comparado com uma árvore real, composta por troncos, ramos e folhas.

A praticidade de se visualizar e interpretar o diagrama obtido da DTC permite que mesmo as pessoas que não possuem muito domínio técnico a respeito da aquisição do conjunto de dados possam interpretar as respostas obtidas a partir desse algoritmo (COURNAPEAU, 2020).

O objetivo da DTC é obter o conjunto de regras mais simples possível para poder classificar corretamente todas as amostras de um conjunto de dados de treinamento. Vale ressaltar que nem sempre essa meta é possível. Para esses casos, o algoritmo visa acertar o maior percentual possível. Essas regras consistem em uma pergunta a respeito de um determinado componente das amostras e com um número limitado de respostas. Entretanto, não necessariamente, as respostas dessas perguntas são exclusivamente sim ou não. Esse caso ocorre usualmente quando é feito comparativo com grandezas numéricas como, por exemplo, perguntar se a temperatura é maior do que 20 °C. Todavia, é possível uma pergunta similar a respeito da temperatura e ter como possíveis respostas: frio, agradável e quente.

A Tabela 4 é utilizada como exemplo para explicar a DTC. Aqui, o conjunto de dados é composto por 14 amostras e simula uma tomada de decisão para o agendamento de uma partida de *golf*. Vale ressaltar que o clima, temperatura, umidade e o vento são parâmetros que são utilizados para a tomada de decisão.

Algumas observações para o conjunto de dados apresentado é que o mesmo é considerado como um conjunto de dados pequeno, uma vez que a DTC é capaz de trabalhar com centenas ou até mesmo milhares de amostras. Todavia, o

tamanho reduzido permite o exemplo ser didático. Outra observação é que existe a ausência de valores numéricos no exemplo. Essa ausência não prejudica na qualidade da explicação do problema. A forma de tratar um valor numérico é similar ao tratar um valor rotulado, com a diferença de que se deve definir um comparativo ou um intervalo para tal operação.

**Tabela 4 – Conjunto de dados para explicar DTC**

<b>Clima</b>	<b>Temperatura</b>	<b>Umidade</b>	<b>Vento</b>	<b>Agendar</b>
Ensolarado	Quente	Alta	Falso	Não
Ensolarado	Quente	Alta	Verdadeiro	Não
Nublado	Quente	Alta	Falso	Sim
Chuvoso	Agradável	Alta	Falso	Sim
Chuvoso	Frio	Normal	Falso	Sim
Chuvoso	Frio	Normal	Verdadeiro	Não
Nublado	Frio	Normal	Verdadeiro	Sim
Ensolarado	Agradável	Alta	Falso	Não
Ensolarado	Frio	Normal	Falso	Sim
Chuvoso	Agradável	Normal	Falso	Sim
Ensolarado	Agradável	Normal	Verdadeiro	Sim
Nublado	Agradável	Alta	Verdadeiro	Sim
Nublado	Quente	Normal	Falso	Sim
Chuvoso	Agradável	Alta	Verdadeiro	Não

Fonte: Adaptado de Eureka (2020).

Um conceito importante deve ser discutido: como determinar quantitativamente o quanto as respostas dadas pela regra gerada são homogêneas. O algoritmo da DTC termina, a princípio e se possível, apenas quando todas essas respostas forem completamente homogêneas, ou seja, todas as amostras do conjunto de treinamento são classificadas corretamente (MIT, 2020a). Vale ressaltar que o programador pode acrescentar outros critérios para finalizar o algoritmo, tais

como tolerância, número de iterações ou memória alocada máxima. Quando uma resposta não for totalmente homogênea, o algoritmo procura colocar outra regra a posteriori a fim de atingir o objetivo.

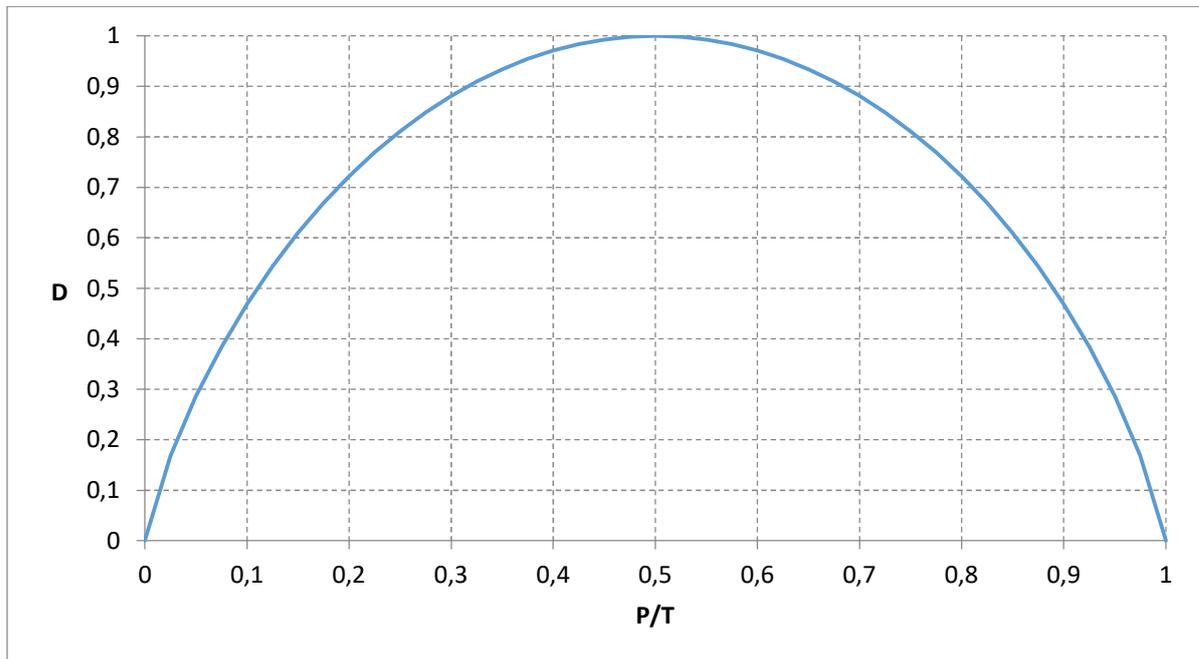
Matematicamente, o conceito a ser apresentado é o de desordem (MIT, 2020a), representado pela letra  $D$  na Equação 1. Os parâmetros  $P$  e  $N$  representam, respectivamente, o número de amostras que pertencem a uma determinada classe e o número de amostras que não pertencem a essa classe. O parâmetro  $T$  representa o número total de amostras, que pode ser obtido a partir da soma de  $P$  e  $N$ .

Nos casos em que a desordem é zero, existe uma indeterminação numérica, todavia, ainda é possível extrair esse valor uma vez que o logaritmo tende ao infinito mais lentamente do que o termo que está fora do logaritmo, que tende a zero. Também, os sinais negativos presentes na Equação 1 servem apenas para compensar o fato de que os logaritmos serem negativos para todos os valores possíveis.

$$D = -\frac{P}{T} \log_2 \frac{P}{T} - \frac{N}{T} \log_2 \frac{N}{T} \quad (1)$$

A desordem é um coeficiente que pode variar de zero até o valor unitário. A Figura 9 mostra o comportamento desse coeficiente em função de  $P$  em relação a um número total de amostras  $T$ . Quando a desordem está em um dos extremos do gráfico, ou seja, igual à zero ou o valor unitário, a mesma indica uma situação em que a resposta é completamente homogênea. Quando a desordem é igual a 0,5, a mesma indica uma situação em que a resposta é completamente heterogênea. A Figura 9 ilustra o valor da desordem.

Figura 9 – Gráfico da desordem em função da relação P/T



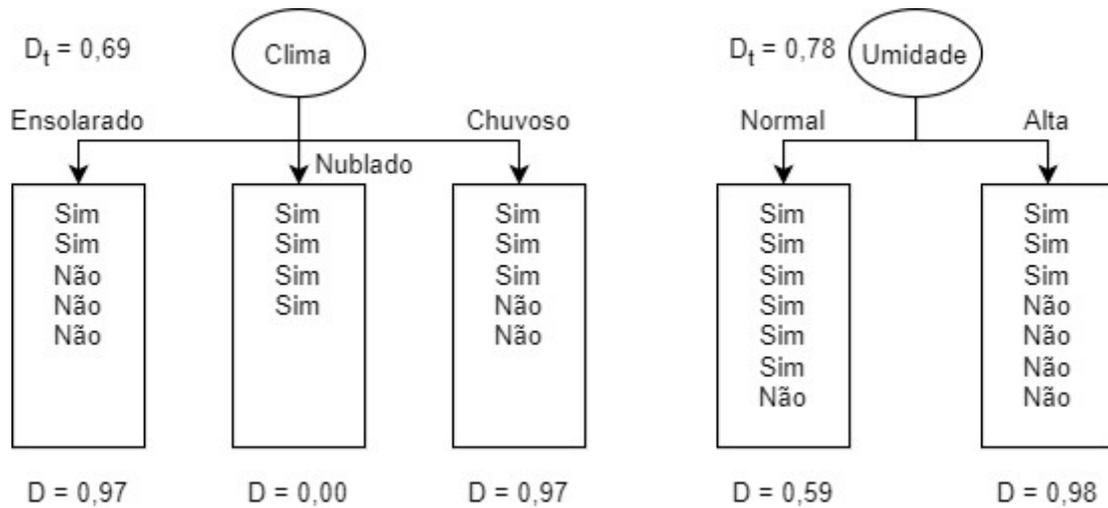
Fonte: Adaptado de MIT OpenCourseWare (2020a).

Eventualmente, também vai ser necessário fazer uma média ponderada  $D_t$  das desordens de uma regra, conforme mostra a Equação 2.  $D_l$  representa a desordem para um determinado rótulo  $l$  (*label*);  $n_l$  representa o número total de amostras para o rótulo  $l$ ;  $l_{max}$  representa o número de rótulos existentes para a regra.

$$D_t = \frac{\sum_{l=1}^{l_{max}} D_l n_l}{\sum_{l=1}^{l_{max}} n_l} \quad (2)$$

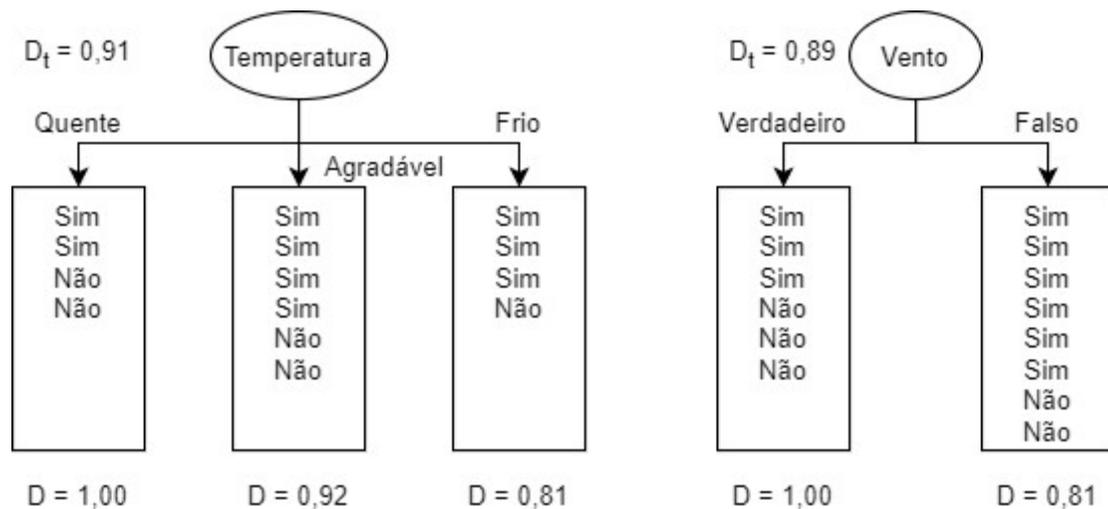
Uma vez apresentado a Equação 1 e a Equação 2, é possível iniciar a explicação da DTC. A primeira etapa desse algoritmo é escolher o tronco da árvore, ou seja, entre as componentes existentes na amostra, escolher aquela que melhor separa os sins dos não, visando homogeneizar todas as respostas possíveis, relacionada com os rótulos de uma determinada componente. A Figura 10 e a Figura 11 mostram os troncos possíveis para a árvore, no exemplo dado. Vale ressaltar que a Equação 1 e Equação 2 já foram utilizadas para obter os valores de  $D$  e  $D_t$ , respectivamente. Para exemplificar como se calcula  $D$ , para o tronco de clima (Figura 10), o rótulo ensolarado possui uma desordem de 0,97 porque  $P$  é igual a 2 por existir dois sins,  $N$  é igual a 3 por existir três não e consequentemente,  $T$  é igual a 5.

**Figura 10 – Definição do tronco da DTC parte A**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

**Figura 11 – Definição do tronco da DTC parte B**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

Feito os cálculos, a escolha do tronco é dada pela menor desordem total. Para o exemplo, é escolhido o clima. Isso finaliza a primeira iteração do algoritmo. Essa iteração permite concluir que o clima é o melhor parâmetro para iniciar a classificação do conjunto de dados. Em contrapartida, também é possível concluir que a temperatura é o pior parâmetro para iniciar a classificação do conjunto de dados por possuir a maior desordem total.

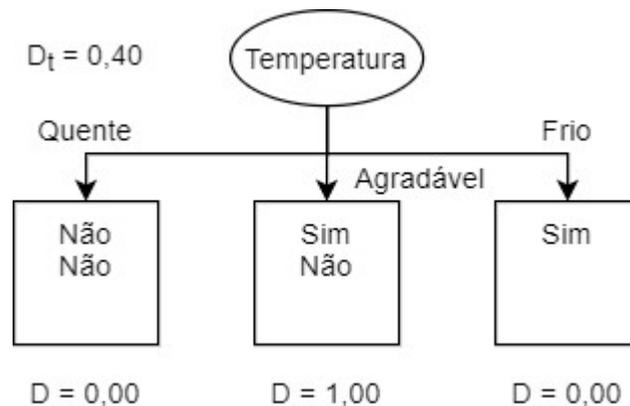
A partir de agora, não são definidos mais troncos, mas sim ramos e folhas. Para cada rótulo não homogêneo existente no tronco de clima (Figura 10), deve-se inserir uma nova regra, sendo essa nomeada de ramo. Em contrapartida,

para cada rótulo homogêneo, deve-se inserir a classificação a qual está associada ao rótulo, uma vez que não é necessária a inserção de uma nova regra. Nesse caso, a árvore ganha uma folha.

Sendo assim, no tronco escolhido, devem ser adicionados ramos aos rótulos de ensolarado e chuvoso e deve ser adicionada uma folha indicando sim ao rótulo de nublado.

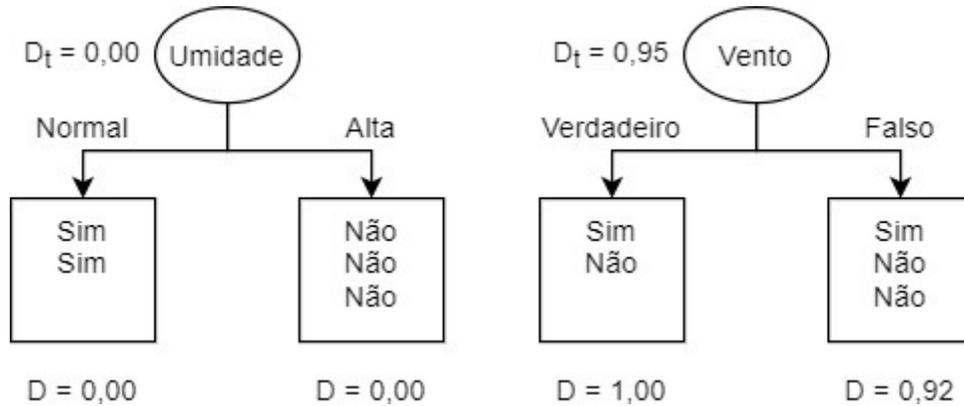
Para cada ramo, deve-se fazer uma análise similar ao que foi feito no tronco. A Figura 12 e a Figura 13 mostram a análise voltada para o ramo de clima ensolarado e a Figura 14 e Figura 15 mostram a análise voltada para o ramo de clima chuvoso. Um detalhe relevante para ser mencionado é que o conjunto de dados passa pelo filtro gerado pelo tronco. Isso justifica o porque existem menos sins e não para essa segunda iteração. Esse padrão de filtragem do conjunto de dados segue de forma análoga para as próximas iterações.

**Figura 12 – Definição do ramo de clima ensolarado da DTC parte A**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

**Figura 13 – Definição do ramo de clima ensolarado da DTC parte B**

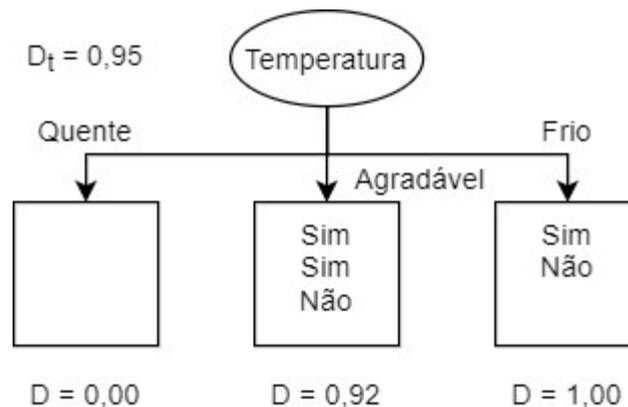


Fonte: Adaptado de MIT OpenCourseWare (2020a).

A respeito da Figura 12 e da Figura 13, é possível concluir que a umidade (Figura 13) é o melhor parâmetro para ser selecionado como ramo. Além disso, é possível visualizar que ambos os rótulos associados à umidade são homogêneos, o que permite a inserção de duas folhas com as respectivas classificações.

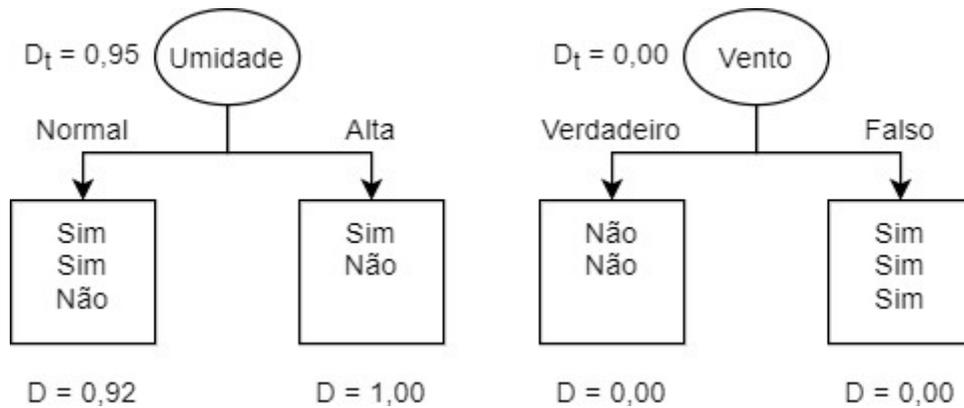
A respeito da Figura 14 e da Figura 15, é possível concluir que o vento (Figura 15) é o melhor parâmetro para ser selecionado como ramo. Além disso, é possível visualizar que ambos os rótulos associados ao vento são homogêneos, o que permite a inserção de duas folhas com as respectivas classificações.

**Figura 14 – Definição do ramo de clima chuvoso da DTC parte A**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

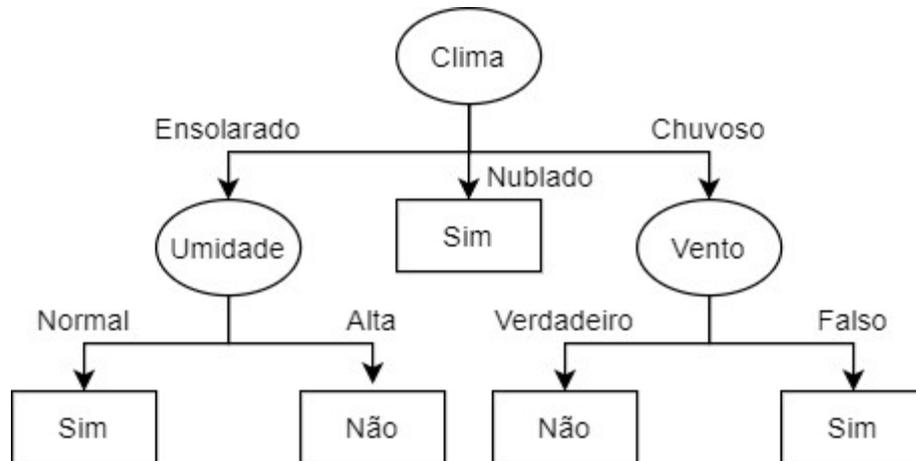
**Figura 15 – Definição do ramo de clima chuvoso da DTC parte B**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

Visto que não há necessidade de inserir mais ramos para poder classificar corretamente todas as amostras do conjunto de treinamento, a Figura 16 mostra o resultado final obtido na DTC. Observe-se que por aqui fica mais fácil de observar que o clima é o tronco da árvore e que a umidade e o vento são ramos da mesma.

**Figura 16 – Árvore de decisão final obtido da DTC**



Fonte: Adaptado de MIT OpenCourseWare (2020a).

A Figura 17 mostra um resumo da explicação da DTC em forma de um pseudocódigo. É evidente que, por se tratar de um pseudocódigo, existem diversas simplificações no código, inclusive a desconsideração de algumas variáveis secundárias que são importantes para o funcionamento da DTC.

Figura 17 – Pseudocódigo da DTC

```

function Arvore = DTC(S)
    % S, um conjunto de dados com amostras rotuladas
    Arvore = um novo nó vazio
    if todas as amostras tem a mesma classe c or
        nenhuma divisão adicional for possível then
        % criar nova folha
        rotulo(Arvore) = c
    else
        % criar novo tronco
        (P,T) = EncontrarMelhorDivisao(S)
        % P, o parâmetro que possui menor desordem
        % T, todos os rótulos pertencentes ao parâmetro P
        for cada t ∈ T do
            S_t = todas as amostras que satisfazem t
            Arvore_t = DTC(S_t)
            ConectarRamo(Arvore, Arvore_t)
        end
    end
end

```

Fonte: Adaptado de Sammut e Webb (2011).

Neste trabalho, optou-se por utilizar o método DTC do módulo *Scikit learn* (COURNAPEAU, 2020). Esta biblioteca é gratuita, de livre distribuição e permite derivativos.

### 2.3.2 Máquina de vetores de suporte (SVM)

*Support Vector Machine* (SVM) é um algoritmo que se baseia primordialmente em vetores (MIT OpenCourseWare, 2020b). Alguns dos conceitos provenientes da álgebra linear devem estar bem entendidos, pois eles são a base do SVM. Cita-se em específico vetores e operação com vetores. Também são utilizadas certas técnicas de otimização e interpolação dentro do SVM. Relacionar otimização e interpolação com ML é bastante usual quando se trata a análise de algoritmos. Assim, para o melhor entendimento da ferramenta SVM, faz-se necessário conhecer, por exemplo, os multiplicadores de Lagrange (KLEIN, 2020).

Ainda vale ressaltar que todos os princípios de ML, que foram apresentados na seção 2.3, são válidos para o SVM. Em outras palavras, o SVM se

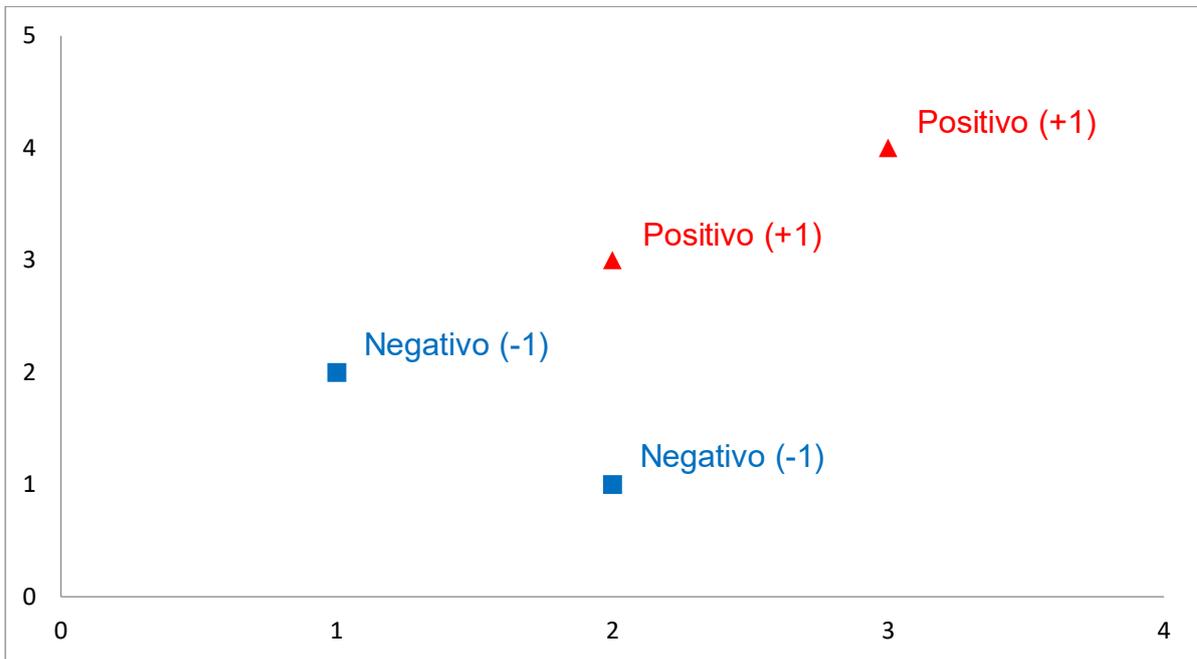
inspira e visa imitar certos aspectos da capacidade cognitiva de um ser vivo, cujo aprendizado é desenvolvido de forma natural ao longo de sua vida.

O principal objetivo do SVM é estabelecer limites de decisão para um determinado conjunto de dados. Em outras palavras, o SVM visa separar um determinado grupo de amostras consideradas como positivas (+1) do restante que são consideradas como negativas (-1). Para esclarecer melhor o que significa estes sinais, o sinal positivo seria o equivalente a afirmar que a amostra em questão pertence a uma determinada classe, ou seja, afirmar que uma determinada amostra pertence à classe da condição saudável da máquina rotativa, por exemplo. O sinal negativo, por outro lado, indica que a amostra em questão não pertence à classe. A utilização dessa nomenclatura para as amostras é bastante usual entre os diversos materiais acadêmicos para a explicação da teoria de SVM.

Essas amostras devem ser definidas dentro de um espaço vetorial. Para o contexto de SVM, o espaço vetorial é um espaço que contém todos os eixos necessários para definir todas as amostras. Cada eixo representa uma variável que ajuda a descrever a amostra.

O limite de decisão que o SVM visa determinar é conhecido como hiperplano. O hiperplano permite estimar uma classificação para futuras amostras. Para ajudar a entender os conceitos apresentados até o momento e, posteriormente, entender o funcionamento do SVM, a Figura 18 ilustra um problema exemplo em suas condições iniciais. O problema exemplo inicia-se apresentado todas as amostras. As amostras positivas são representadas por triângulos vermelhos e as negativas com quadrados azuis.

**Figura 18 – Condição inicial para um problema a ser solucionado por SVM**

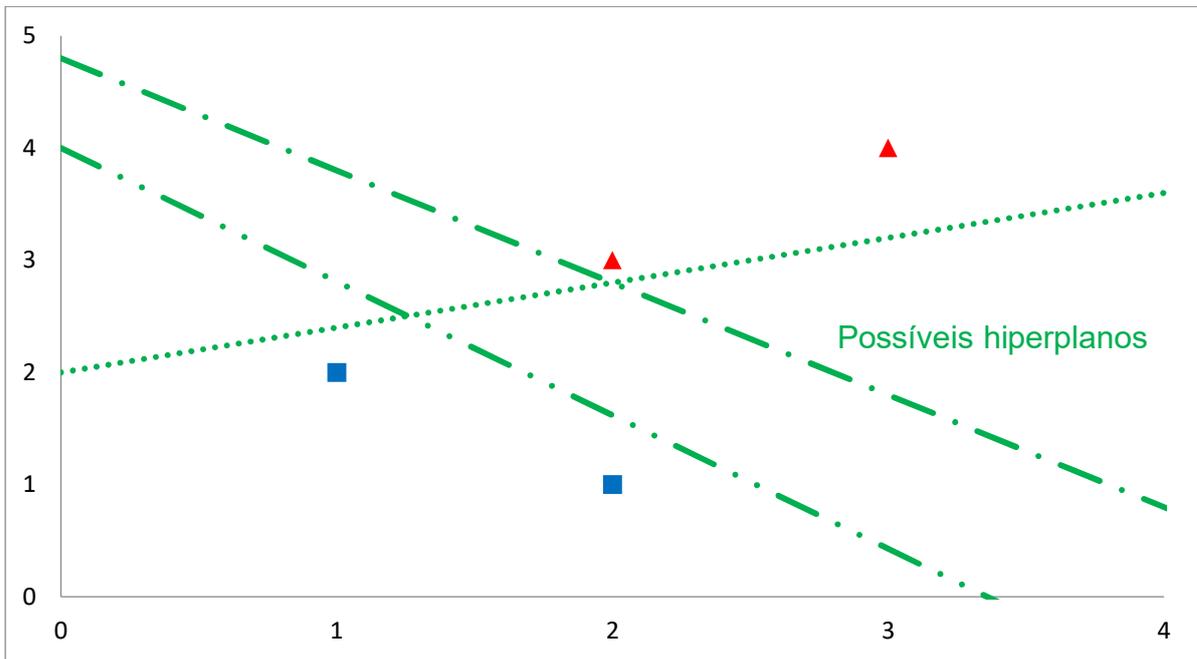


Fonte: Adaptado de MIT OpenCourseWare (2020b).

Uma vez apresentado um conjunto de dados simples, a Figura 19 mostra que existem infinitos limites de decisão possíveis que podem ser utilizados para separar as amostras positivas das negativas, embora seja apresentado apenas três. Para ser considerado como uma possível solução, o limite deve separar todas as amostras de tal forma a não permitir que amostras positivas e negativas fiquem no mesmo lado do limite.

Todavia, o SVM visa em buscar o limite que melhor separa as amostras. Para garantir essa condição de otimalidade, o limite deve estar o mais distante possível de qualquer amostra existente no conjunto de dados. Sendo assim, para cada um dos três limites apresentados é possível visualizar que os mesmos não separam as amostras da melhor forma possível. O limite representado por uma linha pontilhada está muito próximo de uma amostra positiva e também de uma amostra negativa. O limite representado por traço e ponto garante uma boa distância com relação às amostras negativas, todavia, o mesmo ainda fica muito próximo de uma amostra positiva. O limite representado por traço e dois pontos se mostra ser melhor com relação às outras duas apresentadas, todavia, o mesmo não é a melhor solução para o problema.

**Figura 19 – Possíveis limites de decisão para o problema no SVM**



Fonte: Adaptado de MIT OpenCourseWare (2020b).

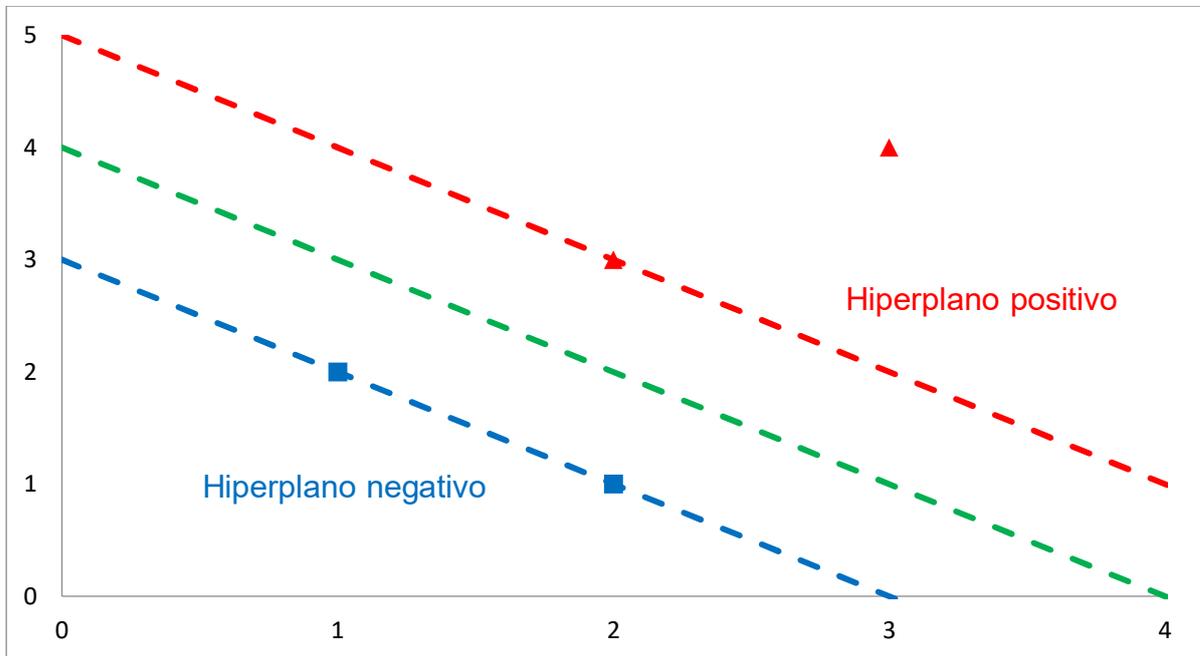
O melhor limite de decisão para o problema é representado pela Figura 20 em cor verde. Observe que, diferente dos outros limites de decisão, esse limite de decisão mantém um equilíbrio entre as amostras positivas e negativas, ao mesmo tempo em que a distância do mesmo até qualquer amostra é maximizada. Vale ressaltar que ainda existem mais duas curvas, em vermelho e azul, que estão em paralelo ao hiperplano e que interceptam as amostras positivas e negativas, respectivamente.

Essas duas curvas possuem a sua utilidade durante a classificação de novas amostras. Para o exemplo retratado na Figura 20, uma amostra que está à direita da curva vermelha é classificada como positiva e uma amostra que está à esquerda da curva azul é classificada como negativa. Essas premissas ainda são reforçadas posteriormente através da apresentação de equações matemáticas.

De fato, esse problema (com quatro amostras e duas classes) foi resolvido facilmente através de uma análise gráfica e de forma manual. Todavia, a complexidade do problema é aumentada quanto mais amostras e classes existirem. Ainda, a obtenção de uma solução – através de análise gráfica e de forma manual – pode ficar ainda mais difícil caso as amostras sejam definidas em um espaço vetorial de três dimensões e impossível para espaços superiores a três. Isso justifica a

necessidade de uma modelagem matemática e da utilização de um poder de processador de uma máquina computacional.

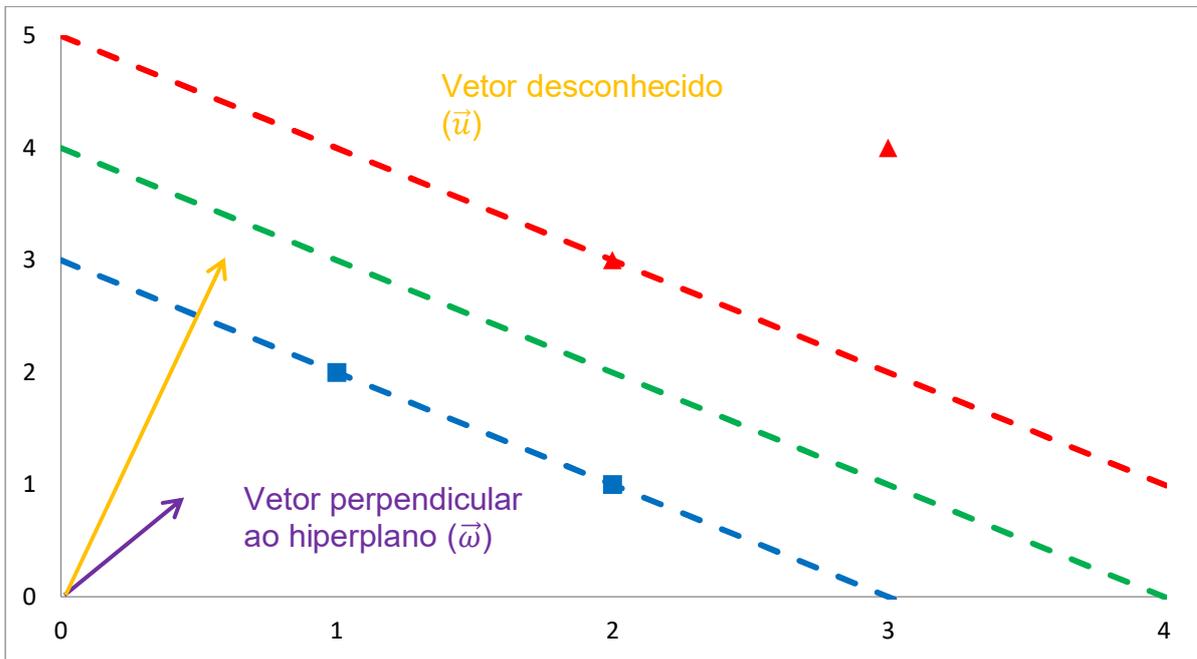
Figura 20 – Apresentação do hiperplano do problema no SVM



Fonte: Adaptado de MIT OpenCourseWare (2020b).

Para a modelagem matemática do SVM, é necessário fazer algumas definições. Para tal, a Figura 21 ajuda a definir dois vetores utilizados no equacionamento. O primeiro é um vetor perpendicular ao hiperplano com um módulo qualquer ( $\vec{w}$ ). O segundo é um vetor desconhecido ( $\vec{u}$ ) o qual deve ser classificado através do SVM. Vale ressaltar que as letras a serem utilizadas durante o equacionamento tem como referência a língua inglesa. Sendo assim, percebe-se que a letra utilizada para o vetor desconhecido vem da palavra inglesa *unknown*.

Figura 21 – Definição de vetores no espaço vetorial do problema no SVM



Fonte: Adaptado de MIT OpenCourseWare (2020b).

Para classificar o vetor desconhecido, deve-se observar que quanto maior for o módulo desse vetor, maior são as chances do mesmo ser classificado como positivo. Também se deve observar que a direção do vetor desconhecido também influencia na classificação da amostra. Sendo assim, é possível inferir que o produto escalar entre os vetores  $\vec{w}$  e  $\vec{u}$  deve ser maior do que uma constante  $c$  para classificar uma amostra como positiva, conforme mostra a Equação 3.

$$\vec{w} \cdot \vec{u} \geq c \quad (3)$$

O produto escalar é um operador vetorial que consiste em um somatório de produtos entre as componentes de dois vetores quaisquer, conforme mostra a Equação 4. Vale ressaltar que, embora seja um operador vetorial, o resultado dessa operação sempre resulta em um número.

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^n a_i \times b_i \quad \forall \vec{u} = (a_1, a_2, \dots, a_n) \wedge \vec{v} = (b_1, b_2, \dots, b_n) \quad (4)$$

De forma alternativa, é possível reescrever a mesma inequação de tal forma a aparecer o termo zero, conforme mostra a Equação 5. É usual a utilização da constante  $b$ , proveniente da palavra inglesa *bias* para tal ação. Vale ressaltar que

$b$  é o negativo de  $c$ . Sendo assim, essa inequação é conhecida como regra de decisão do SVM.

$$\vec{\omega} \cdot \vec{u} + b \geq 0 \quad (5)$$

De fato, essa equação parece estar incompleta, uma vez que  $\vec{\omega}$  e  $b$  são desconhecidos, a princípio. Sendo assim, os próximos equacionamentos visam determinar esses elementos.

Retornando na Figura 20, é possível representar ambos os limites de decisão em vermelho e azul através da inequação apresentada anteriormente, conforme mostra a Equação 6. Para isso, deve-se substituir  $\vec{u}$ , respectivamente, por  $\vec{x}_+$  e  $\vec{x}_-$ .  $\vec{x}_+$  representa uma amostra pertencente ao grupo dos positivos e  $\vec{x}_-$  representa uma amostra pertencente ao grupo dos negativos. Diferente da regra de decisão, a inequação possui  $+1$  e  $-1$ , respectivamente, no lugar de  $0$ . Isso é necessário para poder representar o deslocamento das retas representadas na Figura 20.

$$\begin{aligned} \vec{\omega} \cdot \vec{x}_+ + b &\geq +1 \\ \vec{\omega} \cdot \vec{x}_- + b &\leq -1 \end{aligned} \quad (6)$$

Para continuar o equacionamento do SVM é necessária a inserção de uma variável que represente o comportamento ideal para a tomada de decisão. O comportamento dessa variável depende exclusivamente da amostra, podendo ser um  $\vec{x}_+$  ou  $\vec{x}_-$ , e o valor da mesma é  $+1$  ou  $-1$ , respectivamente, conforme a equação 7.

$$y_i = \begin{cases} +1 \forall \vec{x}_+ \\ -1 \forall \vec{x}_- \end{cases} \quad (7)$$

Quando essa variável é utilizada para multiplicar ambos os lados da desigualdade da regra de decisão aplicadas para  $\vec{x}_+$  e  $\vec{x}_-$ , representada pela Equação 6, é possível obter a Equação 8. Observe-se que a conclusão obtida por ambos os caminhos é a mesma, pois existe a inversão de sinal dada pela Equação 7 para compensar a desigualdades existentes na Equação 6.

$$y_i(\vec{\omega} \cdot \vec{x}_i + b) \geq 1 \quad (8)$$

De forma alternativa, é possível escrever a inequação de tal forma a expor um zero em um dos lados da inequação, conforme mostra a Equação 9. Observe que essa equação consegue representar de forma quantitativa a distância de um  $\vec{x}_i$  qualquer como relação ao hiperplano.

$$y_i(\vec{\omega} \cdot \vec{x}_i + b) - 1 \geq 0 \quad (9)$$

Sendo assim, a inequação anterior permite afirmar um caso específico em que os valores de  $\vec{x}_i$  são pertencentes ao hiperplano, conforme mostra a Equação 10. Retornando a Figura 20, é possível afirmar que as amostras negativas atendem a equação, pois ambas as amostras estão sobre o limite em azul. Todavia, quando se faz uma análise para as amostras positivas, é possível afirmar que apenas uma dessas amostras atende a equação, que é justamente a amostra que está sobre o limite em vermelho.

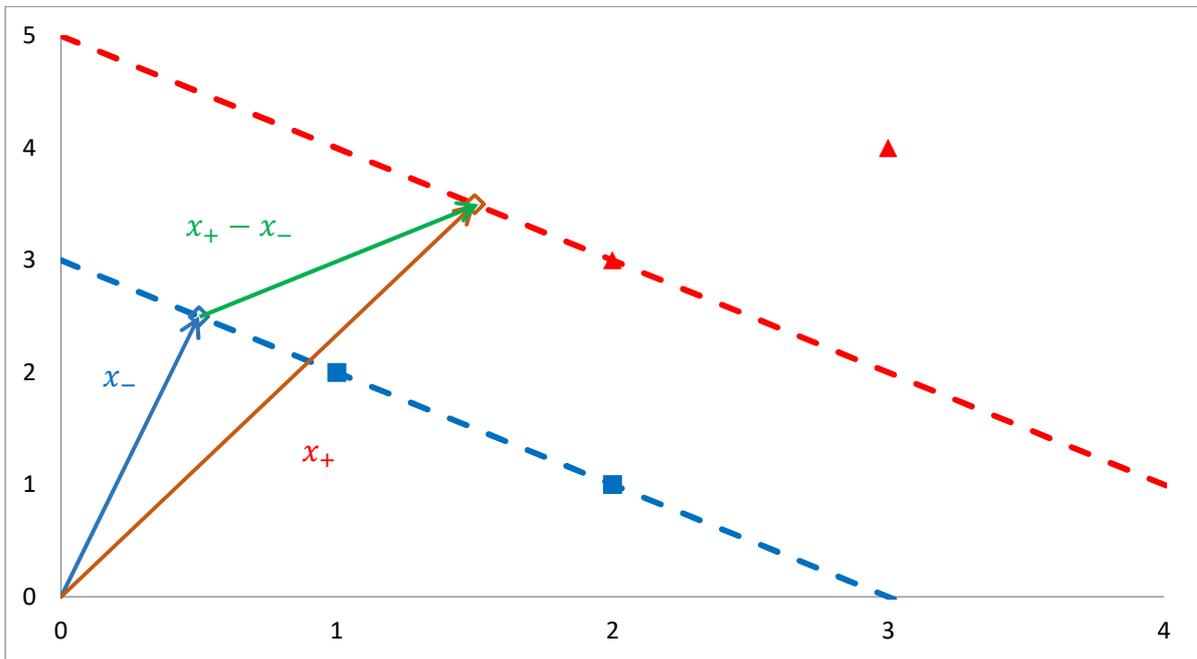
$$y_i(\vec{\omega} \cdot \vec{x}_i + b) - 1 = 0 \quad (10)$$

Embora que, a princípio, a Equação 10 se mostre inútil para a obtenção dos coeficientes  $\vec{\omega}$  e  $b$ , a mesma é utilizada posteriormente para a obtenção desses parâmetros.

Uma vez com a disposição das equações necessárias para a obtenção dos coeficientes  $\vec{\omega}$  e  $b$ , mais uma ilustração (Figura 22) deve ser apresentada. Para encontrar o hiperplano desejado, deve-se primeiramente, definir uma alternativa para calcular a distância entre os limites em azul e em vermelho representado na Figura 19. Isso é visto como necessário, pois o hiperplano tem como características o maior valor possível para essa distância.

Sendo assim, a Figura 22 introduz três vetores linearmente dependentes. O primeiro vetor, representado em vermelho, representa um vetor qualquer que esteja sobre o limite vermelho. Em outras palavras, é um vetor qualquer que indica uma amostra positiva no limite. O segundo vetor, representado em azul, representa outro vetor qualquer que esteja sobre o limite azul. Em outras palavras, é um vetor qualquer que indica uma amostra negativa no limite. O terceiro vetor, em cor verde, é um vetor obtido da subtração entre o vetor vermelho com o azul. É com esse último vetor que é possível obter o parâmetro desejado.

Figura 22 – Definição de vetores para determinar o hiperplano no SVM



Fonte: Adaptado de MIT OpenCourseWare (2020b).

Uma vez apresentado a Figura 22, é possível apresentar a Equação 11, que tem como função medir a distância desejada. Vale ressaltar que apenas a subtração entre os vetores indicados não é o suficiente para definir esse parâmetro, pois a componente que realmente importa desse vetor é aquela que é ortogonal ao vetor  $\vec{\omega}$ . Isso justifica a existência do produto escalar. Todavia, como  $\vec{\omega}$  pode não ser unitário, é necessário dividir o vetor pelo seu módulo. Observe que caso o mesmo fosse unitário, essa operação não seria necessária.

$$width = (\vec{x}_+ - \vec{x}_-) \cdot \frac{\vec{\omega}}{\|\omega\|} \quad (11)$$

A fim de facilitar a explicação a posteriori, a Equação 12 mostra uma forma alternativa de representar a Equação 11. Basicamente, é aplicada a propriedade da distributiva válida para produtos escalares.

$$width = (\vec{x}_+ \cdot \vec{\omega} - \vec{x}_- \cdot \vec{\omega}) \frac{1}{\|\omega\|} \quad (12)$$

Nesse momento, é possível utilizar a Equação 10 para obter uma forma alternativa para representar esse parâmetro. Vale ressaltar que uma das vantagens de se escrever dessa forma é remover a operação de produto escalar. Essa forma alternativa é representada na Equação 13.

$$width = [1 - b - (-1 - b)] \cdot \frac{1}{\|\omega\|} \quad (13)$$

Sendo assim, a Equação 14 mostra a forma definitiva para se calcular a distância desejada. Vale ressaltar que o valor de  $b$  não influencia na largura, o que faz sentido uma vez que esse parâmetro representa apenas um deslocamento do hiperplano com relação à origem.

$$width = \frac{2}{\|\omega\|} \quad (14)$$

Finalmente é obtida a expressão que é necessária para obter o hiperplano. Lembrando que o hiperplano é obtido a partir da maximização dessa margem, um algoritmo de otimização deve ser utilizado. Para facilitar o equacionamento, a Equação 15 indica algumas equivalências que existem na otimização. Algumas propriedades tais como constantes não afetam no  $\|\omega\|$  ótimo e que minimizar uma variável é equivalente a maximizar o inverso da mesma são utilizadas para aferir tais afirmações.

$$\max \frac{2}{\|\omega\|} \rightarrow \max \frac{1}{\|\omega\|} \rightarrow \min \|\omega\| \rightarrow \min \frac{1}{2} \|\omega\|^2 \quad (15)$$

O cuidado que se deve tomar nesse momento é que o problema de otimização a ser resolvido não é, a princípio, um problema irrestrito, ou seja, sem restrições. A restrição desse problema de otimização é dado pela Equação 10, a qual garante que todos os pontos presentes no conjunto de dados devem ser classificados corretamente, independentemente se for uma amostra positiva ou negativa.

Para transformar esse problema de otimização em um problema irrestrito, é possível utilizar o conceito de multiplicadores de Lagrange. Uma das grandes vantagens em transformar o problema em irrestrito é a possibilidade de resolver o mesmo através do método por derivadas parciais. Sendo assim, o problema irrestrito  $L$  (letra escolhida como abreviação de Lagrange no equacionamento) é apresentado na Equação 16.

$$L = \frac{1}{2} \|\omega\|^2 - \sum \alpha_i [y_i (\vec{\omega} \cdot \vec{x}_i + b) - 1] \quad (16)$$

Uma vez definida a Equação 16, a Equação 17 mostra a derivada parcial da função  $L$  com relação a variável  $\vec{\omega}$ . Vale ressaltar que, em condições de ponto ótimo, essa derivada é igual à zero. Sendo assim, a Equação 18 mostra a conclusão obtida a partir desse equacionamento. Vale ressaltar que uma conclusão válida que pode ser extraída a partir desses resultados é que  $\vec{\omega}$  é obtido a partir de uma soma ponderada das amostras do conjunto de dados.

$$\frac{\partial L}{\partial \vec{\omega}} = \vec{\omega} - \sum \alpha_i y_i \vec{x}_i = 0 \quad (17)$$

$$\vec{\omega} = \sum \alpha_i y_i \vec{x}_i \quad (18)$$

A Equação 19 mostra a derivada parcial da função  $L$  com relação a variável  $b$ . Vale ressaltar que, em condições de ponto ótimo, essa derivada também é igual à zero. Sendo assim, a Equação 20 mostra a conclusão obtida a partir desse equacionamento. Uma conclusão válida que pode ser extraída a partir desses resultados é que a somatória dos pesos que são utilizados para a obtenção de  $\vec{\omega}$  na Equação 18 é zero.

$$\frac{\partial L}{\partial b} = - \sum \alpha_i y_i = 0 \quad (19)$$

$$\sum \alpha_i y_i = 0 \quad (20)$$

Uma vez feito as derivadas parciais, é possível aplicar a Equação 18 e a Equação 20 na Equação 16 a fim de obter um valor para  $L$ . Uma vez que o equacionamento é trabalhoso, não são expostos nesse estudo os passos realizados para as simplificações existentes. O estudo não visa se aprofundar na matemática existente. Todavia, o resultado final desse equacionamento é apresentado na Equação 21. O somatório duplo existente mostra a grande dependência de  $L$  dos valores obtidos a partir dos produtos escalares dos pares de amostras.

$$L = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \quad (21)$$

Com o equacionamento apresentado, é possível obter os valores de  $\vec{\omega}$  e  $b$  para modelar o hiperplano do SVM. Todavia, o problema de otimização proposto, obtido a partir da minimização da Equação 21 é complexo demais para ser resolvido

manualmente. Nesse caso, é necessária a utilização de uma máquina computacional.

Ainda é relevante mostrar a regra de decisão para a classificação de novas amostras. A Equação 22 mostra essa regra de decisão para classificar uma amostra como positiva. A dedução para a expressão matemática apresentada não é considerada como relevante para o estudo devido a sua complexidade.

$$\sum \alpha_i y_i \vec{x}_i \cdot \vec{u} + b \geq 0 \quad (22)$$

Com isso, termina-se a exposição matemática envolvida nesse algoritmo de ML. Esses equacionamentos ajudam a entender as mecânicas envolvidas dentro de um algoritmo computacional. Todavia, existem outras variações que envolvem um equacionamento mais complexo.

Esse método apresenta uma boa precisão na generalização dos dados mesmo em caso em que os padrões existentes entre os dados são difíceis de identificar (SAMMUT; WEBB, 2011). Também, o método apresenta um desempenho satisfatório mesmo em espaços de alta dimensão. O algoritmo ainda é eficaz nos casos em que o número de dimensões é maior do que o número de amostras (COURNAPEAU, 2020).

Neste trabalho, optou-se por utilizar o método SVM do módulo *Scikit learn* (COURNAPEAU, 2020). Esta biblioteca é gratuita, de livre distribuição e permite derivativos.

### 2.3.3 One-versus-One (OvO)

O SVM é desenvolvido fundamentalmente para classificar apenas duas classes, todavia, existem variações desse algoritmo que é válido para problemas multiclases – Máquina de Vetores de Suporte Multiclasse (MSVM) (BOHRA; PALIVELA, 2015). Dentre essas variações, existe o aprendizado *One-versus-Rest* (OvR) e o *One-versus-One* (OvO). Essas variações não mudam os equacionamentos apresentados anteriormente, apenas à forma de selecionar as amostras para o treinamento da máquina.

O OvR envolve o treinamento de um único classificador por classe. Nesse contexto, as amostras dessa classe são consideradas como amostras positivas e todas as outras amostras são consideradas como negativas (SHWARTZ; DAVID, 2014).

Um dos problemas com a abordagem OvR é que os conjuntos de treinamento são desequilibrados (SHWARTZ; DAVID, 2014). Por exemplo, em problema de classificação com dez classes, cada uma com um número igual de dados de treinamento, os classificadores individuais serão treinados em conjuntos de dados que incluem 90% de amostras negativas e apenas 10% de amostras positivas, e a simetria do problema original é perdida.

O OvO é um algoritmo que classifica através de agrupamento de pares de classe, fazendo a combinação de todas as classes possíveis. No momento da predição, a classe que receber mais votos é escolhida para classificar a amostra. Devido a esse agrupamento, o OvO é computacionalmente mais custoso com relação ao OvR (GALAR *et al.*, 2014), todavia, esse algoritmo consegue contornar certos problemas existentes no OvR, conforme apontado anteriormente.

O estudo aqui apresentado possui dez classes quando é levado em consideração o tipo e a severidade do desbalanceamento. Sendo assim, enquanto o método OvR executaria a matemática apresentada no SVM 10 vezes, o OvO executaria 45 vezes. O primeiro valor é definido diretamente pelo número de classes existentes no problema. O segundo valor é obtido na comparação par a par de cada uma das dez classes existentes.

A fim de simplificar a nomenclatura, durante a apresentação da metodologia e dos resultados obtidos no estudo, toda vez que for mencionado o algoritmo SVM, este se tratará de sua versão multiclasse (MSVM) através da variação OvR. Todavia, toda vez que for mencionado o algoritmo OvO, deve-se ficar subtendido que se trata da variação do MSVM.

Neste trabalho, optou-se por utilizar o método OvO do módulo *Scikit learn* (COURNAPEAU, 2020). Esta biblioteca é gratuita, de livre distribuição e permite derivativos.

### 3 DESCRIÇÃO DA METODOLOGIA

A metodologia a ser utilizada foi dividida em partes para que a mesma seja mais fácil de ser entendida, executada e interpretada. São elas: aquisição dos sinais, processamento dos dados e a sua interpretação para obter informação útil.

#### 3.1 Aquisição dos sinais

Foi construído um conjunto de dados com a condição saudável da máquina e quando a mesma é sujeita aos desbalanceamentos estático, acoplado e dinâmico. Para poder fazer um teste de versatilidade entre os algoritmos de ML, o conjunto de dados foi construído para três velocidades distintas e três níveis de severidade para os desbalanceamentos.

As três classes de velocidade do rotor foram obtidas com base na variação da frequência de alimentação do motor. O inversor de frequência permite fazer tal variação. Foram utilizadas as velocidades para frequência de 5 Hz, 7,5 Hz e 10 Hz, o que resulta em velocidades de 150 RPM, 225 RPM e 300 RPM, respectivamente, do rotor.

Estes valores são obtidos a partir da velocidade síncrona ( $N_s$ ) de um motor de indução, cuja Equação 23 mostra a dependência desse parâmetro com a frequência de alimentação do motor ( $f$ ) e com o número de polos ( $p$ ) do mesmo. O motor de indução empregado possui quatro polos, o que justifica os valores obtidos anteriormente.

$$N_s = \frac{120 f}{p} \quad (23)$$

Os três níveis de severidade de desbalanceamento foram definidos com base na variação da massa que é inserido no disco de inércia. Foram definidos os níveis 1, 2 e 3, na qual quanto maior o índice, maior é o valor da massa que causa o desbalanceamento. Estes níveis podem ser observados na Figura 23. As massas foram inseridas no disco de inércia através de parafusos. A Figura 24 retrata justamente essa inserção ao disco de inércia para cada um dos níveis de

desbalanceamento. Vale ressaltar ainda que foi certificado de que os parafusos estivessem bem apertados ao furo para garantir que não houvesse vibração. Essa vibração poderia prejudicar a qualidade do conjunto de dados.

A Tabela 5 mostra o resultado de um ensaio realizado para medir a massa de desbalanceamento de cada um dos níveis propostos. Vale ressaltar que conhecer a massa é importante uma vez que elas são utilizadas para emular o comportamento de desbalanceamento sobre o rotor. O ensaio consistiu em medir a massa de cada parafuso, rosca e arruela que são utilizados nessa parte com uma balança de precisão. Foram selecionadas três peças de cada tipo para poder determinar um peso médio para cada elemento. Para fazer o desbalanceamento proposital no eixo do motor, os parafusos apontados na Figura 23 são inseridos no disco de inércia conforme mostra a Figura 24.

Honorato (2017) já havia feito um ensaio parecido para determinar o valor das massas de desbalanceamento, todavia, foi analisado que seria interessante repetir a análise para garantir que os valores obtidos estavam condizentes. A Tabela 6 mostra a comparação entre os valores obtidos.

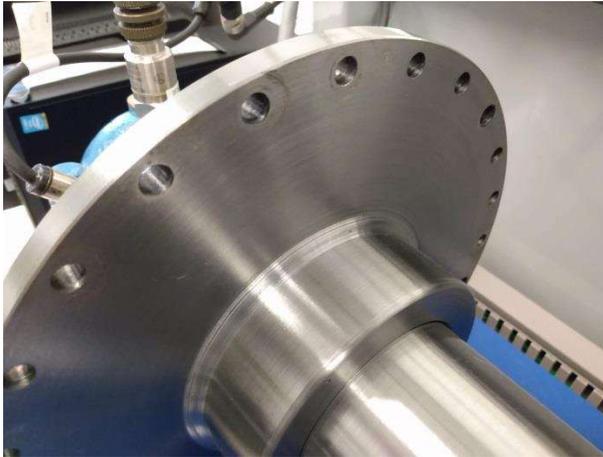
Embora não seja o foco do estudo proposto, esse ensaio simples e rápido de se realizar pode ajudar atividades futuras do PECCE que podem necessitar dessas informações. Pode ser que, futuramente, haja a necessidade de se fazer um desbalanceamento de nível 4. Nesse contexto, a Tabela 6 pode ajudar a gerar uma massa de desbalanceamento ideal para a aplicação do estudo.

**Figura 23 – Parafusos utilizados para o desbalanceamento**

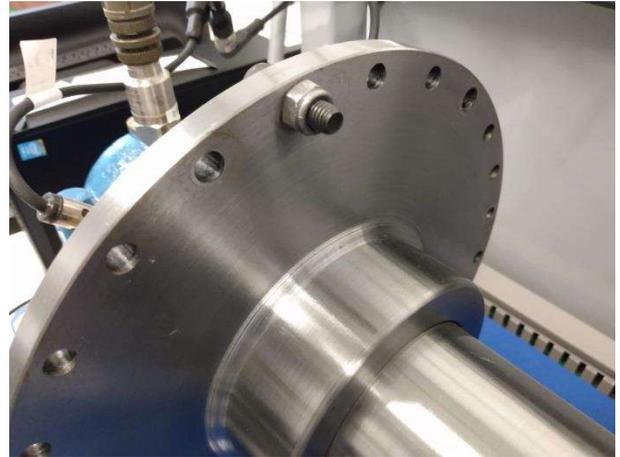


Fonte: Elaboração Própria (2020).

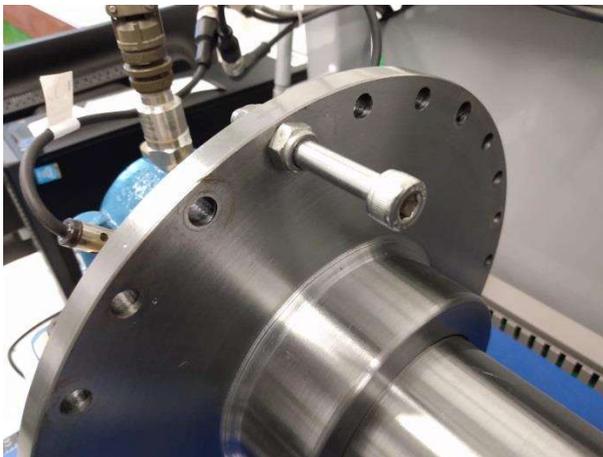
**Figura 24 – Parafuso de desbalanceamento acoplado no disco de inércia**



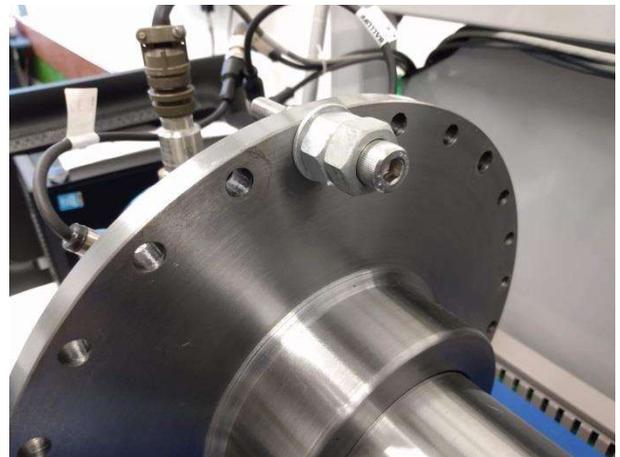
a) Sem desbalanceamento



b) Desbalanceamento nível 1



c) Desbalanceamento nível 2



d) Desbalanceamento nível 3

Fonte: Elaboração Própria (2020).

**Tabela 5 – Massa de parafusos, roscas e arruelas, em gramas**

Tipo	Parafuso		Rosca		Arruela	
	Grande	Pequeno	Grande	Pequena	Grande	Pequena
Medida 1	27,4	14,1	14,7	4,5	15,3	4,9
Medida 2	27,3	13,9	14,7	4,5	14,8	4,8
Medida 3	27,3	13,9	14,7	4,5	15,0	4,9
Média	27,3	14,0	14,7	4,5	15,0	4,9

Fonte: Elaboração Própria (2020).

**Tabela 6 – Massa dos níveis, em gramas, e comparação**

	Valor	Referência	Erro	Descrição
Nível 1	18,5	18,1	2,03%	(1) Parafuso pequeno + (1) rosca pequena
Nível 2	36,3	35,7	1,77%	(1) Parafuso grande + (2) rosca pequena
Nível 3	66,1	65,3	1,23%	(1) Parafuso grande + (1) rosca pequena + (2) rosca grande + (1) arruela pequena

Fonte: Elaboração Própria (2020).

Para a construção do conjunto de dados foi necessária à utilização do *software* Labview. Foi utilizada uma frequência de amostragem de 50.000 amostras por segundo e foram registrados no total de 200.000 pontos por amostra. Isto resulta em intervalo de observação de 4 segundos. Estes valores foram escolhidos para que os resultados obtidos possam ser comparados a outros trabalhos que já foram realizados na mesma bancada como o trabalho defendido por Honorato (2017).

Foi levada em consideração a preocupação de efetuar todos os ensaios no menor período de tempo possível para evitar interferências externas que poderiam contaminar o conjunto de dados. Também foi levada em consideração a preocupação de garantir que o sistema entre em regime antes de aferir as grandezas para evitar alterações no conjunto de dados.

### 3.2 Processamento de dados e sua interpretação

Após a criação do conjunto de dados, foi pensado em como as informações contidas deveriam ser modeladas para que se pudesse usar os algoritmos de ML apresentados. Para isso, foi utilizado o *software Matlab* (MATHWORKS, 2020). A decisão para a utilização deste *software* foi dada pela familiaridade e facilidade para a manipulação das informações que o mesmo fornece.

A utilização de muitos processos para formatar o conjunto de dados podem inutilizar informações ou características do conjunto de dados (BATISTA, 2003). Isso não é interessante para um algoritmo de ML, pois o mesmo teria maiores dificuldades para inferir um diagnóstico com exatidão.

Todavia, um sistema preditivo é considerado eficaz quando o mesmo consegue captar apenas as características mais relevantes do conjunto de dados para inferir um resultado (MOHAMMED; KHAN; BASHIER, 2017). Qualquer característica que não tem relevância pode viciar um sistema preditivo a gerar padrões incongruentes. Esses padrões podem diminuir a eficácia do algoritmo. Sendo assim, utilizar o conjunto de dados exatamente da forma como foi adquirido da bancada não é uma boa opção. O sinal no domínio do tempo possui 200.000 pontos por amostra. Utilizar todos estes pontos é inviável para um algoritmo de aprendizado uma vez que isso iria requerer muito processamento computacional.

Também, para cada amostra, existe uma defasagem angular específica. Essa defasagem iria dificultar muito o algoritmo de aprendizado, pois o mesmo iria tentar buscar algum padrão através da existência dessa defasagem para gerar o diagnóstico. A questão é que não existe nenhuma relação direta com o tipo de desbalanceamento com essa característica. A defasagem apenas teria utilidade na localização da massa desbalanceadora, em outras palavras, esse parâmetro seria útil apenas para determinar a posição exata da massa de balanceamento.

Para solucionar esta dificuldade, foi utilizada a transformada rápida de Fourier (FFT) (DEMAY, 2008), onde o sinal no domínio do tempo é transformado para o domínio da frequência. Isso resolve o problema relacionado com a defasagem angular do sinal. Todavia, ainda existem muitas componentes para representar a amostra. Sendo mais específico, existem 25.000 componentes por amostra, conforme o teorema de Nyquist (NATIONAL INSTRUMENTS, 2020a).

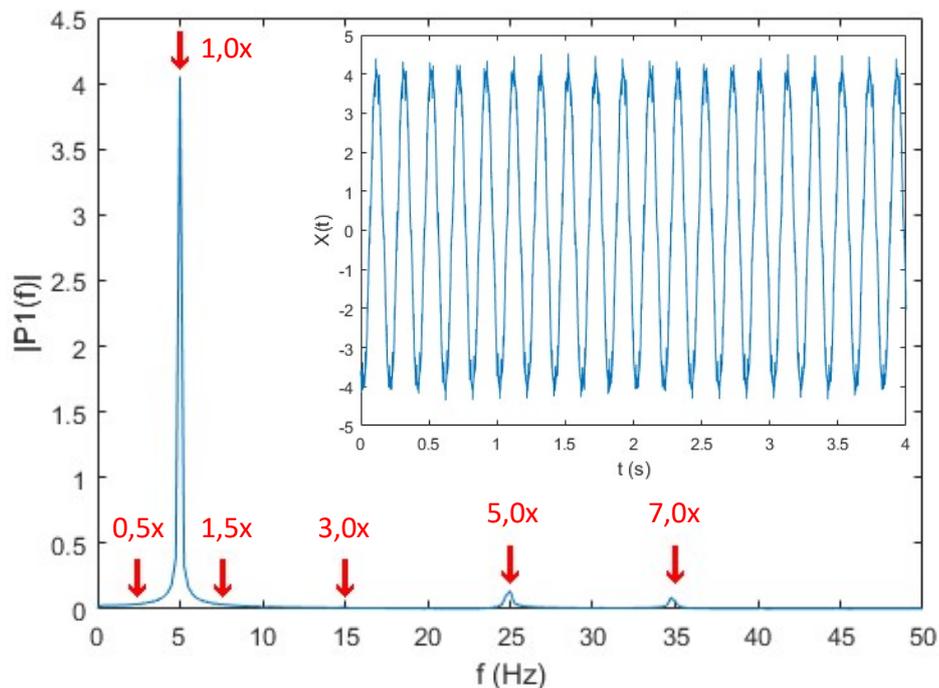
Para facilitar o treinamento do algoritmo, foi utilizado um critério visual que consiste selecionar as componentes que possuem variações na amplitude de acordo com o tipo de desbalanceamento. Para uma frequência fundamental  $f_1$ , que pode ser 5 Hz, 7,5 Hz e 10 Hz, foi selecionado as frequências  $0,5*f_1$ ,  $1,0*f_1$ ,  $1,5*f_1$ ,  $3,0*f_1$ ,  $5,0*f_1$  e  $7,0*f_1$ . As demais componentes foram dispensadas usando dois critérios: ou possuem uma amplitude muito pequena, o que torna mais susceptível a ruídos e, portanto, menor a confiabilidade; ou porque não trazia informações novas para os múltiplos de frequência selecionados.

A Figura 25 mostra em detalhes o processo efetuado para a seleção dos múltiplos de frequência mencionados. É possível observar um sinal de corrente do

motor operando com uma alimentação de 5 Hz no domínio do tempo dentro do gráfico a qual se faz sua transformada de Fourier. As setas em vermelho dentro do gráfico do domínio da frequência indicam as componentes que foram selecionadas para participar do treinamento dos algoritmos de ML.

Uma vez feitas simplificações no conjunto de dados inicial para a utilização futura de ferramentas de ML, existe a necessidade de se estudar métodos para a visualização das informações contidas. Poder visualizar o conjunto de dados é importante para poder detectar se o mesmo está próprio para a utilização de algoritmos de ML.

**Figura 25 – Seleção de componentes no domínio da frequência**



Fonte: Elaboração Própria (2020).

O problema se encontra quando existe a impossibilidade da visualização do conjunto de dados completo devido a diversas variáveis existentes. É possível reduzir a dimensão quando é fixado algum dos parâmetros do conjunto de dados, todavia, esse processo pode impossibilitar certas conclusões caso seja feito de maneira equivocada. Para uma boa representação gráfica, que seria aquela que permite tirar conclusões de forma fácil e evidente, deve-se saber qual parâmetro simplificar e qual ressaltar.

Para fazer as análises de ML, usou-se o *Colaboratory*, uma plataforma do Google (GOOGLE, 2020). Esta plataforma permite trabalhar com a linguagem Python e, dentre dos módulos que existem, possui um que é voltado especificamente para o ML. Este módulo se chama *Scikit learn* (COURNAPEAU, 2020). Outros módulos também são utilizados para o tratamento dos dados, tais como o *Numpy* (OLIPHANT, 2020a), para a realização de operações matriciais e o *Scipy* (OLIPHANT, 2020b), para importar e exportar o conjunto de dados.

Para cada um dos algoritmos de ML, foram feitos os seguintes ensaios: predição para determinar tipo e severidade do desbalanceamento; predição para determinar apenas tipo de desbalanceamento; e predição para determinar apenas severidade de desbalanceamento. A ideia é verificar o quanto que o desempenho dos algoritmos é alterado nestas configurações.

## 4 RESULTADOS OBTIDOS

### 4.1 Sem auxílio de ML

Preliminarmente à aplicação dos algoritmos de ML, foram realizadas análises através de gráficos para buscar evidências de relação entre as variáveis. Esses gráficos determinam o quanto que as classes estão bem definidas e a discrepância dos dados. Também, esta análise permite observar as características relevantes que são necessárias para fazer o treinamento dos algoritmos de ML. Dentre das observações obtidas sem o auxílio de ferramentas de ML estão:

- a) É possível obter a mesma conclusão independente da fase a qual é feita a análise, mesmo com a pequena variação existente nos valores entre as fases;
- b) Um desbalanceamento qualquer, independente do tipo ou da severidade, possui a capacidade de mudar a amplitude da corrente elétrica do motor;
- c) Quanto maior for a severidade do desbalanceamento maior tende a ser a variação da amplitude da corrente com relação a um valor referencial de corrente saudável da máquina;
- d) Quando se analisa a corrente de uma máquina na forma de distribuição normal, a fim de avaliar a influência de desbalanceamentos sobre a amplitude da corrente, o desvio padrão das amostras que pertencem a qualquer tipo de desbalanceamento é maior do que o desvio padrão na condição saudável da máquina. Isso evidencia a incerteza que o valor da amplitude da corrente pode assumir quando o motor está sujeito a qualquer tipo de desbalanceamento. Vale ressaltar que isso é justificado quando se percebe que existe uma vasta gama de severidades possíveis de desbalanceamento;
- e) Quando o sinal de corrente é decomposto através de uma FFT, quando maior for a amplitude de uma componente, mais relevante é a mesma para poder classificar os tipos de desbalanceamento. Isso

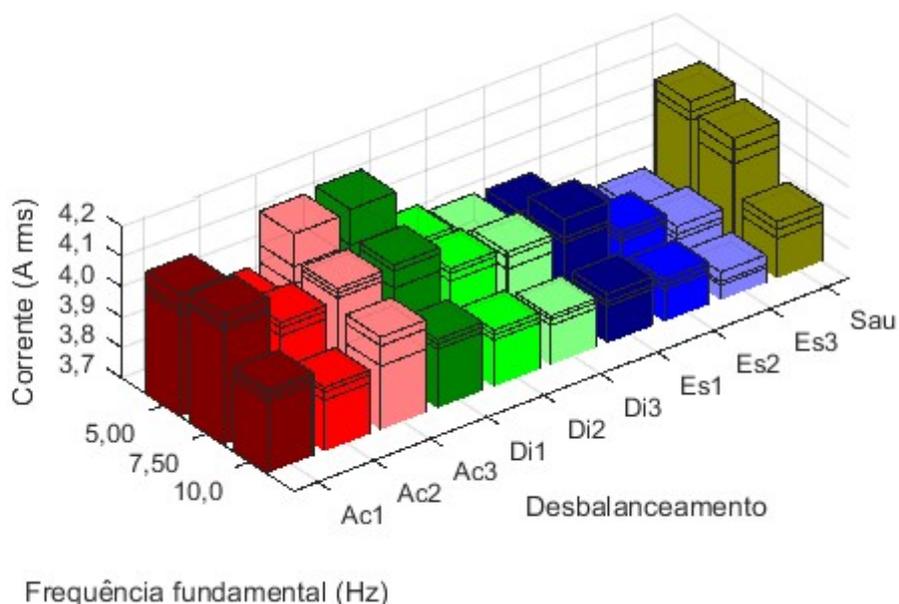
permite afirmar que a componente fundamental é a componente mais importante para a classificação de desbalanceamentos em máquinas rotativas;

- f) Componente com baixa amplitude está mais sujeito a ruídos, o que pode atrapalhar na classificação correta de desbalanceamento;
- g) Assim como o desbalanceamento dinâmico pode ser decomposto, em termos de massa de desbalanceamento, na somatória de um desbalanceamento estático com um desbalanceamento acoplado, o efeito da variação de corrente do desbalanceamento dinâmico também pode ser dividido de forma similar;
- h) Desbalanceamento estático tende a reduzir a amplitude da componente fundamental da corrente, assim como a amplitude das componentes 3 vezes a fundamental, 5 vezes a fundamental e 7 vezes a fundamental, e essa redução é maior quanto maior for a severidade;
- i) Em uma máquina elétrica rotativa com desbalanceamento acoplado, quanto menor for a frequência fundamental de alimentação, em outras palavras, quanto menor a rotação do eixo do motor, menor tende a ser a amplitude da componente fundamental, o que torna mais fácil de detectar visualmente o desbalanceamento através de plotagem de gráficos específicos.

Para exemplificar o citado acima, apresentam-se algumas figuras.

A Figura 26 mostra um gráfico de barras tridimensional a qual se foca a mostrar a componente fundamental da fase T para cada tipo de desbalanceamento. Para cada tipo de desbalanceamento, existe uma cor associada. Ressalta-se que para cada uma das 30 barras tridimensionais existem três marcações: a primeira marcação representa o menor valor obtido de corrente entre as repetições da aquisição de dados; a segunda marcação representa a mediana das repetições; a terceira marcação representa o maior valor obtido.

Figura 26 – Visão geral da componente fundamental na fase T



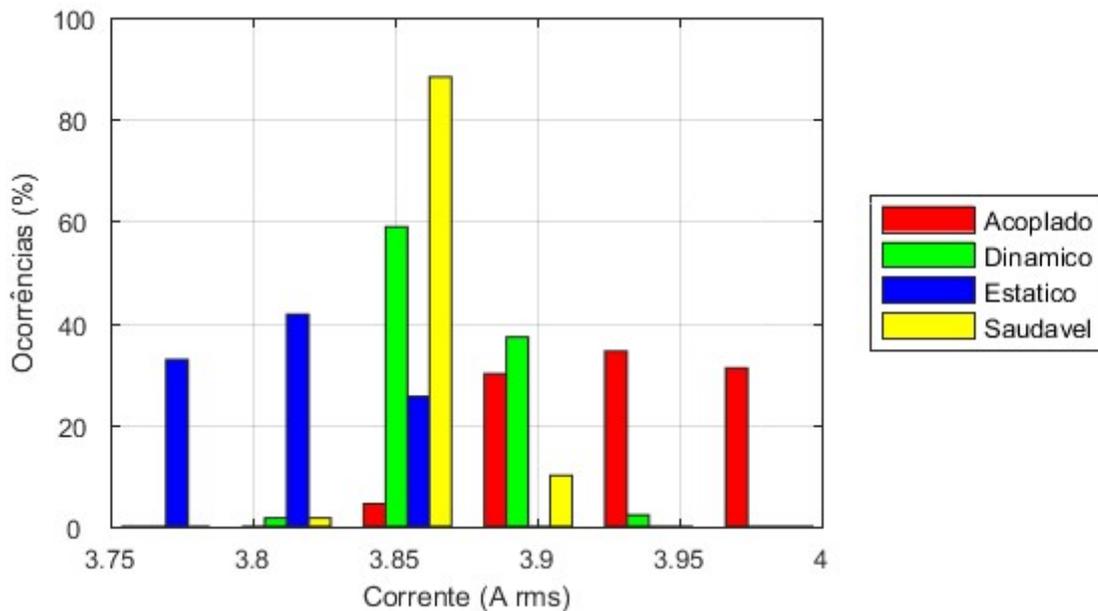
Fonte: Elaboração Própria (2020).

A Figura 27 mostra um histograma de ocorrência normalizada para a componente fundamental de 10 Hz na fase R. Na abscissa é definida a corrente, em ampères eficazes, e na ordenadas é definida a ocorrência, em percentual. O histograma possui seis conjuntos de barras, sendo que cada conjunto é composto por quatro barras que possuem cores diferentes. Cada cor está associada com um tipo de desbalanceamento, conforme mostra a legenda. A severidade do desbalanceamento não é considerada em primeiro momento para poder facilitar a visualização dos elementos já existentes.

A Figura 28 mostra um histograma que também visa mostrar a severidade do desbalanceamento para a componente 0,5x fundamental de 5 Hz na fase T. Na abscissa é definida a corrente, em ampères eficazes, todavia, na ordenadas é definida a ocorrência, por unidade. O valor máximo de ocorrências para cada classe existente no gráfico é de 60, ou seja, esse valor máximo está relacionado com as repetições das amostras coletadas. O histograma possui seis conjuntos de barras, sendo que cada conjunto é composto por dez barras que possuem cores diferentes. Conforme já havia comentado anteriormente, cada cor está associada com um tipo de desbalanceamento e a sua severidade, conforme mostra a legenda. Esse gráfico

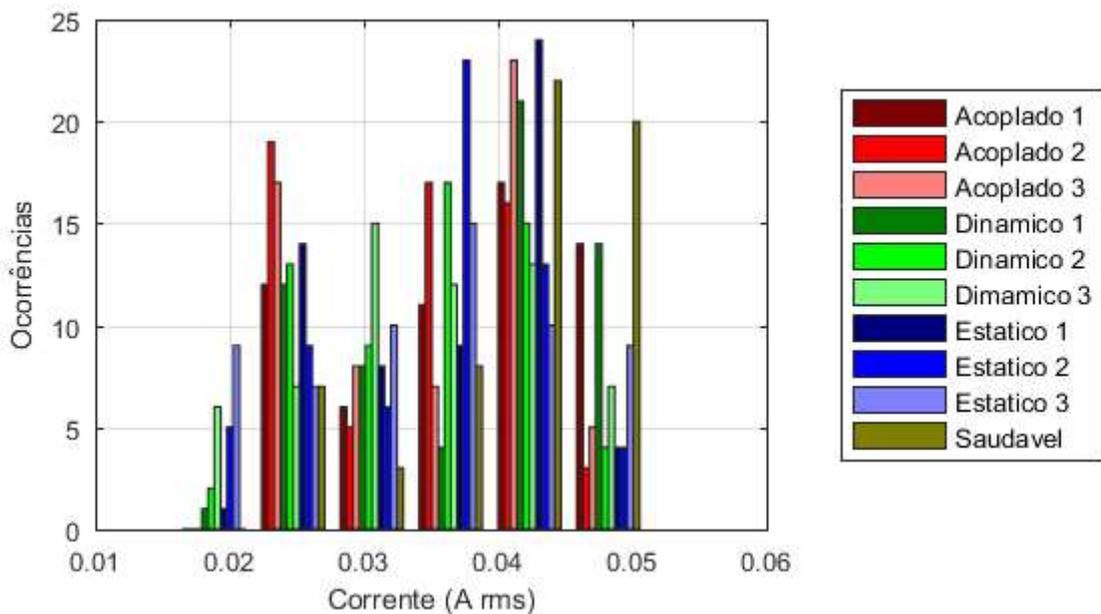
permite evidenciar melhor o efeito da severidade de um desbalanceamento sobre a amplitude da componente.

**Figura 27 – Histograma simplificado da componente fundamental**



Fonte: Elaboração Própria (2020).

**Figura 28 – Histograma detalhado da componente 0,5x fundamental**

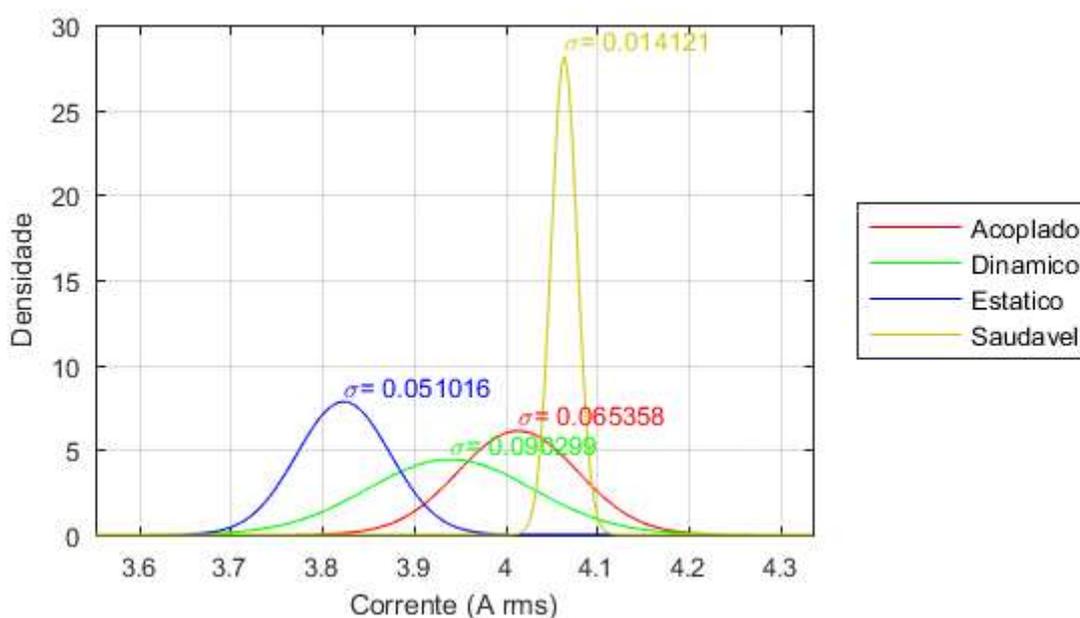


Fonte: Elaboração Própria (2020).

A Figura 29 mostra um gráfico de distribuição normal para a componente fundamental de 5 Hz na fase R. Para se plotar uma curva de distribuição normal, é

necessária conhecer o valor médio e a variância da série de amostras. A média é um parâmetro que representa o valor a qual estão concentradas as amostras de uma determinada série. A variância é um parâmetro de dispersão que determina o quanto as amostras estão distante do valor médio. Vale ressaltar que esse gráfico permite visualizar com mais clareza os valores médios de cada classe existente. A variância de cada curva está explícita no gráfico para facilitar a visualização e a interpretação dos dados.

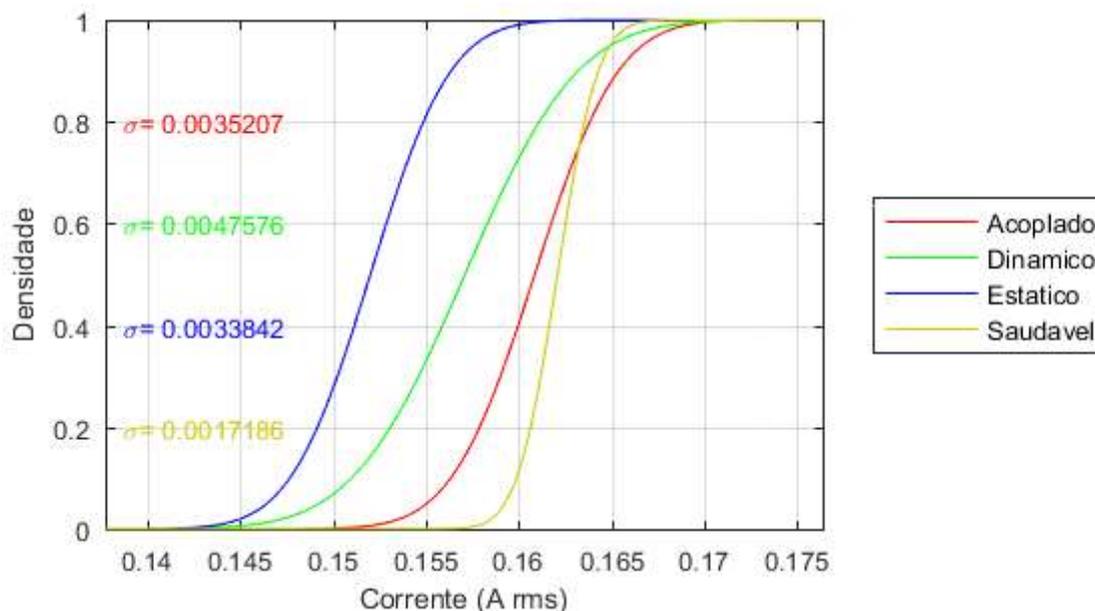
**Figura 29 – Distribuição de probabilidade da componente fundamental**



Fonte: Elaboração Própria (2020).

A Figura 30 mostra um gráfico de distribuição cumulativa para a componente 5,0x fundamental de 7,5 Hz na fase S. Aqui são utilizadas curvas de distribuição cumulativa. Na abscissa é definida a corrente, em ampères eficazes, todavia, na ordenadas é definida a densidade, cujo parâmetro pode variar dentro do intervalo delimitado pelos valores 0 e 1 e que representa o quanto de amostras pertencem ao intervalo definido entre a origem até o valor que está sendo definido. Cada curva presente no gráfico representa um tipo de desbalanceamento. Quanto maior for a inclinação da curva em um determinado ponto maior é a densidade de amostras. Quanto maior for a distância entre as curvas, mais fácil é a distinção entre as classes existentes. A interseção entre as curvas mostra o quanto que as classes podem ser confundidas entre si.

Figura 30 – Distribuição cumulativa da componente 5,0x fundamental

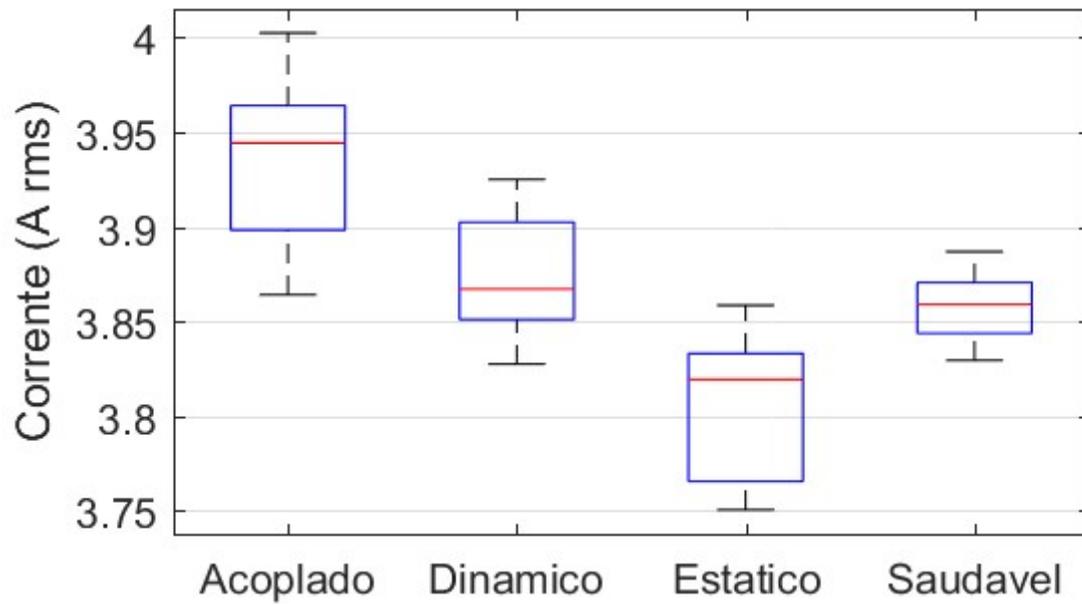


Fonte: Elaboração Própria (2020).

A Figura 31 mostra um gráfico de *boxplot* que não inclui a severidade do desbalanceamento para a componente fundamental de 10 Hz na fase S. Esse gráfico trabalha com conceito de quantil, mais especificamente, de quartis. Na abscissa é definido o tipo de desbalanceamento e na ordenadas é definida a corrente, em ampères eficazes. O quartil permite visualizar como as amostras de uma série estão distribuídas. Para isso, o quartil organiza todas as amostras em ordem crescente e, posteriormente, divide a série em quatro conjuntos com o mesmo número de amostras. Sendo assim, é possível afirmar que o segundo quartil define também a mediana das amostras e o mesmo está destacada em cor vermelha no gráfico.

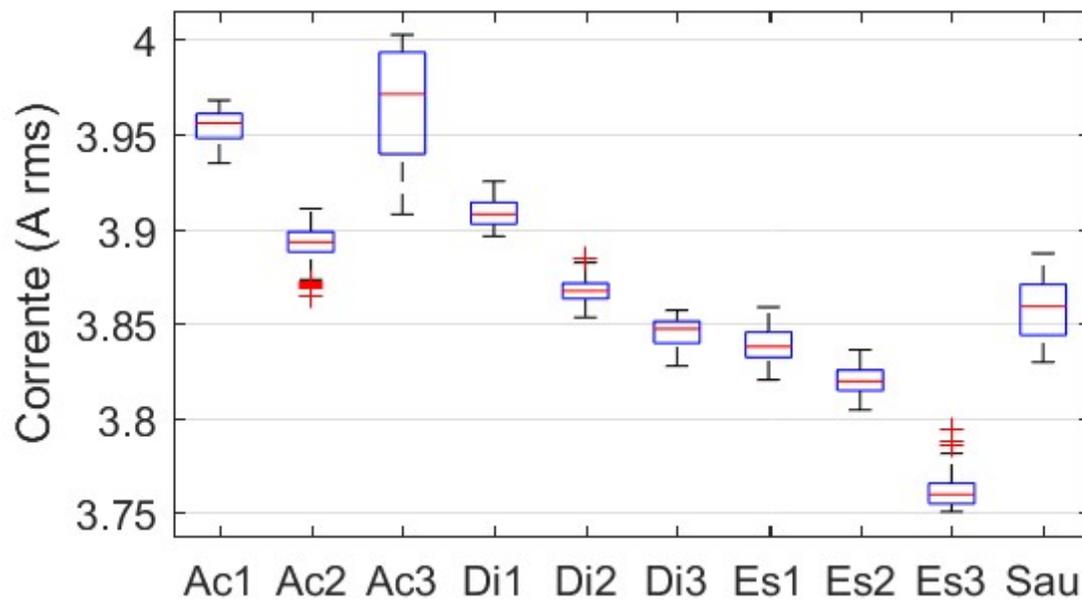
A Figura 32 mostra um gráfico de *boxplot* com severidade do desbalanceamento para a componente fundamental de 10 Hz na fase S. Os pontos representados em vermelho são amostras discrepantes do conjunto de dados. Os mesmos não são utilizados para determinar os quantis necessários para a plotagem do *boxplot*.

Figura 31 – *Boxplot* simplificado da componente fundamental



Fonte: Elaboração Própria (2020).

Figura 32 – *Boxplot* detalhado da componente fundamental



Fonte: Elaboração Própria (2020).

## 4.2 Com auxílio de ML

A utilização de algoritmos de ML permite obter resultados que não poderiam ser obtidos simplesmente com a observação de gráficos de distribuição normal, conforme foi apresentado na Figura 29, por exemplo. Com o uso de ML, têm-se as seguintes conclusões:

- a) Dentre todos os algoritmos de ML testados, o DTC apresentou o melhor desempenho em termos de precisão e tempo de execução, todavia, existem outros critérios que não foram considerados e que possuem grande relevância ao avaliar o desempenho de um classificador, tais como a memória alocada e a robustez;
- b) Dentre os algoritmos de DTC testados, aquele que é baseado no critério de entropia possui maior tempo de execução com relação ao que opera com o critério de gini. A diferença entre estes critérios presentes no módulo do *Scikit Learn* está na forma que a equação da desordem, apresentada na Equação 1, é definida. Ademais, as precisões de ambos classificadores demonstram ser similares na maior parte das análises efetuadas. Outra observação que é relevante de ser ressaltado porque pode afetar o tempo de execução do algoritmo é que não foi definida uma profundidade máxima para nenhuma das árvores modeladas no estudo;
- c) Dentre os possíveis *kernel* que podem ser utilizados para modelar o hiperplano e que estão disponíveis no módulo *Scikit Learn*, apenas o polinomial apresentou um desempenho satisfatório em uma análise inicial. Define-se como desempenho satisfatório toda execução de algoritmo que possui uma precisão significativa em função do tempo de execução. Nessa definição, algoritmos com baixa precisão ou que levam muito tempo para conseguir gerar um resultado são considerados como insatisfatórios em termos de desempenho;
- d) Variar a modelagem do SVM em OvR e OvO não trouxe nenhuma variação significativa em termos de precisão ou tempo de execução dos algoritmos. A princípio, não possui nenhuma explicação clara do porquê disto, todavia, é possível levantar certas hipóteses. Uma hipótese bem plausível é que a diferenciação destes algoritmos só

seria mais evidente se fosse utilizar um conjunto de dados maior e utilizando um *kernel* polinomial de maior grau. Ao fazer isto, o tempo de treinamento fica mais demorado, podendo evidenciar melhor o tempo de execução dos algoritmos, e também permite maior flexibilidade para que os algoritmos possam se comportar de forma diferente, podendo refletir na assertividade;

- e) Com relação ao grau do polinômio a ser utilizado no algoritmo de SVM, é recomendado valores menores para máquinas rotativas que operam em baixa rotação e valores maiores para altas rotações. Essa conclusão é feita pelo fato de que o desempenho dos algoritmos com polinômio de 20º grau tende a serem melhores, em termos de precisão e tempo de execução, do que os algoritmos com polinômio de 10º grau quando o motor era alimentado com uma frequência de 10 Hz, todavia a situação se inverte quando o motor é alimentado com uma frequência de 5 Hz. Vale ressaltar ainda que existe uma chance considerável de que uma alteração não muito significativa do grau do polinômio não afete no desempenho do classificador. Um exemplo de alteração significativa seria impor um novo valor de polinômio que seja 10 vezes maior do que o valor de referência;
- f) As análises apontam que todos os seis múltiplos de fundamental de corrente utilizados possui certa contribuição para a obtenção de melhores resultados nos classificadores testados. Quanto mais múltiplos são removidos para o aprendizado dos algoritmos, mais demorado fica o processo e pior tende a serem os resultados obtidos pelo classificador;
- g) Analisar o tipo e a severidade do desbalanceamento separadamente, ou seja, utilizar dois classificadores distintos, a fim de compor em conjunto um diagnóstico completo, só é compensatório quando o classificador não possui muitos graus de liberdade para poder diferenciar cada classe existente, o que ocorre, por exemplo, em classificador por SVM polinomial de baixíssimo grau, tais como 2º ou 3º grau. Boa parte dos resultados obtidos mostra que, se é desejada obter um diagnóstico completo, com tipo e severidade de

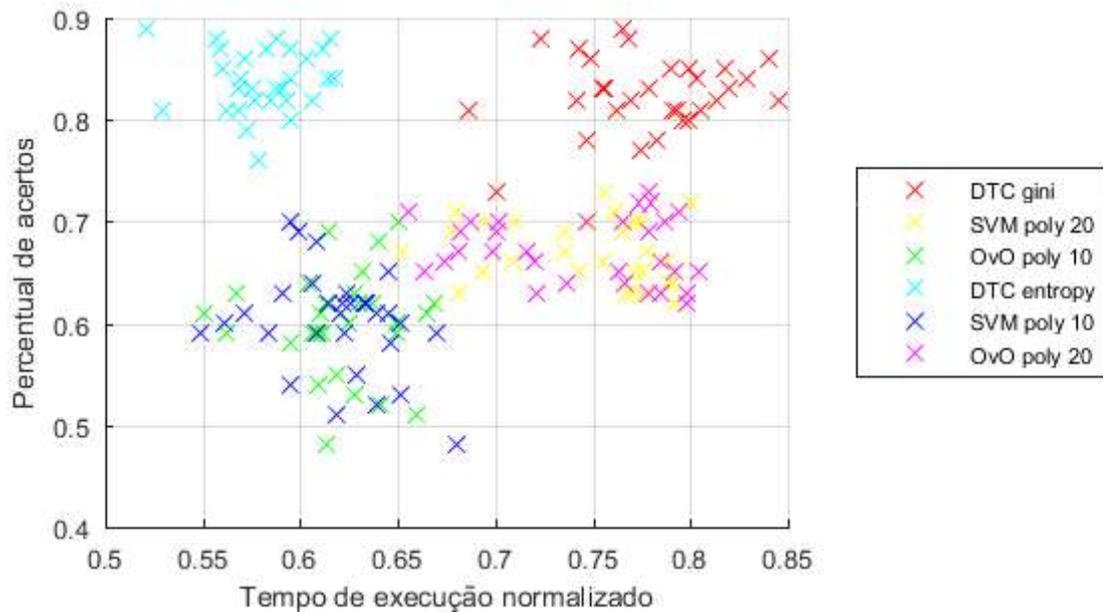
desbalanceamento, é melhor utilizar um único classificador que consiga distinguir adequadamente cada classe com os graus de liberdade;

- h) Os estudos apontaram que existem quatro tipos de classificação de erros que podem ser identificados em uma matriz de confusão de detecção de desbalanceamento de máquinas elétricas rotativas com base na leitura da corrente: erro aleatório, erro de severidade, erro de composição e erro de detecção. Vale ressaltar que cada tipo está associado com uma justificativa diferente para justificar a existência dos falsos positivos e negativos existentes na matriz de confusão. A justificativa de cada classe é mostrada posteriormente.

A Figura 33 mostra um exemplo de diagrama de pontos construído para avaliar o desempenho dos algoritmos. Nas ordenadas é definida a precisão, que é obtida a partir do percentual de acertos do classificador. O valor unitário para esse eixo representa que o algoritmo consegue acertar todas as predições e o valor nulo indica que o algoritmo não consegue fazer qualquer predição correta. A abscissa representa um indicativo de tempo, obtido através do inverso do tempo de execução. O valor unitário para esse eixo representa a amostra que obteve o menor tempo de execução das séries plotadas e o valor nulo representa um algoritmo que nunca consegue entregar uma predição, ou seja, tempo de execução infinito. Cada algoritmo possui uma cor diferente no diagrama e a legenda localizada a direita ajuda identificar qual cor está relacionada com qual algoritmo.

Uma vez que o diagrama de pontos pode dificultar a visualização de uma região a qual os resultados se concentram, foi buscada uma forma alternativa para a visualização do desempenho dos algoritmos. A Figura 34 mostra a plotagem de elipse do mesmo diagrama apresentada anteriormente, na Figura 33. O significado dos eixos ainda é o mesmo. Cada elipse indica uma zona a qual tende a se concentrar boa parte das amostras de um determinado algoritmo. O centro da elipse também é esboçado e visa mostrar o ponto mais provável que representa o desempenho de cada classificador.

Figura 33 – Desempenho na detecção de tipo e severidade (ponto)

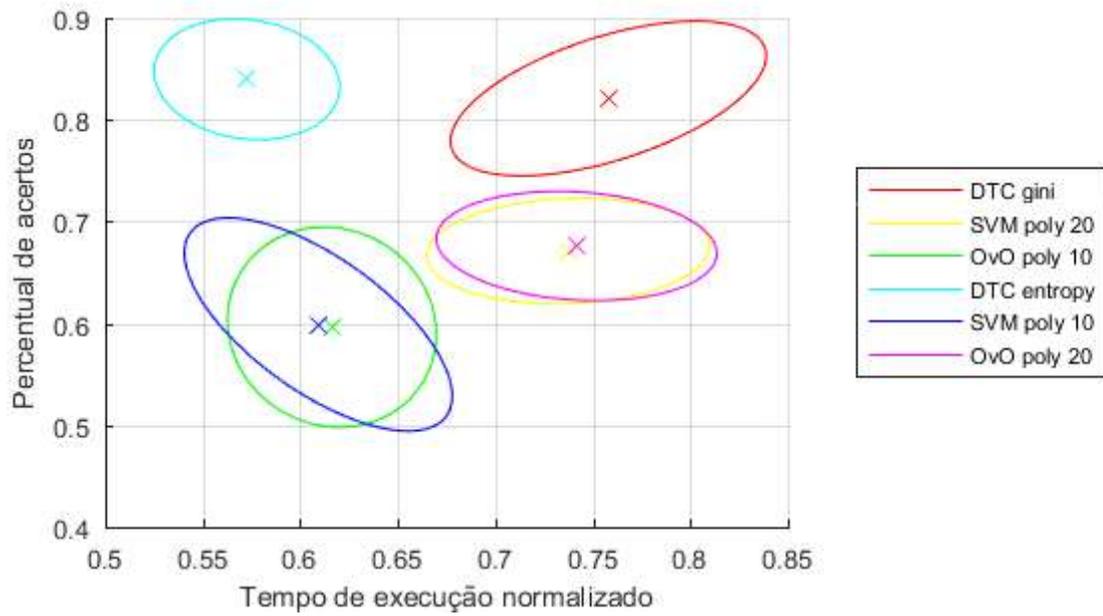


Fonte: Elaboração Própria (2020).

Embora que o algoritmo da DTC com critério gini seja o mais interessante para se utilizar na predição, apresentando melhor precisão e tempo, é interessante apresentar um critério de decisão para a seleção do melhor algoritmo. Eventualmente, o algoritmo da DTC com critério gini pode não ser uma opção para uma determinada aplicação ou, futuramente, podem surgir outros algoritmos que façam levantar a dúvida do melhor algoritmo.

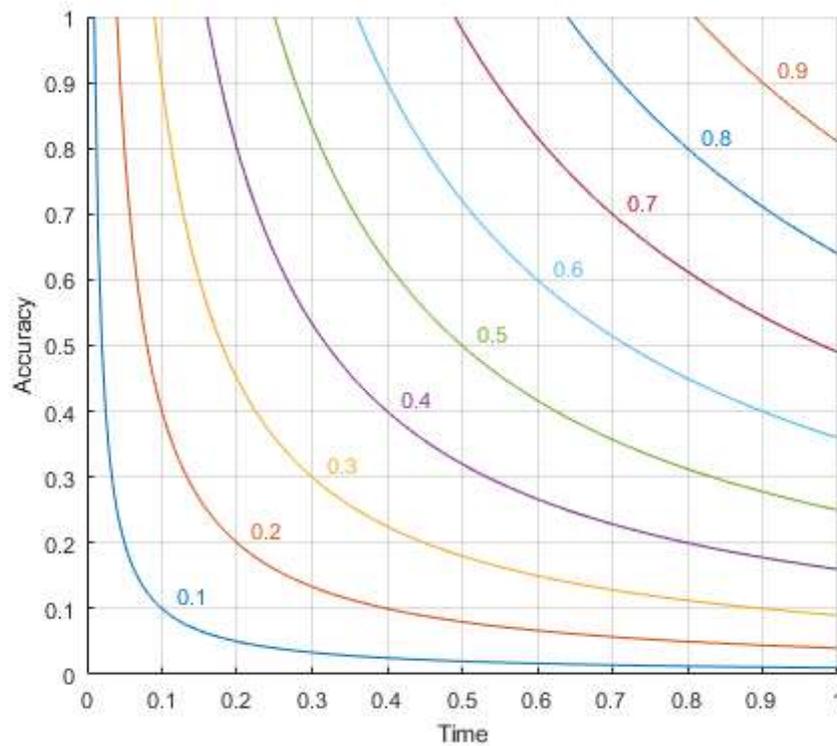
A sugestão deixada pelo estudo é a utilização da média geométrica ponderada entre os critérios de avaliação. A ponderação permite ressaltar uma determinada característica como mais relevante com relação às outras. A Figura 35 mostra um exemplo em que é dado o mesmo peso para a precisão e o tempo do algoritmo.

Figura 34 – Desempenho na detecção de tipo e severidade (elipse)



Fonte: Elaboração Própria (2020).

Figura 35 – Exemplo de critério de avaliação para o melhor algoritmo



Fonte: Elaboração Própria (2020).

O erro aleatório é toda predição feita erroneamente pelo classificador que não possui nenhuma justificativa aparente. É um erro relativamente comum e que costuma ocorrer de forma isolada e não com muita frequência. A própria forma de se modelar computacionalmente o classificador pode gerar alguns desses erros no momento da aprendizagem ou predição do algoritmo. Esse erro não costuma prejudicar muito a predição, pois o mesmo pode facilmente ser contornado com tratamento estatístico que atenuam o impacto do mesmo no desempenho do classificador.

O erro de severidade está associado na incapacidade do classificador em conseguir diferenciar adequadamente a severidade do desbalanceamento. Esse erro caracteriza os casos em que um classificador detecta que a máquina rotativa está sobre um desbalanceamento bem acentuado (nível 3) só que na verdade o desbalanceamento é mais sutil (nível 2). Esse erro jamais ocorre com tipos de desbalanceamentos diferentes. Assim como o erro aleatório, esse erro não costuma prejudicar muito a predição, embora que seja ligeiramente mais severo. A ação de correção sobre a máquina ainda é feita para o tipo de desbalanceamento correto e a realização de novas leituras de corrente após a correção pode contornar esse problema.

O erro de composição está associado ao conceito de que o efeito da variação de corrente do desbalanceamento dinâmico pode ser decomposto de em uma soma de efeitos de desbalanceamento estático com acoplado. Nesse erro, o classificador não consegue observar todas as características do desbalanceamento dinâmico, classificando-o em desbalanceamento estático ou acoplado. Também, esse erro inclui os casos em que o classificador capta informações inexistentes do desbalanceamento estático ou acoplado, classificando-o como desbalanceamento dinâmico. Esse erro não inclui os casos em que o classificador confunde o desbalanceamento estático com acoplado e vice-versa, uma vez que é difícil que esses casos ocorram e, quando ocorre, são classificados como um erro aleatório. O erro de composição é um erro que prejudica o classificador como todo, pois o mesmo pode induzir ação de correção que não vai resolver o problema de desbalanceamento na totalidade, uma vez que o tipo de desbalanceamento predito e real é diferente.

O erro de detecção está associado na incapacidade do classificador de conseguir diferenciar uma máquina saudável com uma máquina desbalanceada a partir do sinal de corrente. Esse erro costuma ter certa frequência nas matrizes de confusão desse tipo, pois a transformada rápida de Fourier de um sinal de corrente de uma máquina rotativa sem desbalanceamento costuma possuir amplitudes muito próximas de uma máquina desbalanceada, principalmente quando a severidade do desbalanceamento é pequena. O estudo aponta que é mais difícil identificar a existência de um desbalanceamento do que o classificar como estático, acoplado ou dinâmico.

A Tabela 7 mostra um exemplo de matriz de confusão obtida no estudo. Todos os valores nulos na matriz foram removidos para facilitar a visualização dos dados que realmente possuem relevância ao estudo. A legenda e as cores utilizadas nos gráficos ajudam a reforçar os conceitos vistos a respeito dos tipos de erros.

Os resultados até o momento apontam que seria interessante fazer futuramente outra análise focada apenas na detecção de desbalanceamento, indiferente do tipo ou severidade do mesmo. Isso se deve ao fato de que as matrizes de confusão obtidas mostram um número considerável de falsos positivos e falsos negativos quanto à detecção de desbalanceamento. A ideia é que, com estes ensaios, seria possível entender melhor o porquê da ocorrência destes erros.

Tabela 7 – Matriz de confusão do algoritmo DTC com critério gini em 5,0 Hz

		Diagnóstico predito										Legenda	
		Ac1	Ac2	Ac3	Di1	Di2	Di3	Es1	Es2	Es3	Sau		
Diagnóstico verdadeiro	Ac1	118		54	22							82	Acerto
	Ac2		221	4	1	30	7	16					
	Ac3	71	1	127	90	1		2				44	Erro aleatório
	Di1	36	1	102	118							37	
	Di2		48	3		124	25	109					Erro de severidade
	Di3		7			53	233	40	3				
	Es1		24			87	22	143					Erro de composição
	Es2						3		188	70			
	Es3						2		87	196			Erro de detecção
	Sau	111		47	28							162	

Fonte: Elaboração Própria (2020).

## 5 CONCLUSÃO

Com relação a todos os objetivos específicos mencionados, o trabalho conseguiu atingir cada um deles com êxito. O conjunto de dados coletados permite uma diferenciação clara para cada tipo e severidade de desbalanceamento. Os gráficos apresentados permitem uma clara visualização desta diferenciação. Esta diferenciação foi reconhecida pelos algoritmos de ML estudados, uma vez que os mesmos conseguiram converter o conjunto de dados em um sistema preditivo que possui certa eficácia, conforme mostra os resultados obtidos.

O estudo apresentado também teve sua relevância reconhecida pela comunidade acadêmica e industrial. Um artigo científico derivado deste TCC foi aceito e apresentado no IX Encontro Nacional de Máquinas Rotativas (ENAM), realizada no ano de 2020. O ENAM é promovido pelo Comitê Nacional Brasileiro de Produção e Transmissão de Energia Elétrica (CIGRE). O ENAM, com periodicidade bianual, apresenta as tendências de tecnologias para máquinas rotativas. O Apêndice A é o artigo científico aprovado. O Apêndice B mostra o comprovante de apresentação, com o link para assistir no Youtube. O Apêndice C mostra os slides utilizados para defender o trabalho. Ainda, o Anexo A mostra o *folder* de chamada do congresso.

O trabalho demonstra que é possível detectar desbalanceamento em rotores através da interpretação do sinal de corrente da máquina rotativa por algoritmos de ML. Todavia, este estudo não traz uma comparação do desempenho de se utilizar esta técnica com o método de detecção por sensores de vibração, também utilizando ML. Sendo assim, uma proposta possível para uma atividade futura é desenvolver um estudo que confronte estes métodos. Esta análise poderia gerar resultados sólidos para apontar qual dos dois métodos mencionados é mais eficaz na predição e também indicar em quais contextos cada método é mais vantajoso.

## REFERÊNCIAS

- BALLUFF. **BAW M12MG2-ICC20B-BP00,2-GS04**. Disponível em: <https://www.alliedelec.com/m/d/1f2c5b0392e509c616ae4c4bf211a665.pdf>. Acesso em: 21 ago. 2020a.
- BALLUFF. **BOS 18KF-PA-1LQP-S4-C**. Disponível em: <http://publications.balluff.com/pdfengine/pdf?type=pdb&id=269146&con=pt&ws=approval>. Acesso em: 21 ago. 2020b.
- BATISTA, Gustavo Enrique de Almeida Prado Alves. **Pré-processamento de Dados em Aprendizado de Máquina Supervisionado**. 2003. 204 f. Tese (Doutorado) - Curso de Ciência da Computação e Matemática Computacional, Instituto de Ciências Matemáticas e de Computação, São Carlos, 2003. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-06102003-160219/publico/TeseDoutorado.pdf>. Acesso em: 22 ago. 2020.
- BOHRA, Prayashi; PALIVELA, Hemant. Understanding and formulation of various kernel techniques for suport vector machines. **2015 IEEE International Conference On Computational Intelligence And Computing Research (ICCIC)**, [s.l.], dez. 2015.
- COURNAPEAU, David. **Scikit-learn**. Disponível em: <https://scikit-learn.org/>. Acesso em: 10 mar. 2020.
- DEMAY, M. B. **Desenvolvimento de sistemas para medição de frequência rotacional em compressores herméticos**. Florianópolis: Universidade Federal de Santa Catarina, 2008.
- DINEVA, Adrienn et al. Fault Diagnosis of Rotating Electrical Machines Using Multi-Label Classification. **Applied Sciences**, [S.L.], v. 9, n. 23, p. 1-18, nov. 2019. Disponível em: <https://www.mdpi.com/2076-3417/9/23/5086/htm>. Acesso em: 04 jan. 2021.
- Edureka. **Decision Tree Algorithm | Decision Tree in Python | Machine Learning Algorithms | Edureka**. Disponível em: <https://www.youtube.com/watch?v=qDcl-FRnwsU>. Acesso em: 21 maio 2020. Vídeo online (46'37")
- FINLEY, William R.; SAUER, Barton J.; LOUTFI, Moheb. Motor Vibration Problems: how to diagnose and correct vibration errors. : How to Diagnose and Correct Vibration Errors. **IEEE Industry Applications Magazine**, [s.l.], v. 21, n. 6, p. 14-28, nov. 2015.
- GALAR, M. *et al.* Enhancing difficult classes in one-vs-one classifier fusion strategy using restricted equivalence functions. **17th International Conference on Information Fusion (FUSION)**. Salamanca: IEEE, 2014.
- GOOGLE. Colaboratory. Disponível em: <https://colab.research.google.com/>. Acesso em: 22 ago. 2020.

HEY, Jonathan *et al.* Online Monitoring of Electromagnetic Losses in an Electric Motor Indirectly Through Temperature Measurement. **IEEE Transactions On Energy Conversion**, [s.l.], v. 31, n. 4, p. 1347-1355, dez. 2016.

HONORATO, Jonas Fortunato *et al.* **Análise de grandezas elétricas para identificação do desbalanceamento de massa em cargas acopladas a motores de indução trifásicos alimentados por inversor**. 2017. 148 f. TCC (Graduação) - Curso de Engenharia Elétrica, Departamento Acadêmico de Eletrotécnica, Instituto Federal de Santa Catarina, Florianópolis, 2017.

IMI SENSORS. **Model 601A01**: Installation and Operating Manual. Disponível em: [https://www.pcb.com/contentstore/docs/PCB\\_Corporate/IMI/Products/Manuals/601A01.pdf](https://www.pcb.com/contentstore/docs/PCB_Corporate/IMI/Products/Manuals/601A01.pdf). Acesso em: 21 ago. 2020.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 1940-1**: Balance quality requirements for rotors in a constant (rigid) state. 2 ed. Genebra: Iso, 2003.

KLEIN, Dan. **Lagrange Multipliers without Permanent Scarring**. Disponível em: <https://people.eecs.berkeley.edu/~klein/papers/lagrange-multipliers.pdf>. Acesso em: 22 ago. 2020.

LEM. **Current Transducer LA 25-NP**. Disponível em: <https://pdf1.alldatasheet.com/datasheet-pdf/view/81509/LEM/LA25-NP.html>. Acesso em: 21 ago. 2020a.

LEM. **Voltage Transducer LV 20-P**. Disponível em: <https://pdf1.alldatasheetpt.com/datasheet-pdf/view/828813/LEM/LV20-P.html>. Acesso em: 21 ago. 2020b.

MATHWORKS. **Matlab**. Disponível em: <https://www.mathworks.com/products/matlab.html>. Acesso em: 10 mar. 2020.

MATSUO, Tiago Kaoru *et al.* **Método de análise de vibrações em máquinas rotativas no processo de geração de energia elétrica, validado com dados históricos e simulação**. 2017. 158 f. Dissertação (Mestrado) - Curso de Mecatrônica, Instituto Federal de Santa Catarina, Florianópolis, 2017.

MATSUSHITA, O. *et al.* **Vibrations of Rotating Machinery**. [s.l.]: Springer, 2017.

MCKINNEY, Wes. **Pandas**. Disponível em: <https://pandas.pydata.org/>. Acesso em: 10 mar. 2020.

MEAN WELL. **RT 65 series**: 65W Triple Output Switching Power Supply. Disponível em: <https://www.meanwell-web.com/content/files/pdfs/productPdfs/MW/Rt-65/RT-65-spec.pdf>. Acesso em: 21 ago. 2020.

MESHGIN-KELK, Homayoun *et al.* **Mechanical vibration — Balance quality requirements for rotors in a constant (rigid) state —**. Boca Raton: CRC Press, 2017. 272 p.

MINIPA. **Fonte de alimentação DC regulada**. Disponível em: <http://www.minipa.com.br/images/Manual/MPL-1303M-1305M-3303M-3305M-1104-BR.pdf>. Acesso em: 21 ago. 2020.

MIT OPENCOURSEWARE. **Learning: Identification Trees, Disorder**. Disponível em: [https://www.youtube.com/watch?v=SXBG3RGr\\_Rc&list=PLUI4u3cNGP63gFHB6xb-kVBiQHye\\_4hSi&index=11](https://www.youtube.com/watch?v=SXBG3RGr_Rc&list=PLUI4u3cNGP63gFHB6xb-kVBiQHye_4hSi&index=11). Acesso em: 21 maio 2020a. Vídeo online (49'36")

MIT OPENCOURSEWARE. **Learning: Support Vector Machines**. Disponível em: [https://www.youtube.com/watch?v=\\_PwhiWxHK8o](https://www.youtube.com/watch?v=_PwhiWxHK8o). Acesso em: 21 maio 2020b. Vídeo online (49'33")

MOHAMMED, Mohssen; KHAN, Muhammad Badruddin; BASHIER, Eihab Bashier Mohammed. **Machine Learning: algorithms and applications**. Boca Raton: Crc Press, 2017.

MOTOROLA. **Octal High Voltage, High Current Darlington Transistor Arrays**. Disponível em: <https://pdf1.alldatasheet.com/datasheet-pdf/view/12687/ONSEMI/ULN2803.html>. Acesso em: 21 ago. 2020.

NATIONAL INSTRUMENTS. **DAQ M Series: NI USB-621x user manual**. Austin: NI, 2009. Disponível em: <https://www.ni.com/pdf/manuals/371931f.pdf>. Acesso em: 21 ago. 2020.

NATIONAL INSTRUMENTS. **Aquisição de sinais analógicos: largura de banda, teorema de amostragem de Nyquist e aliasing**. Disponível em: <https://www.ni.com/pt-br/innovations/white-papers/06/acquiring-an-analog-signal--bandwidth--nyquist-sampling-theorem-.html>. Acesso em: 10 mar. 2020a.

NATIONAL INSTRUMENTS. **Labview**. Disponível em: <https://www.ni.com/ptbr/shop/labview.html>. Acesso em: 10 mar. 2020b.

OLIPHANT, Travis; **Numpy**. Disponível em: <https://numpy.org/>. Acesso em: 10 mar. 2020a.

OLIPHANT, Travis; PETERSON, Pearu; JONES, Eric. **Scipy**. Disponível em: <https://www.scipy.org/>. Acesso em: 10 mar. 2020b.

PHOENIX CONTACT. **MACX MCR-UI-UI-UP(-SP)(-NC)**. Disponível em: <https://media.digikey.com/pdf/Data%20Sheets/Phoenix%20Contact%20PDFs/MACX%20MCR-UI-UI-UPx.pdf>. Acesso em: 21 ago. 2020.

ROCKAH, Lior; MAIMON, Oded. **Data mining with decision trees: theory and applications**. 2. ed. Singapura: World Scientific, 2014. (Machine Perception and Artificial Intelligence).

RUSSELL, Rudolph. **Machine Learning: step-by-step guide to implement machine learning algorithms with python**. Pdf Drive, 2018. Disponível em: <https://www.pdfdrive.com/machine-learning-step-by-step-guide-to-implement-machine-learning-algorithms-with-python-e158324853.html>. Acesso: 04 jan. 2021.

SAMMUT, Claude; WEBB, Geoffrey I. **Encyclopedia of Machine Learning**. Nova York: Springer, 2011.

SANTOS, H. F. dos *et al.* Non-invasive monitoring system of synchronous generator using external field. **Journal Of Microwaves, Optoelectronics And Electromagnetic Applications**, [s.l.], v. 16, n. 1, p. 70-89, mar. 2017.

SHWARTZ, Shai Shalev; DAVID, Shai Ben. **Understanding Machine Learning: from theory to algorithms**. Nova York: Cambridge University Press, 2014.

SMITH, Amelia. **Learning Each Function with Machine Learning**. Disponível em: <https://medium.com/@ameliasmithml/learning-each-function-with-machine-learning-264dbaae0e20>. Acesso em 22 ago. 2020.

THOMSON, William T.; CULBERT, Ian. **Current Signature Analysis for Condition Monitoring of Cage Induction Motors**: industrial application and case histories. Series: IEEE Press Series On Power Engineering. Wiley-IEEE - Press, 2017.

## **APÊNDICES**

## APÊNDICE A – Artigo completo

**IDENTIFICAÇÃO DO TIPO E DA SEVERIDADE DO DESBALANCEAMENTO DE ROTORES POR MEIO DA ANÁLISE DA ASSINATURA DE CORRENTE COM APRENDIZADO DE MÁQUINA MULTICLASSES**

R. N. Tominaga

F. S. Pacheco

C. A. Penz

S. L. Avila

IFSC

IFSC

IFSC

IFSC

**Resumo**

A saúde de uma máquina elétrica rotativa pode ser avaliada através do monitoramento de grandezas elétricas e mecânicas associadas à sua operação. Quanto mais informação estiver disponível, mais fácil pode ser diagnosticado o seu estado operacional. Para se estudar esse tema, foi construída uma bancada de teste que permite inserir desbalanceamentos propositalmente sobre o eixo de um motor de indução. No presente trabalho, fica-se restrito a diagnósticos utilizando os harmônicos de corrente do motor. A partir dessa assinatura, e entendendo-se que se trata de um problema com características não-lineares e multiclasse, fez-se o uso de ferramentas computacionais para proceder a identificação (tipo) e a classificação (severidade) dos desbalanceamentos. Os algoritmos estudados aqui foram Máquina de Vetores de Suporte (SVM), Árvore de Decisão (DTC) e *One-vs-One* (OvO). Os resultados obtidos atestam a eficiência da metodologia proposta. A comparação das estratégias de aprendizado de máquina, bem como os cuidados necessários quanto ao processamento de sinais, também são contribuições deste trabalho para a engenharia de manutenção.

**Palavras chave**

Máquinas elétricas rotativas; Análise de sinais de corrente; Identificação de desbalanceamento; Aprendizado de máquina.

**1 INTRODUÇÃO**

O diagnóstico eficaz de falhas, por meio da instrumentação de monitoramento em tempo real, pode melhorar a segurança e a confiabilidade na operação de máquinas rotativas [1]. Ele também pode fornecer uma base sólida nos fundamentos da causa e efeito da falha. A maioria das técnicas existentes para monitoramento de condições e diagnóstico de falhas é baseada em medições de corrente [2], vibração [3], campo eletromagnético externo [4] e/ou temperatura [5].

A análise de assinatura de corrente do motor (MCSA) tem sido usada para monitoramento de condições em aplicações industriais [2]. Uma vantagem do MCSA em relação a outras técnicas é a simplicidade de sua instrumentação. Não há requisitos para ajustes mecânicos complexos, como os necessários para sensores de vibração; ou mesmo atenção para a complexidade do ambiente, como a presença de campos eletromagnéticos externos. Entretanto, cabe ressaltar que a interpretação da MCSA não é evidente, devido principalmente aos ruídos inerentes ao sinal em ambiente industrial. Assim, MCSA necessita de algoritmos complementares que em geral são custosos computacionalmente e dependem da experiência de um profissional.

O principal benefício do emprego de algoritmos de aprendizado de máquina (ML) é a menor dependência da experiência no problema para se obter resultados satisfatórios. A aplicação de técnicas de ML para o diagnóstico de comportamento de máquinas rotativas se tornou alvo de muitas pesquisas nos últimos anos.

É consenso que, quanto mais informações se têm, mais fácil é o gerenciamento da integridade de um sistema. Apesar disso, diferentemente dos ambientes controlados de laboratório, em plantas industriais pode ser difícil realizar a instalação de sensores em local e/ou em condições adequadas.

Nesse contexto, realizou-se este estudo. Para uma questão industrial relevante e bastante comum, o processo de desalinhamento do rotor de um motor elétrico, e com uma instrumentação de baixo custo, simples e fácil de instalar (MCSA), desenvolveu-se uma metodologia que usa ML para identificar o tipo e a severidade de falhas de desbalanceamento.

A classificação multiclasse significa identificar três ou mais tipos de ocorrências. De acordo com a ISO 1940-1 [6], existem quatro tipos de falhas de desbalanceamento com três níveis de gravidade cada.

Qualquer máquina deve ser parada para manutenção caso seja desejável estender a sua vida útil. Quando o equipamento retorna ao estado natural de operação, o transitório da corrente indicará a condição saudável da máquina. Conseqüentemente, recomenda-se um procedimento de ML dito supervisionado: a comparação direta entre dados ditos saudáveis e dados ainda não classificados. Entre as técnicas mais relevantes, destaca-se Máquina de Vetores de Suporte (SVM), Árvore de Decisão (DTC) e *One-vs-One* (OvO). Elas se destacam, pois resolvem com eficiência uma classificação multiclasse e não linear. Não linear devido à disparidade da operação das condições do motor, por exemplo, variações de carga e fonte de alimentação.

Este artigo apresenta a metodologia de diagnóstico desenvolvida e realiza um comparativo entre as três técnicas de ML citadas.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Desbalanceamentos em rotores

Define-se desbalanceamento do motor como o desequilíbrio de massa existente com relação ao eixo de rotação do motor. Alternativamente, o desbalanceamento é a distância existente entre o centro de massa com o eixo de rotação do motor. Assimetrias, tolerâncias dos desvios de forma e imperfeições da montagem são responsáveis pela existência do desbalanceamento em motores [6][7].

Existem quatro tipos de desbalanceamento quando é levada em consideração a posição relativa entre os eixos de rotação e de massa [8]: estático, quase estático, acoplado e dinâmico. A Figura 1 representa esses tipos de desbalanceamentos.

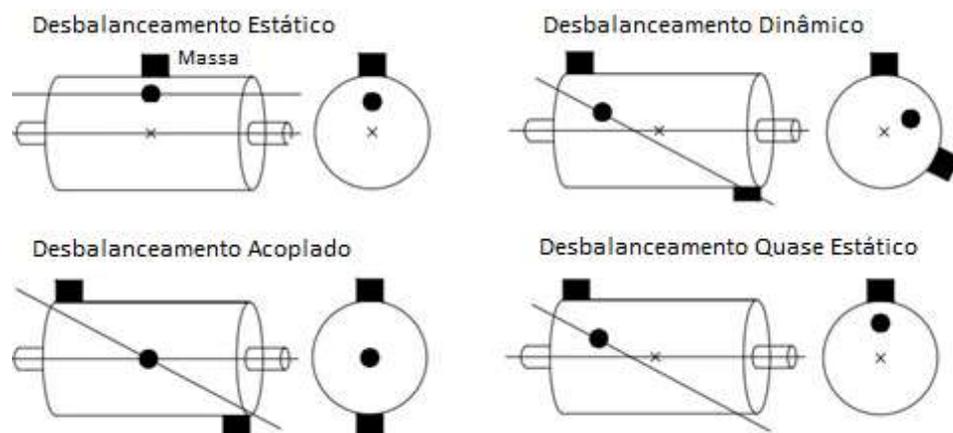


Figura 1: Tipos de desbalanceamento em rotores

O desbalanceamento estático ocorre quando o eixo de massa está deslocado paralelamente ao eixo de rotação. Esta denominação se deve ao fato de que quando a máquina é desligada, o eixo tende a girar até que o ponto de maior massa esteja para baixo, tornando-o facilmente detectável. Este desbalanceamento pode ser facilmente corrigido com a adição de uma massa de balanceamento diametralmente oposta ao centro de gravidade em um plano perpendicular ao eixo de rotação.

No desbalanceamento acoplado o eixo de massa intercepta o eixo de rotação no centro de gravidade do rotor. Este desbalanceamento é equivalente a duas massas em dois planos diferentes, deslocadas de  $180^\circ$  entre si. A amplitude desse desbalanceamento é determinada pelas massas responsáveis pelo desbalanceamento, pelo raio existente entre o eixo de rotação e as massas e pela distância entre elas. Este desbalanceamento pode ser corrigido com a adição de duas massas de balanceamento também em planos opostos e deslocadas de  $180^\circ$  entre si, de tal forma que as massas fiquem diametralmente opostas.

O desbalanceamento dinâmico ocorre quando o eixo de massa não é paralelo ao eixo de rotação e nem o intercepta. Este é o desbalanceamento mais comum, e pode ser corrigido com a adição de pelo menos duas massas em planos perpendiculares ao eixo de rotação. Em um rotor real, é comum que ocorra uma série de pequenos desbalanceamentos que podem ser definidos por um desbalanceamento dinâmico equivalente. Este desbalanceamento pode ser representado pela combinação de um desbalanceamento estático e um acoplado.

No desbalanceamento quase estático, o eixo de massa intercepta o eixo de rotação em um ponto diferente do centro de gravidade. Este desbalanceamento pode ser considerado um caso especial do desbalanceamento dinâmico.

## 2.2 Ferramentas de aprendizado de máquina

O ML consiste no desenvolvimento de um algoritmo capaz de executar uma tarefa sem a utilização de instruções explícitas. Esses algoritmos executam as tarefas com bases em padrões e inferências de um conjunto de dados. Aqui serão utilizadas as estratégias de aprendizado OvO, DTC e SVM.

O OvO é um algoritmo que classifica através de agrupamento de pares de classe. Usualmente, o OvO é utilizado para comparar apenas duas classes. Contudo, é possível fazer um agrupamento deste método permitindo a sua transformação em um classificador multiclasse. No momento da predição, a classe que receber mais votos é escolhida para classificar a amostra. Devido a esse agrupamento, o OvO é computacionalmente mais custoso que os demais [9].

O DTC é um algoritmo que busca uma generalização de regras para a classificação das amostras. Essas regras visam descrever as características mais marcantes de um conjunto até chegar a uma classificação. DTC é hábil tanto para problemas binários quanto para problemas multiclases [10].

O SVM é um algoritmo que visa distinguir classes através da modelagem de um hiperplano que melhor separa essas classes. O hiperplano é gerado através da maximização das distâncias mínimas geradas para cada uma das classes a partir do mesmo. Existe uma variação desse algoritmo que é válido para problemas multiclases, que a Máquina de Vetores de Suporte Multiclasse (MSVM) [11].

## 3 BANCADA DE DESBALANCEAMENTO

A bancada de desbalanceamento pode ser dividida em duas partes: a primeira relativa ao aspecto eletromecânico da bancada; e a segunda referente à instrumentação que é utilizada para a aquisição dos sinais.

### 3.1 Aspecto eletromecânico

A bancada de desbalanceamento consiste em uma estrutura metálica sobre cavalete metálico com *vibrastop*. A estrutura metálica facilita no aterramento da instrumentação. O *vibrastop* ajuda no amortecimento de certos ruídos que são indesejados na aquisição dos sinais.

Sobre a estrutura metálica foi instalado um motor de indução trifásico WEG 220/380 V; 60 Hz; 4 polos; 0,75 HP. O motor escolhido é de pequeno porte para emular as principais aplicações industriais.

Para o acionamento do motor é utilizado um inversor da fabricante WEG CFW500. Este inversor possui faixas de tensão de 0 a 600 V, corrente máxima de 200 A e frequência de 0 a 500 Hz. A utilização do inversor para a bancada foi necessária devido à necessidade de flexibilidade em alterar a velocidade do motor.

Ao eixo do motor é acoplado a uma peça metálica simétrica e sem desequilíbrio de massa. Esta peça possui três discos de inércia com 24 furos igualmente espaçados para permitir a inserção de desbalanceamentos propositais. O acoplamento foi projetado para minimizar o desalinhamento do sistema. Esta peça é apoiada por dois mancais de rolamentos idênticos com base rígida.

A Figura 2 apresenta a bancada. Além dos componentes mencionados anteriormente, a foto permite apresentar o como é feito a inserção dos desbalanceamentos (A). Também, são apresentados o sensor óptico (B), os acelerômetros (C) e os sensores indutivos (D). Os detalhes do senso óptico e da instrumentação utilizada para adquirir os sinais de corrente são mostrados na próxima seção.

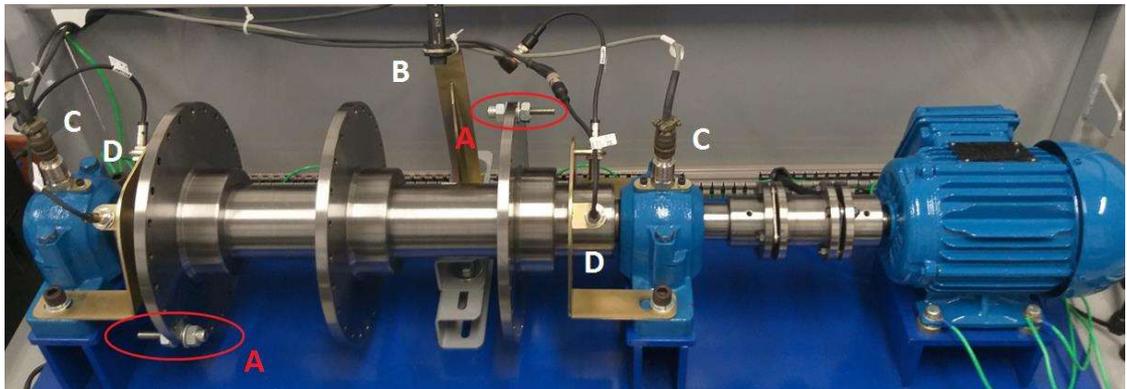


Figura 2: Estrutura metálica da bancada de desbalanceamento

### 3.2 Instrumentação

Para a escolha do dispositivo de aquisição dos sinais, foi considerado a afinidade para trabalhar com o mesmo e o seu poder de processamento. Neste contexto, o escolhido foi a placa USB 6212 da *National Instruments*. O equipamento possui 16 entradas analógicas com 4 faixas de entrada programáveis ( $\pm 0,2$  V a  $\pm 10$  V) com resolução de 16 bits, taxa de amostragem de 400 kS/s, 2 saídas analógicas, 32 linhas de E/S digital. A entrada USB do dispositivo é utilizada tanto para alimentação do módulo quanto para transferir os dados para o computador.

Para aumentar a resolução da leitura foi optado pela utilização de entradas diferenciais, ou seja, duas entradas analógicas são utilizadas para a leitura de uma grandeza. Devido a este fato, foi necessária a utilização de dois dispositivos de aquisição. Foi definido que um desses dispositivos estaria destinado a coletar grandezas de natureza mecânica, provenientes de sensores indutivos, acelerômetros e sensor óptico e o outro destinado para coletar grandezas de natureza elétrica, provenientes dos transdutores de corrente.

O modelo dos sensores indutivos utilizados é BAW M12MG2-ICC20B-BP00,2-GS04. Possui uma alimentação de 24 Vcc e é capaz de representar a faixa de 0,5 a 2 mm em um sinal de corrente de 4 a 20 mA. Para cada mancal, existem dois sensores indutivos perpendiculares entre si. Para que o sinal possa ser interpretado pelo dispositivo de aquisição dos sinais, para cada sensor indutivo é utilizado um resistor de 240  $\Omega$ .

Os acelerômetros utilizados são do modelo 601A02, pertencente ao fabricante IMI Sensors. Este modelo possui uma sensibilidade de 10,2 mV/(m/s<sup>2</sup>), faixa de medição de  $\pm 490$  m/s<sup>2</sup> e resposta em frequência de 0,27 a 10.000 Hz. Os mesmos ficam acoplados aos mancais existentes na bancada. Para que o sinal possa ser interpretado pelo dispositivo de aquisição dos sinais, para cada acelerômetro foi utilizado um condicionador de sinal. O modelo do condicionador é MACXMCR-UI-UI-UP, pertencente ao fabricante Phoenix Contact.

O sensor óptico utilizado é do modelo BOS 18KF-PA-1LQP-S4-C, pertencente ao fabricante Balluff. Pode ser alimentado de 10 a 30 V em corrente contínua e o mesmo possui um contato normalmente aberto e outro fechado. Este foi instalado próximo ao eixo metálico que possui os discos de inércia. Para que o sinal possa ser interpretado pelo dispositivo de aquisição dos sinais, é utilizado um divisor de tensão composto por dois resistores de 10 k $\Omega$ .

O modelo do transdutor de corrente utilizado é LA 25-NP, pertencente ao fabricante LEM. É um transdutor de efeito Hall e possui isolamento galvânica. Este transdutor necessita de uma alimentação simétrica de 15 Vcc. O mesmo foi inserido em uma placa eletrônica de fabricação caseira de tal forma que o transdutor interprete o sinal de corrente de até 5 A e converta para um sinal de 25 mA. O sinal é convertido em tensão através de um resistor também inserido na placa.

A placa eletrônica, ainda, possui dois relés que permitem a isolamento do transdutor de corrente. Os relés são comandados por um circuito integrado ULN2803, que possui um conjunto de transistores Darlington, e o mesmo recebe um sinal da saída digital do dispositivo de aquisição dos sinais. Na bancada são utilizados três transdutores de corrente a fim de mensurar a corrente que percorre os três enrolamentos do motor de indução.

## 4 AQUISIÇÃO DOS SINAIS

Foi construído um conjunto de dados com a condição saudável da máquina e quando a mesma é sujeita aos desbalanceamentos estático, acoplado e dinâmico. Para poder fazer um teste de versatilidade entre os algoritmos de ML, o conjunto de dados foi construído para três velocidades distintas e três níveis de severidade para os desbalanceamentos.

As três classes de velocidade do rotor foram obtidas com base na variação da frequência de alimentação do motor. O inversor de frequência permite fazer tal variação. Foram utilizadas as velocidades para frequência de 5 Hz, 7,5 Hz e 10 Hz, o que resulta em velocidades de 150 RPM, 225 RPM e 300 RPM, respectivamente, do rotor.

Os três níveis de severidade de desbalanceamento foram definidos com base na variação da massa que é inserido no disco de inércia. Foram definidos os níveis 1, 2 e 3, na qual quanto maior o índice, maior é o peso da massa que causa o desbalanceamento. As massas foram inseridas no disco de inércia através de um parafuso e foram utilizadas algumas porcas para poder aumentar a massa nos pontos desejados. Vale ressaltar ainda que foi certificado de que os parafusos estivessem bem apertados ao furo para garantir que não houvesse vibração. Essa vibração poderia prejudicar a qualidade do conjunto de dados.

Para a construção do conjunto de dados foi necessária a utilização do software Labview [12]. Para cada amostra foi utilizado uma frequência de amostragem de 50.000 amostras por segundos e foram registrados no total de 200.000 pontos por amostra. Isto resulta em uma janela de 4 segundos. Estes valores foram escolhidos para que os resultados obtidos possam ser comparados a outros trabalhos que já foram realizados na mesma bancada.

Foi levada em consideração a preocupação de efetuar todos os ensaios no menor período possível para evitar interferências externas que possam estar poluindo o conjunto de dados. Também foi levada em consideração a preocupação de garantir que o sistema entre em regime antes de aferir as grandezas para evitar alterações no conjunto de dados.

## 5 PROCESSAMENTO DE DADOS

Após a criação do conjunto de dados, foi pensado em como as informações contidas deveriam ser modeladas para que pudesse usar os algoritmos de ML apresentados. Para isso, foi utilizado o software Matlab [13]. A decisão para a utilização deste software foi dada pela familiaridade e facilidade para a manipulação das informações que o mesmo fornece.

A utilização de muitos processos para formatar o conjunto de dados podem inutilizar informações ou características do conjunto de dados. Isso não é interessante para um algoritmo de ML, pois o mesmo teria maiores dificuldades para inferir um diagnóstico com exatidão.

Todavia, um sistema preditivo é considerado eficaz quando o mesmo consegue captar apenas as características mais relevantes do conjunto de dados para inferir um resultado. Qualquer característica que não tem relevância pode viciar um sistema preditivo a gerar padrões incongruentes. Esses padrões podem diminuir a eficácia do algoritmo. Sendo assim, utilizar o conjunto de dados exatamente da forma como foi adquirido da bancada não é uma boa opção. O sinal no domínio do tempo possui 200.000 pontos por amostra. Utilizar todos estes pontos é inviável para um algoritmo de aprendizado uma vez que isso iria requerer muito processamento computacional.

Também, para cada amostra, existe uma defasagem angular específica. Essa defasagem iria dificultar muito o algoritmo de aprendizado, pois o mesmo iria tentar buscar algum padrão através da existência dessa defasagem para gerar o diagnóstico. O problema é que não existe nenhuma relação com o tipo de desbalanceamento com essa característica.

Para solucionar esse problema, foi optada a utilização da transformada rápida de Fourier (FFT). Sendo assim, o sinal no domínio do tempo é transformado para o domínio da frequência. Isso resolve o problema relacionado com a defasagem angular do sinal. Todavia, ainda existem muitas componentes para representar a amostra. Sendo mais específico, existem 25.000 componentes por amostra, conforme o teorema de Nyquist.

Para facilitar o treinamento do algoritmo, foi utilizado um critério visual que consiste selecionar as componentes que possuem variações na amplitude de acordo com o tipo de desbalanceamento. Para uma frequência fundamental  $f_1$ , sendo esse valor podendo ser 5 Hz, 7,5 Hz e 10 Hz, foi selecionado as frequências  $0,5*f_1$ ,  $1,0*f_1$ ,  $1,5*f_1$ ,  $3,0*f_1$ ,  $5,0*f_1$  e  $7,0*f_1$ . As demais componentes foram dispensadas ou porque possuem uma amplitude muito pequena, o que torna mais susceptível a ruídos e, portanto, menor a confiabilidade, ou porque não trazia informações novas para os múltiplos de frequência selecionados.

Para confirmar que os múltiplos de frequências selecionados realmente possuem a capacidade de poder distinguir o conjunto de dados entre as possíveis classificações de desbalanceamento, foi desenvolvido histogramas para averiguar variações nas amplitudes de acordo com o tipo de desbalanceamento.

A Figura 3 mostra parte dos histogramas construídos para a análise de dados. Foi adotado um critério de cores que facilita a visualização do tipo de desbalanceamento conforme mostra a legenda localizada no canto direito.

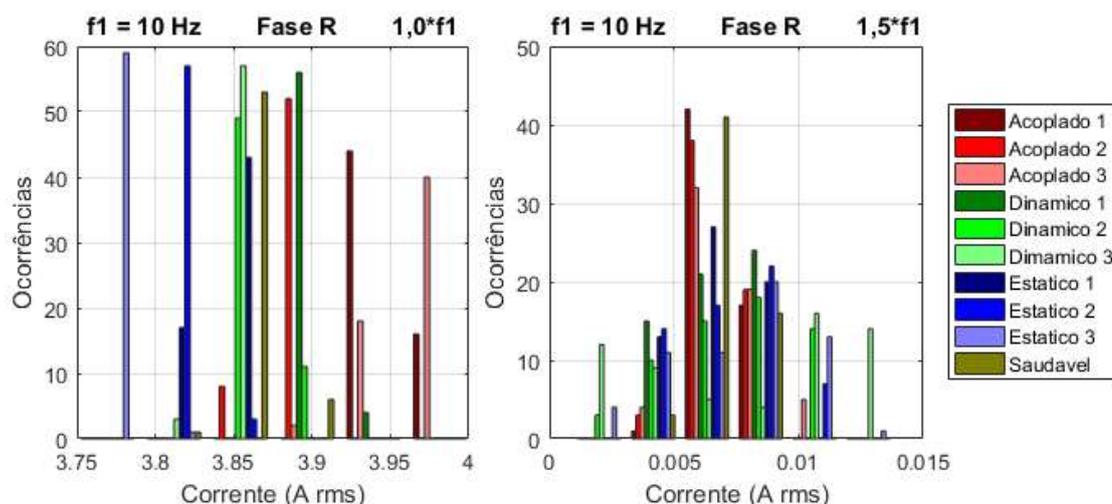


Figura 3: Histograma da corrente para diferentes casos de desbalanceamento

É possível visualizar que no desbalanceamento estático existe uma redução na amplitude da componente fundamental do sinal quando comparado com a condição saudável da máquina. Esta redução é proporcional à intensidade do desbalanceamento.

No desbalanceamento acoplado é possível visualizar que existe um aumento na amplitude da componente fundamental do sinal quando comparado com a condição saudável da máquina. Este aumento é proporcional à intensidade do desbalanceamento.

Para o desbalanceamento dinâmico é difícil fazer uma distinção em relação com a condição saudável da máquina, pois a amplitude da fundamental em ambos os casos é parecida. Todavia, quando é feita uma análise visual na componente que é 1,5 vezes a componente fundamental, o cenário muda. É possível visualizar que existe uma oscilação na amplitude nessa componente no caso do desbalanceamento dinâmico enquanto que, na condição saudável, não é possível visualizar essa oscilação.

Vale ressaltar ainda que a componente que é 1,5 vezes a componente fundamental possui certa dificuldade para distinguir o desbalanceamento estático e ainda possui mais dificuldade para distinguir o desbalanceamento acoplado.

Os outros múltiplos de frequência que não foram mostrados também possuem certa contribuição para a distinção dos tipos de desbalanceamento e também é possível fazer uma análise visual semelhante ao que foi mostrado para os múltiplos  $1,0 \cdot f_1$  e  $1,5 \cdot f_1$ .

Outra conclusão que foi obtida na inspeção visual é o fato de que é mais fácil reconhecer o tipo de desbalanceamento quando o mesmo ocorrer em maior intensidade e frequência.

Para fazer as análises de ML, usou-se o *Colaboratory*, uma plataforma do Google. Este software permite trabalhar com a linguagem Python e, dentre dos módulos que existem, possui um que é voltado especificamente para o ML. Este módulo se chama *Scikit learn* [14].

Também foram utilizados outros módulos para as análises dos dados. São eles: *Numpy* [15], para a realização de operações matriciais; *Scipy* [16], para importar o conjunto de dados; e o *Pandas* [17], para melhorar a visualização dos resultados.

Foi tomado como critério destinar 5/6 do conjunto de dados para o treinamento dos algoritmos de aprendizado e o restante para avaliar a eficácia do treinamento. Foi assegurado que a mesma divisão fosse adotada para todos os algoritmos, de forma a permitir uma comparação direta entre eles.

Para cada um dos algoritmos de ML, foram feitas as seguintes análises: predição para determinar tipo e severidade do desbalanceamento; predição para determinar apenas tipo de desbalanceamento; e predição para determinar apenas severidade de desbalanceamento.

As análises têm como objetivo testar a eficácia de cada um dos algoritmos em cada uma das situações propostas e concluir se é mais interessante fazer uma análise simultânea de tipo e severidade ou se é mais interessante fazer as outras duas análises em cascata.

Após estas análises, foi plotada uma matriz de confusão para facilitar a visualização de erros cometidos pelos algoritmos [18]. Esta matriz confronta os diagnósticos reais dos dados de teste com os diagnósticos gerados pelo algoritmo. Os elementos pertencentes à diagonal principal desta matriz informam os casos em que o algoritmo realizou uma predição correta. Os demais elementos mostram casos em que o algoritmo se equivocou na predição. A ideia desta providência é facilitar a visualização dos valores que são relevantes para o estudo.

O Quadro 1 mostra a matriz de confusão referente ao algoritmo de SVM, aplicada sobre o sinal de corrente da fase R e quando o motor operava com uma frequência de alimentação de 5,0 Hz. Considerando a mesma semente para a separação do conjunto de dados em treino e teste utilizada para definir o Quadro 1, quando é feita uma análise de tipo e severidade de desbalanceamento simultânea, o percentual de acerto é de apenas 56%. Quando é feita uma análise de tipo de desbalanceamento apenas, o percentual de acerto é de 56% também. Quando é feita uma análise de severidade de desbalanceamento apenas, o percentual de acerto é de 50%.

O Quadro 2 mostra outra matriz de confusão gerada através do algoritmo de OvO, aplicada sobre o sinal de corrente da fase S e quando o motor operava com uma frequência de alimentação de 7,5 Hz. Considerando a mesma semente para a separação do conjunto de dados em treino e teste utilizada para definir o Quadro 2, quando é feita uma análise de tipo e severidade de desbalanceamento simultânea, o percentual de acerto é de 74%. Quando é feita uma análise de tipo de desbalanceamento apenas, o percentual de acerto é de 77%. Quando é feita uma análise de severidade de desbalanceamento apenas, o percentual é de 51%.

O Quadro 3 mostra mais uma matriz de confusão proveniente do algoritmo de DTC, aplicada sobre o sinal de corrente da fase T e quando o motor operava com uma frequência de alimentação de 10,0 Hz. Considerando a mesma semente para a separação do conjunto de dados em treino e teste utilizada para definir o Quadro 3, quando é feita uma análise de tipo e severidade de desbalanceamento simultânea, o percentual de acerto é de 80%. Quando é feita uma análise de tipo de desbalanceamento apenas, o percentual de acerto é de 88%. Quando é feita uma análise de severidade de desbalanceamento apenas, o percentual de acerto é de 83%.

		Matriz de confusão de tipo e severidade – SVM										Apenas tipo				
		A1	A2	A3	D1	D2	D3	E1	E2	E3	Sa	Ac	Di	Es	Sa	
Diagnóstico verdadeiro	A1	2		3							5	Ac	17	12		
	A2		6			2						Di	7	15	9	
	A3	4		6							1	Es		8	24	
	D1	2		5	1						2	Sa	8			0
	D2		2			4	3	5				Apenas severidade				
	D3						7						N1	N2	N3	Sa
	E1		1			1	1	7				N1	18	10		2
	E2								8	2		N2		29	3	
	E3									12		N3	9	19	2	
	Sa	2		3							3	Sa	7			1
		Diagnóstico predito														

Quadro 1: Matriz de confusão do SVM na fase R com frequência fundamental de 5,0 Hz

		Matriz de confusão de tipo e severidade – OvO										Apenas tipo				
		A1	A2	A3	D1	D2	D3	E1	E2	E3	Sa	Ac	Di	Es	Sa	
Diagnóstico verdadeiro	A1	9		1								Ac	25	3		1
	A2		8									Di	5	20	2	4
	A3			5	5						1	Es		3	29	
	D1	1	1	3	2						3	Sa	2	3		3
	D2					10		4				Apenas severidade				
	D3						7					N1	N2	N3	Sa	
	E1		1					8			1	N1	20	2	8	
	E2								10			N2		14	18	
	E3									12		N3	10	3	17	
	Sa			3	2						3	Sa	5		3	0
		Diagnóstico predito														

Quadro 2: Matriz de confusão do OvO na fase S com frequência fundamental de 7,5 Hz

		Matriz de confusão de tipo e severidade - DTC										Apenas tipo				
		A1	A2	A3	D1	D2	D3	E1	E2	E3	Sa	Ac	Di	Es	Sa	
Diagnóstico verdadeiro	A1	10										Ac	28			1
	A2		7								1	Di	7	22	2	
	A3	3		6	2							Es		1	31	
	D1		2		8							Sa	1			7
	D2		2			8		2			2	Apenas severidade				
	D3		1			1	5					N1	N2	N3	Sa	
	E1						1	8	1			N1	25	2	2	1
	E2						1		9			N2	2	28	1	1
	E3									12		N3	5	2	23	
	Sa					1					7	Sa		1		7
		Diagnóstico predito														

Quadro 3: Matriz de confusão da DTC na fase T com frequência fundamental de 10,0 Hz

## 6 DISCUSSÃO DOS RESULTADOS E CONCLUSÃO

Independente do algoritmo de aprendizado que é utilizado, os melhores resultados são encontrados nas situações em que o motor opera com maior frequência de alimentação. Com maior velocidade, mais vibração, mais fácil o diagnóstico de desbalanceamento.

No algoritmo da DTC, aplicada na fase R e buscando uma predição de nível e severidade simultânea, o percentual de acertos nas frequências fundamentais 5,0, 7,5 e 10,0 Hz são, respectivamente, 54%, 78% e 82%. Para os algoritmos de SVM e OvO, o kernel que resultou em melhores resultados foi o polinomial. Todavia, o grau de polinômio a ser utilizado possui grande influência tanto no percentual de acertos quanto na velocidade de processamento. No algoritmo do SVM, aplicada na fase S e buscando uma predição de nível e severidade simultânea, na frequência de 5,0 Hz, o percentual de acertos utilizando polinômio de 5°, 20° e 50° grau são, respectivamente, 30% (tempo de execução: 30 ms), 55% (30 ms) e 58% (18,5 s). Já para a frequência de 7,5 Hz, o percentual de acertos utilizando polinômio de 5°, 20° e 50° grau são, respectivamente, 30% (30 ms), 60% (30 ms) e 74% (300 ms). Por fim, para a frequência de 10,0 Hz, o percentual de acertos utilizando polinômio de 5°, 20° e 50° grau são, respectivamente, 20% (30 ms), 69% (30 ms) e 91% (125 ms). Isso permite concluir que o

aumento do grau do polinômio é mais relevante quando o motor é alimentado com uma maior frequência de alimentação.

Ainda a respeito dos algoritmos de SVM e OvO, existe uma relação entre o grau de polinômio a ser utilizado com a escolha da análise a ser feita. Caso seja utilizado um polinômio de grau baixo, recomenda-se fazer a classificação de tipo e severidade de desbalanceamento separadamente, colocando as duas análises em cascata. Isso gera um percentual de acerto maior em relação à predição de tipo e severidade simultânea.

As análises mostram que existem poucas variações na predição do desbalanceamento quando é feita apenas a alternância das fases, independente da rotação do motor e do algoritmo de aprendizado de máquina a ser utilizado. Contudo, ao combinar os resultados obtidos sobre cada uma das fases, é possível fazer um algoritmo que apresenta resultados melhores.

Outra análise feita foi a remoção, individual, de um múltiplo de frequência para a realização do treinamento dos algoritmos de aprendizado. Foi detectada que a componente fundamental é a mais importante para o aprendizado, uma vez que a remoção dessa componente no treinamento, independente do algoritmo e frequência de rotação, faz com que o percentual de acerto tenha uma queda considerável.

Através de uma análise similar, também foi registrado que a componente de 1,5 vezes a componente fundamental possui certa relevância no treinamento e que a sua remoção no aprendizado piora a predição. A remoção exclusiva dessa componente, não possui tanto impacto na predição quanto a remoção da componente fundamental, todavia, se ambas as componentes forem removidas no aprendizado de máquina, fica evidente uma diminuição no percentual de acerto.

Apesar da maior importância da componente fundamental e da componente que é 1,5 vezes a fundamental, todas as outras componentes utilizadas apresentam certas informações que ajudam na predição correta do desbalanceamento. Isto é provado quando se percebe que a predição com apenas as duas componentes mais importantes é pior quando comparado na situação em que todas as componentes são utilizadas.

Entre os três métodos de aprendizado de máquina utilizados, os melhores resultados foram obtidos com o DTC. Isto se deve a dois fatos: 1) o percentual de acerto deste método é maior do que os demais; e 2) o tempo de processamento para este algoritmo é menor do que os demais. Os resultados obtidos entre os dois outros algoritmos são similares.

Os resultados obtidos atestam a validade da metodologia proposta. A comparação das estratégias de ML, bem como os cuidados necessários quanto ao processamento de sinais, também são contribuições deste trabalho para a engenharia de manutenção.

Futuramente, podem ser desenvolvidos estudos que comparam os sinais de corrente com as grandezas mecânicas que também foram coletados durante a aquisição dos sinais. Esse estudo é interessante pois pode mostrar uma relação de equivalência ou complementaridade para a de predição de desbalanceamento em máquinas rotativas.

## **Bibliografia**

[1] H. Toliyat et. al, Electric machines: modelling, condition monitoring, and fault diagnosis, 2 ed., CRC Press (2017) ISBN 9781138073975.

[2] W. T. Thomson and I. Culbert, Current signature analysis for condition monitoring of cage induction motors: industrial application and case histories. Series: IEEE Press Series on Power Engineering, 1 ed., Wiley-IEEE Press (2017) ISBN 9781119029595.

[3] W. R. Finley et. al, Motor vibration problem: How to diagnose and correct vibration errors. IEEE Industrial Application Magazine, 21(6) (2015) 14–28. doi: 10.1109/MIAS.2014.2345811.

- [4] M. Rigoni et. al, Non-invasive monitoring system of synchronous generator using external field. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications* 16 (2017) 70-89. doi: 10.1590/2179-10742017v16i1872.
- [5] J. Hey et. al, Online monitoring of electromagnetic losses in an electric motor indirectly through temperature measurement. *IEEE Transactions on Energy Conversion*, 31(4) (2016) 1347-1355. doi: 10.1109/TEC.2016.2562029.
- [6] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 1940-1: Mechanical vibration — Balance quality requirements for rotors in a constant (rigid) state —. 2 ed. Geneva: Iso, 2003.
- [7] CHAPMAN, S. J. *Fundamentos de Máquinas Eléctricas*. 5. ed. Porto Alegre: AMGH, 2013.
- [8] O. Matsushita et al. *Vibrations of rotating machinery*, 1 ed., Springer (2017) ISBN 9784431554561.
- [9] M. Galar et. Al. Enhancing difficult classes in one-vs-one classifier fusion strategy using restricted equivalence functions. 17th International Conference on Information Fusion (FUSION) (2014). ISBN: 978-8-4901-2355-3.
- [10] L. Rockah, O. Maimon, *Data mining with decision trees theory and applications*, Machine Perception Artificial Intelligence, vol. 69, 2014.
- [11] P. Bohra and H. Palivela, Understanding and formulation of various kernel techniques for support vector machines. *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (2015). doi: 10.1109/ICCIC.2015.7435804.
- [12] National Instruments. Labview, 2020 [Online]. Disponível em: <https://www.ni.com/pt-br/shop/labview.html>
- [13] Mathworks. Matlab, 2020. [Online]. Disponível em: <https://www.mathworks.com/products/matlab.html>
- [14] Scikit-learn, 2020. [Online]. Disponível em: <https://scikit-learn.org/>
- [15] Numpy, 2020. [Online] Disponível em: <https://numpy.org/>
- [16] Scipy, 2020. [Online] Disponível em: <https://www.scipy.org/>
- [17] Pandas, 2020. [Online] Disponível em: <https://pandas.pydata.org/>
- [18] SAMMUT, Claude; WEBB, Geoffrey I.. *Encyclopedia of Machine Learning*. Nova York: Springer, 2011.

## APÊNDICE B – Comprovante de apresentação

O presente estudo foi defendido no IX ENAM através de uma apresentação online. Foi feita uma gravação da apresentação e mesma foi disponibilizada pelos organizadores do evento no endereço eletrônico vinculado a este apêndice. A defesa do trabalho apresentado inicia em 01:06:54 e termina em 01:22:39.

**CIGRE. IX ENAM - Encontro Nacional de Máquinas Rotativas (17.01.2020) - Parte 1.** Disponível em: <https://www.youtube.com/watch?v=fJEUwVEnZjk>. Acesso em 04 jan. 2021.

## APÊNDICE C – Slides apresentados no congresso



**Identificação do tipo e da severidade do desbalanceamento de rotores por meio da análise da assinatura de corrente com aprendizado de máquina multiclases**

Rafael N. Tominaga      Fernando S. Pacheco  
Cesar A. Penz              Sergio L. Avila

Instituto Federal de Santa Catarina (IFSC)

*Florianópolis, 30 de março a 1º de abril de 2020*





### Sumário

- Introdução
- Fundamentação teórica
- Bancada de desbalanceamento
- Aquisição de sinais
- Processamento de sinais
- Utilização de Aprendizado de Máquina
- Conclusões





2

## Introdução – Justificativa



- Problema de desbalanceamento em rotores
  - Reduz vida útil da máquina rotativa
  - Reduz a eficiência do processo produtivo
- Método de prevenção
  - Monitoramento das condições operativas da máquina
- Variáveis possíveis para o monitoramento
  - Vibração
  - Temperatura
  - Campo eletromagnético
  - Corrente
    - Análise de assinatura de corrente do motor (MCSA)

3

## Introdução – Definição do problema

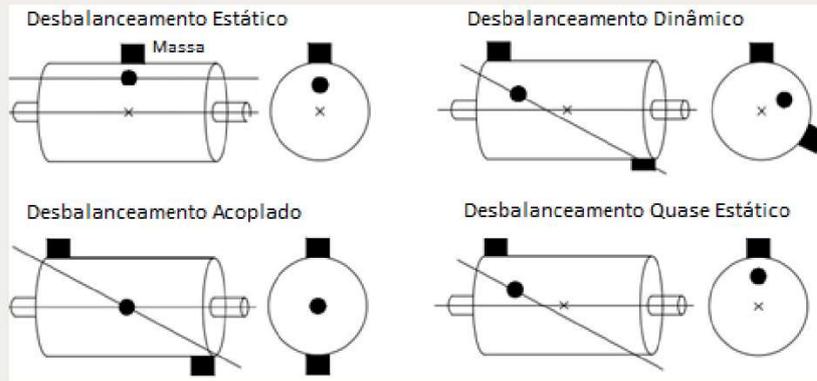


- Obter um sistema preditivo a partir do MCSA
  - Utilizando algoritmos de aprendizado de máquina (ML)
    - Árvore de decisão (DTC)
    - Máquina de vetores de suporte (SVM)
    - *One-vs-one* (OvO)



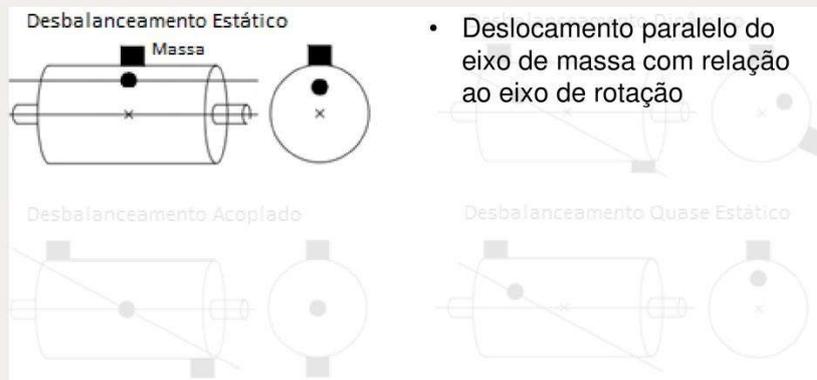
4

## Fundamentação teórica – Desbalanceamentos



5

## Fundamentação teórica – Desbalanceamentos



- Deslocamento paralelo do eixo de massa com relação ao eixo de rotação

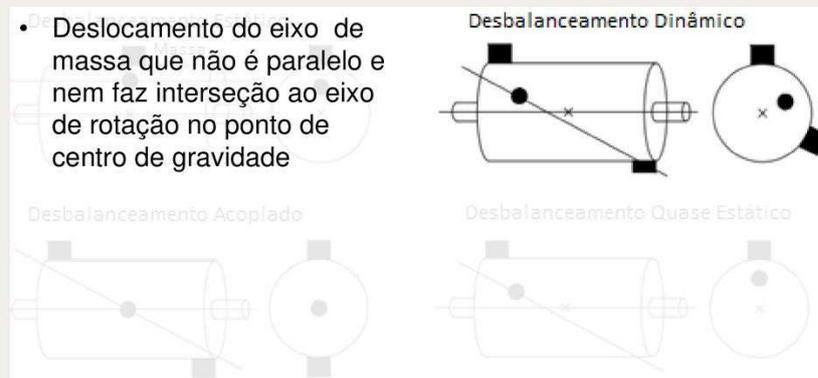
6

## Fundamentação teórica – Desbalanceamentos



7

## Fundamentação teórica – Desbalanceamentos



8

## Fundamentação teórica – Desbalanceamentos

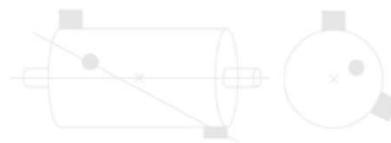


- Deslocamento do eixo de massa que faz interseção ao eixo de rotação em um ponto diferente do centro de gravidade

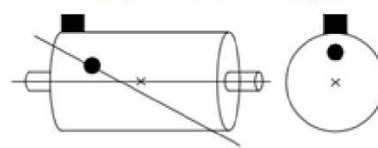
Desbalanceamento Acoplado



Desbalanceamento Dinâmico



Desbalanceamento Quase Estático

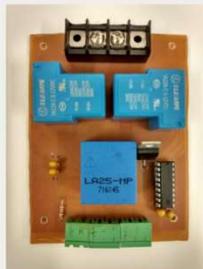


9

## Fundamentação teórica – MCSA



- Principais características:
  - Não há requisitos para ajustes mecânicos complexos
  - Resistente a interferência de campos magnéticos externos
  - A interpretação dos resultados não é evidente



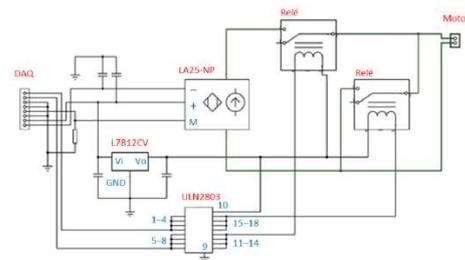
Modelo: LA25-NP

Fabricante: LEM

Erro máximo:  $\pm 0,5\%$  (25 °C)

Tempo resposta:  $< 1 \mu s$  (90% In)

Medição: 5 a 25 A (ajustável)



10

## Fundamentação teórica – Árvore de decisão



- Principais características:
  - Visa buscar uma árvore com o menor número de decisões possíveis
  - Fornece um diagrama com fácil compreensão
  - Consegue trabalhar tanto com números quanto com palavras

Clima	Temper.	Umidade	Vento	Agendar
Ensolar.	Quente	Alta	Falso	<b>Não</b>
Ensolar.	Quente	Alta	Verdadei.	<b>Não</b>
Nublado	Quente	Alta	Falso	<b>Sim</b>
Chuvoso	Agradável	Alta	Falso	<b>Sim</b>
Chuvoso	Frio	Normal	Falso	<b>Sim</b>
Chuvoso	Frio	Normal	Verdadei.	<b>Não</b>

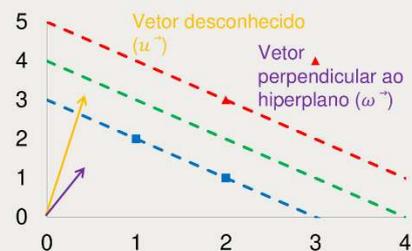
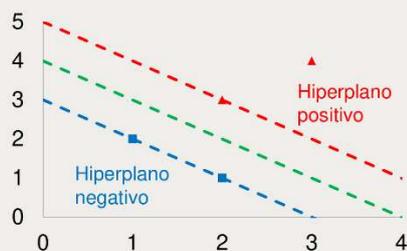


11

## Fundamentação teórica – Máquina de vetores de suporte

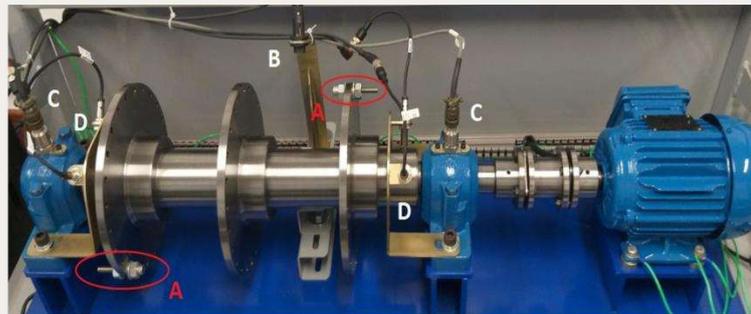


- Principais características:
  - Utiliza-se de conceitos de álgebra linear
  - Visa na obtenção do hiperplano através de um algoritmo de otimização
  - O hiperplano separa todas as classes da melhor forma possível



12

### Aquisição de sinais – Visão geral

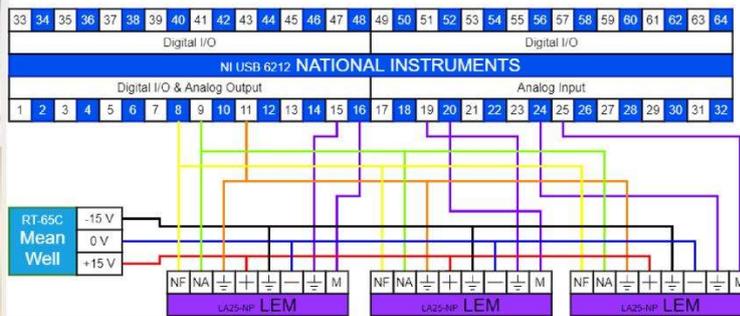


- A – parafusos para desbalanceamento
- B – sensor óptico
- C – acelerômetros
- D – sensores indutivos

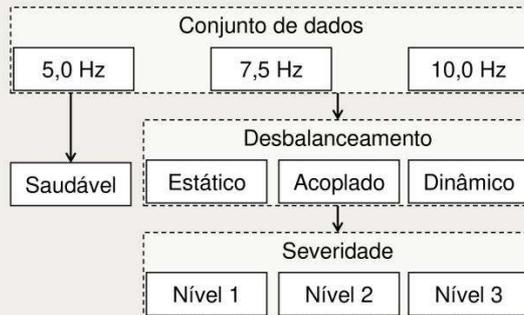
### Aquisição de sinais – Visão geral



Modelo: USB-6212      Fabricante: NI      Resolução: 16 bits  
 Amostragem: 400 kS/s      Medição:  $\pm 0,2 \text{ V}$  a  $\pm 10 \text{ V}$



## Aquisição de sinais – Visão geral

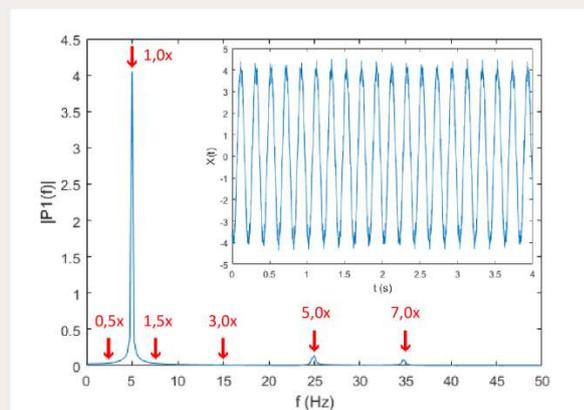


15

## Processamento dos sinais – Visão geral

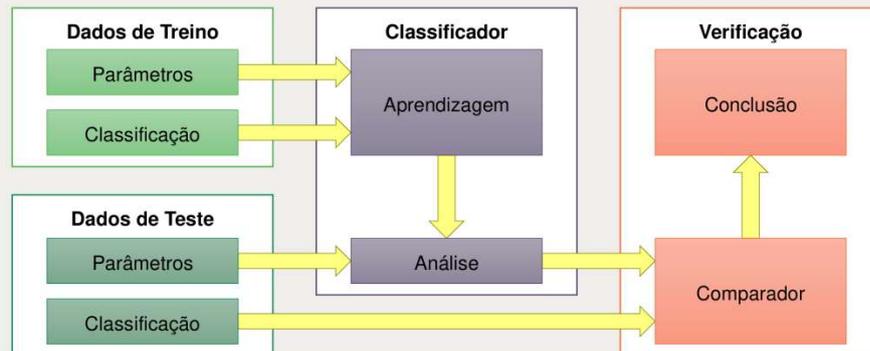


- Sinal no domínio do tempo
  - Aplicação da FFT
  - Seleção de componentes



16

## Utilização de Aprendizado de Máquina – Visão geral

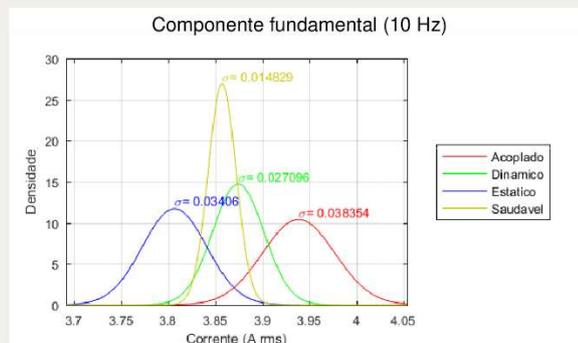


17

## Conclusões – A respeito do conjunto de dados



- Assim como o desbalanceamento dinâmico pode ser decomposto, em termos de massa de desbalanceamento, na somatória de um desbalanceamento estático com um desbalanceamento acoplado, o efeito da variação de corrente do desbalanceamento dinâmico também pode ser dividido de forma similar

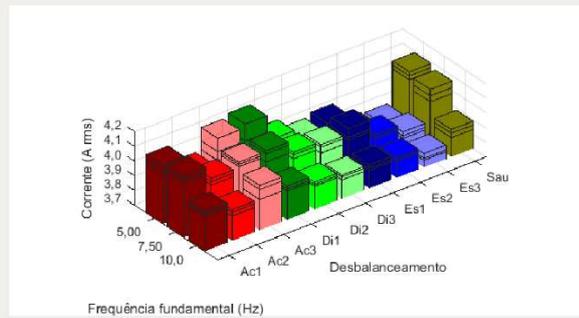


18

## Conclusões – A respeito do conjunto de dados



- Desbalanceamento estático tende a reduzir a amplitude da componente fundamental da corrente e essa redução é maior quanto maior for a severidade

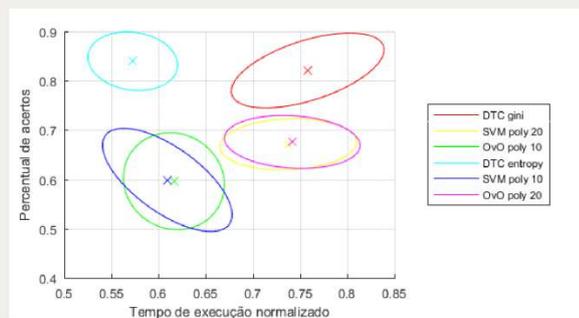


19

## Conclusões – A respeito dos algoritmos de ML



- Dentre todos os algoritmos de ML testados, o DTC apresentou o melhor desempenho em termos de precisão e tempo de execução



20

## Conclusões – A respeito da matriz de confusão



- Os estudos apontaram que existem quatro tipos de classificação de erros que podem ser identificados em uma matriz de confusão de detecção de desbalanceamento de máquinas elétricas rotativas com base na leitura da corrente

		Diagnóstico predito										Legenda	
		Ac1	Ac2	Ac3	Di1	Di2	Di3	Es1	Es2	Es3	Sau		
Diagnóstico verdadeiro	Ac1	118		54	22							82	<div style="display: flex; flex-direction: column; gap: 5px;"> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #90EE90; border: 1px solid black;"></span> Acerto</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #FF6347; border: 1px solid black;"></span> Erro aleatório</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #FFFF00; border: 1px solid black;"></span> Erro de severidade</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #FFB6C1; border: 1px solid black;"></span> Erro de composição</div> <div><span style="display: inline-block; width: 15px; height: 15px; background-color: #FF0000; border: 1px solid black;"></span> Erro de detecção</div> </div>
	Ac2		221	4	1	30	7	16					
	Ac3	71	1	127	90	1		2				44	
	Di1	36	1	102	118							37	
	Di2		48	3		124	25	109					
	Di3		7			53	233	40	3				
	Es1		24			87	22	143					
	Es2						3		188	70			
	Es3						2		87	196			
	Sau	111		47	28							162	

21



**ANEXOS**

## ANEXO A – Folder de chamada do congresso




**IX ENAM**  
**Atenção! Nova data**

O EVENTO SERÁ ONLINE  
 Encontro Nacional de Máquinas Rotativas  
 De 16 a 18 de Novembro de 2020

[Resumos aprovados](#)
[Template IT](#)
[Template PPT](#)

O ENAM, na sua nona edição, é um evento consolidado no âmbito do sistema interligado nacional e na incorporação de novas tecnologias de geradores e motores integrando as experiências de fabricantes, empresas geradoras, universidades, centros de pesquisa e empresas de consultoria. A primeira edição ocorreu em 2002 (coordenada pela CHESF), seguida de outras e coordenadas pela Eletronorte, Itaipu, Copel, Eletronuclear, Alstom, FURNAS, CEPEL e agora ELETROSUL.



**Data do evento: 16 a 18 de novembro de 2020**

O evento tem como alvo Agentes dos setores elétrico e acadêmico nas áreas de operação e manutenção de grandes máquinas rotativas, engenharia de projetos, modernização e monitoramento, fabricantes, empresas geradoras, universidades, centros de pesquisa e empresas de consultoria.

### TEMAS PARA TRABALHOS

A apresentação dos trabalhos terá como enfoque os seguintes assuntos: pesquisa, desenvolvimento, projeto, fabricação, inspeção convencional e robótica, operação convencional e intermitente e/ou como compensador síncrono, análise de defeitos: manutenção, especificação, modernização ou substituição, oscilações torcionais, tecnologia de sensores, experiência com sistemas de monitoramento com aplicação de algoritmos de diagnóstico e prognóstico, sistemas de proteção aplicáveis às máquinas rotativas, técnicas de ensaio, aumento de eficiência, grandes motores e motores de alto rendimento, pequenas centrais hidrelétricas, máquinas para geração distribuída e energia renovável, usinas reversíveis, sistemas de controle de excitação e de velocidade, sistemas de proteção de máquinas rotativas, impacto dos códigos de rede sobre a performance e custos das máquinas e aspectos ambientais.

**EVENTO ONLINE**

**INSCRIÇÕES**

Inscrições gratuitas

Realização:



Patrocínio Ouro:



Patrocínio Prata:



Patrocínio Bronze:

