



INSTITUTO TECNOLÓGICO VALE



**Programa de Pós-Graduação em Instrumentação, Controle e  
Automação de Processos de Mineração (PROFICAM)  
Escola de Minas, Universidade Federal de Ouro Preto (UFOP)  
Associação Instituto Tecnológico Vale (ITV)**

**Dissertação**

**Detecção Automática de Defeitos em Rolos de Transportadores de Correia  
Utilizando Sensoriamento Ultrassônico**

**Daniel Rodrigues Ericeira**

**Ouro Preto  
Minas Gerais, Brasil  
2019**

**Daniel Rodrigues Ericeira**

**Detecção Automática de Defeitos em Rolos de Transportadores de Correia  
Utilizando Sensoriamento Ultrassônico**

Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração da Universidade Federal de Ouro Preto e do Instituto Tecnológico Vale, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Controle e Automação.

Orientador: Prof. Gustavo Pessin, D.Sc.

Coorientador: Andrea G. Campos Bianchi, D.Sc.

Ouro Preto

2019

## SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

E68d Ericeira, Daniel Rodrigues .  
Detecção automática de defeitos em rolos de transportadores de correia utilizando sensoriamento ultrassônico. [manuscrito] / Daniel Rodrigues Ericeira. - 2019.  
82 f.

Orientador: Prof. Dr. Gustavo Pessin.  
Coorientadora: Profa. Dra. Andrea G. Campos Bianchi.  
Dissertação (Mestrado Profissional). Universidade Federal de Ouro Preto.  
Programa de Mestrado Profissional em Instrumentação, Controle e Automação de Processos de Mineração. Programa de Instrumentação, Controle e Automação de Processos de Mineração.  
Área de Concentração: Engenharia de Controle e Automação de Processos Mineraiis.

1. Correias transportadoras. 2. Aprendizado do computador. 3. Ultrassom. I. Ericeira, Daniel Rodrigues . II. Bianchi, Andrea G. Campos. III. Pessin, Gustavo. IV. Universidade Federal de Ouro Preto. V. Título.

CDU 681.5:622.2

Mestrado Profissional em Instrumentação, Controle e Automação de Processos  
de Mineração - PROFICAM

Detecção Automática de Defeitos em Rolos de Transportadores de Correia  
Utilizando Sensoriamento Ultrassônico

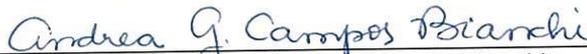
Daniel Rodrigues Ericeira

Dissertação defendida e aprovada em 26 de agosto de 2019 pela banca  
examinadora constituída pelos professores:



---

D.Sc. Gustavo Pessin  
Orientador – Instituto Tecnológico Vale Mineração (ITV)



---

D.Sc. Andrea Gomes Campos Bianchi  
Coorientadora – Universidade Federal de Ouro Preto (UFOP)



---

D.Sc. Bruno Nazário Coelho  
Membro interno – Universidade Federal de São João del-Rei (UFSJ)

---

D.Sc. Geraldo Pereira Rocha Filho  
Participação por videoconferência  
Membro externo – Universidade de Brasília (UNB)



---

MSc. Filipe Augusto Santos Rocha  
Membro interno – Instituto Tecnológico Vale Mineração (ITV)

## **Agradecimentos**

Agradeço a meus familiares, especialmente à minha mãe Vilma e à minha esposa Emmanuela. A elas dedico esta conquista.

Agradeço aos professores do curso, especialmente aos meus orientadores pelo apoio.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001; do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG); e da Vale SA.

## Resumo

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração da Escola de Minas/UFOP e do ITV como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

### DETECÇÃO AUTOMÁTICA DE DEFEITOS EM ROLOS DE TRANSPORTADORES DE CORREIA UTILIZANDO SENSORIAMENTO ULTRASSÔNICO

Daniel Rodrigues Ericeira

Novembro, 2019

Orientadores: Gustavo Pessin

Andrea G. Campos Bianchi

Transportadores de correia são os principais equipamentos que compõem a logística de um terminal portuário. As partes rolantes do transportador de correia podem falhar principalmente devido a rolos danificados, que podem causar avarias graves à planta, como rasgar a correia e causar incêndios. Atualmente, a proteção do transportador é feita por sensores que indicam uma anormalidade já ocorrida, ou com inspeção humana que utiliza experiência prática em busca de sinais visuais, sonoros e de temperatura excessiva que indiquem falhas iminentes. Com objetivo de auxiliar o atual sistema corretivo e de inspeção local com análise de dados para detectar falhas mecânicas iminentes, é proposto um modelo de classificação de defeitos em rolos. Foram realizadas gravações de ultrassom em rolos sem ruídos perceptíveis, classificados como “não-defeituosos” e em rolos que apresentavam ruídos característicos de falhas já perceptíveis, classificados como “defeituosos”. A base de dados montada foi utilizada para treinamento e teste de desempenho de algoritmos de aprendizado de máquina do tipo *Random Forest* e *Multilayer Perceptron*. Foram elaborados quatro tipos de experimentos para teste, dois usando dados no domínio do tempo e dois usando dados no domínio da frequência, com atributos estatísticos diferentes. Os resultados obtidos em testes de classificação mostraram que existe

um padrão característico na faixa de ultrassons que difere os rolos não-defeituosos de defeituosos, conforme pré-avaliados com métodos tradicionais de inspeção humana. No melhor caso, o experimento usando média móvel dos dados no domínio da frequência apresentou média de classificação correta de rolos de 83,68%, tendo o melhor resultado com taxa de acerto de 90%.

**Palavras-chave:** Transportadores de Correia, Aprendizado de Máquinas, Ultrassom.

**Macrotema:** Logística; **Linha de Pesquisa:** Instrumentação no Processamento de Minérios;

**Tema:** Transportadores de Correia.

## **Abstract**

Abstract of Dissertation presented to the Graduate Program on Instrumentation, Control and Automation of Mining Process of the Escola de Minas/UFOP and ITV as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

### AN ULTRASONIC SENSING SYSTEM FOR AUTOMATIC DEFECT DETECTION ON CONVEYOR BELTS IDLERS

Daniel Rodrigues Ericeira

November, 2019

Advisors: Gustavo Pessin

Andrea G. Campos Bianchi

Conveyor belts are the main class of machinery that compose the logistics of a port terminal. The rolling components of the conveyor belt may fail mainly due to damaged idlers, which may cause serious industrial breakdown, such as belt ripping or fires of major proportions. Nowadays, the equipment protection is done by a set of sensors that indicate an already existing abnormality, or by human inspection applying empirical experience in search of visual, sound or temperature signatures of imminent failure. Aiming to aid the current corrective system and local diagnostic inspection with data analysis to detect approaching mechanical failures, a model for classifying defects on the conveyor's idlers is proposed. Ultrasound recordings were conducted on idlers that didn't present any perceptible abnormalities, labeled as non-defective, and on idlers that displayed typical failure noise, labeled as defectives. The dataset collected was used for training and testing of Random Forest and Multilayer Perceptron machine learning algorithms. Four types of experiments were devised to test the methodology, two of them using time domain data and two of them using frequency domain data, with different statistical attributes. The results achieved in different classification experiments showed that there is a distinctive pattern on the ultrasound spectrum that differs non-defective from defective idlers,

as pre-evaluated by traditional methods of human inspection. In the best case, the experiment that used moving average on the frequency domain data presented an average of 83,68% of correctly classified idlers, having as best result a success rate of 90%.

**Keywords:** Conveyor Belts, Machine Learning, Ultrasound.

**Macrotheme:** Logistics; **Research Line:** Instrumentation in Mineral Processing; **Theme:** Conveyor Belts.

## Lista de Figuras

|   |    |
|---|----|
| Figura 1: Transportador de correia no pátio do Terminal de Ponta da Madeira. ....   | 15 |
| Figura 2: Partes componentes de um transportador de correia. Fonte: adaptado de<br>CONDITION MONITORING, 2015.....                              | 16 |
| Figura 3: Seção transversal do transportador, com os roletes de carga e retorno.<br>Fonte: adaptado de INGENIUM DESIGN, 2019. ....              | 17 |
| Figura 4: Corte de um rolete de carga típico, componente principal de um<br>transportador de correia. Fonte: adaptado de AKT KUPPER, 2019. .... | 18 |
| Figura 5: Vista aérea do TPM, com 92km de transportadores (Fonte: Google<br>Earth).....   | 19 |
| Figura 6: Detectora de rasgo de correia.....  | 20 |
| Figura 7: Chave de desalinhamento instalada ao lado da correia.....   | 20 |
| Figura 8: Metodologia desenvolvida no trabalho. ....  | 24 |
| Figura 9: Comparação entre comandos de entrada e dados de entrada [20]. ....  | 32 |
| Figura 10: Diagrama esquemático de uma RF com 3 árvores de decisão. Fonte: DOS<br>SANTOS et al., 2019. ....                                     | 34 |
| Figura 11: Diagrama esquemático de uma MLP com duas camadas. Fonte: DOS<br>SANTOS et al., 2019. ....  | 35 |
| Figura 12: Metodologia aplicada a uma base de dados. Fonte: adaptado de STACK<br>EXCHANGE, 2019.....  | 37 |
| Figura 13: Metodologia aplicada a uma base de dados. ....   | 38 |
| Figura 14: Sensor ultrassônico Ultraprobe 10000 (UE Systems Inc.).....  | 39 |
| Figura 15: Gravação de ultrassom de um rolo em um transportador do TPM.....   | 40 |
| Figura 16: Tela do usuário no Ultraprobe 10000. ....  | 41 |
| Figura 17: Transportador TR-315K-09, um dos transportadores utilizado para<br>captação de dados. ....   | 41 |

|   |    |
|---|----|
| Figura 18: Gráficos amplitude x tempo: à esquerda, rolos sem defeitos; à direita, rolos defeituosos. ....   | 43 |
| Figura 19: Gráficos magnitude x frequência: à esquerda, rolos sem defeitos; à direita, rolos defeituosos. ....  | 44 |
| Figura 20: Cada amostra de meio segundo de áudio gerou um vetor de média, mediana, desvio padrão e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso). ....   | 45 |
| Figura 21: Cada amostra de um segundo de áudio gerou um vetor de média, mediana, desvio padrão para cada quarto de segundo, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso). ....  | 46 |
| Figura 22: Cada FFT de um arquivo de áudio foi dividida em 40 partes, gerando para cada parte um vetor com valores de média, mediana, desvio padrão, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso). ....   | 47 |
| Figura 23: Para cada 5 segundos de áudio foi aplicada a FFT e esta foi dividida em 40 partes, gerando um vetor de 40 valores médios para cada FFT, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso). ....   | 48 |
| Figura 24: Comparação de gráficos de dispersão para os atributos do Experimento 1. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP). .... | 51 |
| Figura 25: Comparação de gráficos de dispersão para os atributos do Experimento 2. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP). .... | 53 |
| Figura 26: Comparação de gráficos de dispersão para os atributos do Experimento 3. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP). .... | 54 |

|  |    |
|--|----|
| Figura 27: Comparação de gráficos de dispersão para os atributos do Experimento 4. Cada par de <i>boxplots</i> apresenta a comparação entre a dispersão de rolos não-defeituosos (amarelo) e defeituosos (azul). ..... | 55 |
| Figura 28: Matriz de confusão com as medidas de desempenho. ....   | 56 |
| Figura 29: Matriz de confusão para o melhor caso do Experimento 1.....   | 58 |
| Figura 30: Resultados obtidos no Experimento 1.....  | 59 |
| Figura 31: Matriz de confusão para o melhor caso do Experimento 2.....   | 60 |
| Figura 32: Resultados obtidos no Experimento 2.....  | 60 |
| Figura 33: Matriz de confusão para o melhor caso do Experimento 3.....   | 61 |
| Figura 34: Resultados obtidos no Experimento 3.....  | 62 |
| Figura 35: Matriz de confusão para o melhor caso do Experimento 4.....   | 63 |
| Figura 36: Resultados obtidos no Experimento 4.....  | 63 |
| Figura 37: Taxas de acerto de todos os 4 experimentos. ....  | 64 |
| Figura 38: Análise de desempenho de todos os 4 experimentos. ....  | 65 |

## **Lista de Tabelas**

|   |    |
|---|----|
| Tabela 1: Total de correias do TPM.....   | 18 |
| Tabela 2: Plano de inspeção sensitiva mecânica de material de desgaste do TPM. .... | 21 |
| Tabela 3: Limites técnicos para troca de material de desgaste do TPM. ....          | 21 |

## **Lista de Siglas e Abreviaturas**

AE – *Acoustic Emission*

TMPM – Terminal Marítimo de Ponta da Madeira

CEMA – *Conveyor Belt Manufacturing Association*

US – Ultrassom

IA – Inteligência Artificial

ML – *Machine Learning*

RF – *Random Forest*

RNA – Redes Neurais Artificiais

MLP – *Multilayer Perceptron*

TDF – Transformada Discreta de Fourier

FFT – Transformada Rápida de Fourier

TF – Transformada de Fourier

## Sumário

|  |           |
|--|-----------|
| <b>1. Introdução .....</b>   | <b>15</b> |
| 1.1. Justificativa .....   | 19        |
| 1.2. Objetivos .....   | 23        |
| 1.3. Organização do Trabalho .....   | 24        |
| 1.4. Contribuições do trabalho .....   | 25        |
| 1.4.1. Contribuições técnicas .....  | 25        |
| 1.4.2. Contribuições científicas .....   | 26        |
| 1.4.3. Contribuições industriais .....   | 26        |
| 1.4.4. Contribuições de segurança .....  | 26        |
| <b>2. Referencial Teórico e Fundamentação Científica.....</b>                          | <b>27</b> |
| 2.1. Sensoriamento por vibração .....  | 27        |
| 2.2. Sensoriamento acústico por fibra óptica .....                                     | 28        |
| 2.3. Aplicação de sensores de som e ultrassom .....                                    | 29        |
| <b>3. Inteligência Artificial e Métodos de Aprendizado de Máquina.....</b>             | <b>31</b> |
| 3.1. Aprendizado de Máquina .....  | 31        |
| 3.1.1. Aprendizado supervisionado .....  | 32        |
| 3.1.2. Aprendizado não-supervisionado .....  | 32        |
| 3.1.3. Random Forests .....  | 32        |
| 3.1.4. Redes Neurais Artificiais .....   | 34        |
| <b>4. Materiais e Métodos .....</b>  | <b>37</b> |
| 4.1. Gravações de ultrassom de rolos realizadas em campo .....                         | 39        |
| 4.2. Pré-processamento dos arquivos de áudio obtidos e criação do banco de dados ..... | 42        |
| 4.3. Métodos estatísticos e classificação das amostras .....                           | 44        |
| 4.3.1. Experimento 1 – Vetor de atributos de meio segundo no domínio do tempo ....     | 44        |

|           |  |           |
|-----------|--|-----------|
| 4.3.2.    | Experimento 2 – Vetor de atributos de 1 segundo no domínio do tempo .....                                      | 45        |
| 4.3.3.    | Experimento 3 – FFT do arquivo de áudio completo dividida em 40 partes .....                                   | 46        |
| 4.3.4.    | Experimento 4 – média móvel da FFT de cada 5 segundos de áudio .....   | 47        |
| 4.4.      | Treinamento e teste de desempenho do classificador .....   | 49        |
| <b>5.</b> | <b>Resultados e Discussão .....</b>  | <b>50</b> |
| 5.1.      | Análise de dispersão dos atributos de entrada .....  | 50        |
| 5.2.      | Avaliação de desempenho com matriz de confusão .....   | 55        |
| 5.2.1.    | Sensibilidade.....   | 56        |
| 5.2.2.    | Especificidade.....  | 57        |
| 5.2.3.    | Precisão.....  | 57        |
| 5.2.4.    | Valor preditivo negativo .....   | 57        |
| 5.2.5.    | Acurácia.....  | 58        |
| 5.3.      | Investigação com Aprendizado de Máquina no domínio do tempo .....  | 58        |
| 5.3.1.    | Experimento 1 – Utilizando amostras de meio segundo com média, mediana e desvio padrão.....                    | 58        |
| 5.3.2.    | Experimento 2 – Utilizando amostras um segundo com vetor de 12 valores de média, mediana e desvio padrão ..... | 59        |
| 5.4.      | Investigação com Aprendizado de Máquina no domínio da frequência.....  | 61        |
| 5.4.1.    | Experimento 3 – Utilizando vetor de 3 valores para FFT .....   | 61        |
| 5.4.2.    | Experimento 4 – Utilizando média móvel das FFT de amostras de 5 segundos .....                                 | 62        |
| 5.5.      | Resultados gerais dos experimentos .....   | 63        |
| <b>6.</b> | <b>Conclusão e Trabalhos Futuros.....</b>  | <b>67</b> |
|           | <b>Referências Bibliográficas .....</b>  | <b>69</b> |
|           | <b>Anexo A: Código fonte para extração de dados no domínio do tempo .....</b>                                  | <b>73</b> |
|           | <b>Anexo B: Cabeçalho do dataset dos experimentos .....</b>  | <b>75</b> |
|           | <b>Anexo C: Dataset do Experimento 1 (20 primeiras linhas) .....</b>   | <b>76</b> |

|  |           |
|--|-----------|
| <b>Anexo D: Dataset do Experimento 2 (20 primeiras linhas) .....</b> | <b>77</b> |
| <b>Anexo E: Dataset do Experimento 3 (20 primeiras linhas) .....</b> | <b>79</b> |
| <b>Anexo F: Dataset do Experimento 4 (20 primeiras linhas).....</b>  | <b>80</b> |

## 1. Introdução

Transportadores de correia são os principais equipamentos que compõem a planta de um terminal portuário. São equipamentos que fazem o transporte contínuo do material a granel descarregado no terminal, até a fase de empilhamento nos pátios de estocagem para posterior recuperação e embarque nos carregadores. Toda a movimentação e logística do terminal portuário é dependente do correto funcionamento deste equipamento. A Figura 1 exhibe um transportador no pátio do Terminal Marítimo de Ponta da Madeira (TMPM), em São Luís.



Figura 1: Transportador de correia no pátio do Terminal de Ponta da Madeira.

O transportador de correia é composto por uma correia de borracha tensionada e apoiada sobre estruturas que comportam partes rolantes (roletes e tambores), movidas por sistemas de acionamento compostos comumente por motores elétricos ligados a acoplamentos e redutores de velocidade. O tambor ligado ao sistema motor é chamado de tambor de acionamento ou motriz, enquanto o tambor posicionado no extremo oposto é chamado de tambor de retorno. A

correia deve ser mantida tensionado pelo sistema esticador. O sistema esticador, normalmente atuado por gravidade (contrapeso) ou sistema hidráulico, é responsável por absorver as variações de operação da correia e mantê-la sempre aderente aos roletes de carga. A Figura 2 exibe os principais componentes de um transportador de correia típico.

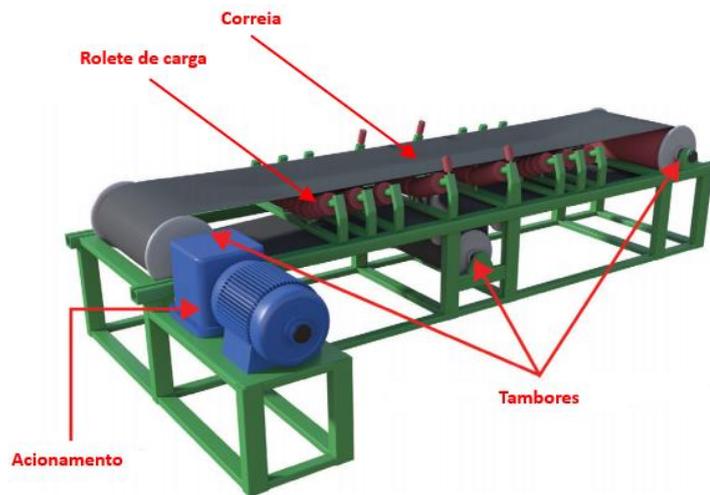


Figura 2: Partes componentes de um transportador de correia. Fonte: adaptado de CONDITION MONITORING, 2015.

A Figura 3 mostra um corte transversal de um transportador típico. Os roletes de carga normalmente são instalados em conjuntos de três, apoiados por uma estrutura chamada cavalete. Após a passagem pelo tambor da extremidade, a correia passa pela parte inferior da estrutura, apoiado por roletes maiores chamados de roletes de retorno.

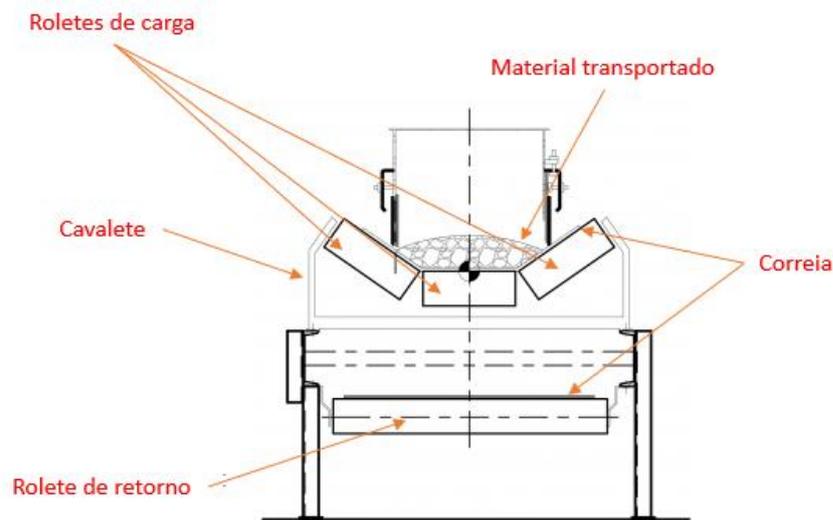


Figura 3: Seção transversal do transportador, com os roletes de carga e retorno. Fonte: adaptado de INGENIUM DESIGN, 2019.

As partes rolantes do transportador de correia podem falhar principalmente devido a: rolamentos danificados, abrasão por contato com a correia ou defeitos de fabricação. Em uma operação portuária, também existe a possibilidade de acidentes devido a impactos, material não-especificado, lotes defeituosos, entre outros. Os danos nas partes rolantes do transportador podem causar avarias graves à planta, como rasgar a correia e causar incêndios que podem se alastrar para os demais equipamentos.

A Figura 4 exibe uma visão de corte de um rolete típico, item que compõe a vasta maioria das partes rolantes de um transportador de correia. De acordo com Lodewijks (2003) e Wheeler e Ausling (2007), os maiores modos de falha de roletes são rolamentos danificados, que podem ter como causas principais: desgaste, contaminação por pó ou água, subdimensionamento ou acidentes operacionais. É estimado que 43% das falhas em rolamentos ocorrem por umidade ou outros contaminantes, efeitos de falhas na vedação (FLEXCO INSIGHT, 2011). As falhas neste componente podem fazer com que o rolete trave, deixando de exercer sua função e potencialmente causando danos à correia ou incêndios por sobreaquecimento (neste caso, o atrito pode aumentar a temperatura de um rolete travado em até 400°C (KRÓL e ZIMROZ, 2011). Roletes travados também aumentam o consumo de energia do transportador devido ao aumento da resistência ao movimento.

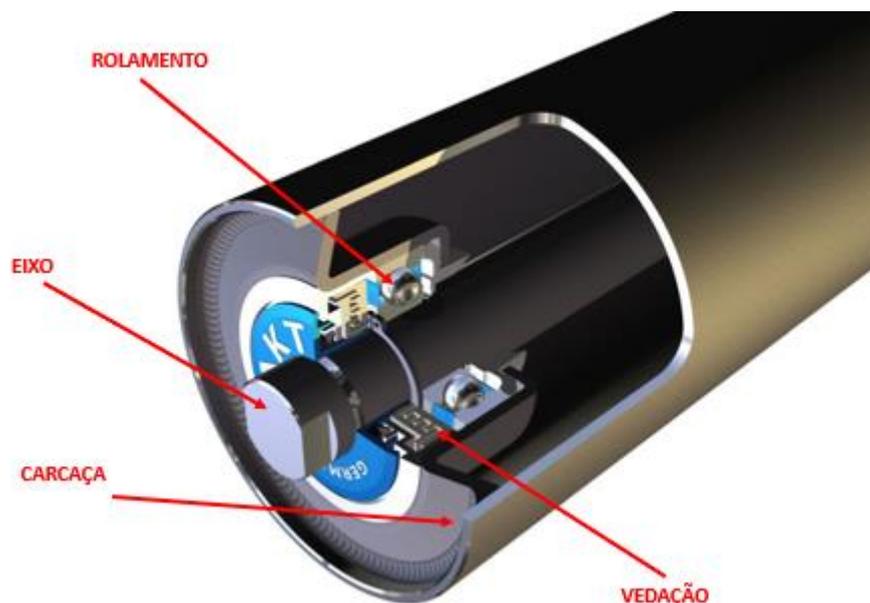


Figura 4: Corte de um rolete de carga típico, componente principal de um transportador de correia.  
 Fonte: adaptado de AKT KUPPER, 2019.

De acordo com dados do Plano Diretor de Manutenção do TMPM, ciclo 2014/2018, o terminal possui as quantidades descritas na Tabela 1.

Tabela 1: Total de correias do TMPM.

| Unidade | Quantidade de transportadores | Comprimento total de correias (km) | Tipo de correia (cabo de aço/lona) |
|---------|-------------------------------|------------------------------------|------------------------------------|
| TMPM    | 149                           | 92,054                             | 75,258 / 16,796                    |

Considerando a distância padrão dos projetos do TMPM entre cavaletes, pode-se estimar a quantidade total de rolos de carga em aproximadamente 368000 rolos de carga em 92,054 km de transportador (observar Figura 5).



Figura 5: Vista aérea do TMPM, com 92km de transportadores (Fonte: Google Earth).

As normas da Conveyor Equipment Manufacturers Association (CEMA) estabelecem os padrões que devem ser seguidos por fabricantes de componentes de transportadores. Esta norma define que os fabricantes de rolos devem definir um número para a expectativa de vida em porcentagem de falhas esperadas. Por exemplo, o número  $L_{10}$  define o tempo de operação em que, de um determinado lote de rolos fornecidos, espera-se que 10% tenha falhado. Para os fornecedores de rolos do TMPM, o número definido para  $L_{10}$  é de 30000 horas de operação.

### 1.1. Justificativa

Atualmente, a proteção do transportador é feita por sensores que indicam uma anormalidade já ocorrida (chaves de desalinhamento, detectores de correia rasgada, etc.) A chave detectora de rasgo (Figura 6) tem como objetivo parar o equipamento após a ocorrência de um rasgo na correia, que irá derramar material sobre uma bandeja localizada abaixo do

centro. A chave de desalinhamento (Figura 7) tem como objetivo parar o equipamento quando a correia estiver rodando excessivamente fora de sua posição central.



Figura 6: Detetora de rasgo de correia.

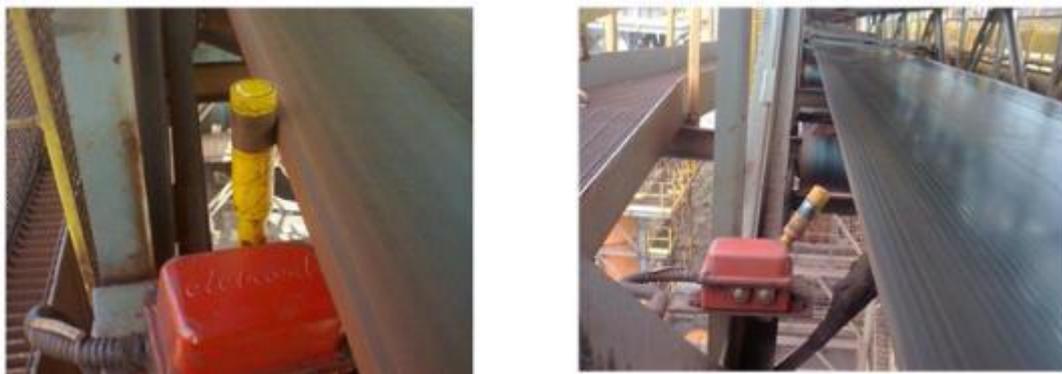


Figura 7: Chave de desalinhamento instalada ao lado da correia.

Existe também o processo de inspeção humana, com objetivo de percorrer toda a extensão do equipamento, em busca de sinais visuais, sonoros e de temperatura excessiva nas partes rolantes (através de pirômetro ou câmera termográfica) que indiquem falhas iminentes. Os inspetores utilizam de sua experiência para, através da característica do som, identificar e local e qual tipo de falha está ocorrendo: ruídos periódicos de batidas para danos de rolamentos no rolete, ruído de abrasão para rolos travados atritando com a correia, ruído característico de contato metal com metal, etc.

A Tabela 2, obtida do Plano Diretor de Manutenção 2014-2018 do TMPM, informa a rotina definida para inspeção dos rolos nos transportadores de correia. Está definido que, a cada 3 a 7 dias, os inspetores de itens de desgaste devem percorrer sua rota para inspeção sensível próximo ao equipamento.

Tabela 2: Plano de inspeção sensível mecânica de material de desgaste do TMPM.

| <b>Item a verificar</b> | <b>Mão de obra</b>   | <b>Criticidade</b> | <b>Frequência (dias)</b> |
|-------------------------|--|--------------------|--------------------------|
| Rolos                   | Mínimo de 02 executantes com formação técnica na área de mecânica e com experiência de execução. | 1                  | 3 a 7                    |

A Tabela 3, obtida do Plano Diretor de Manutenção 2014-2018 do TMPM, informa a quantidade de itens de desgaste prevista para troca a cada ano. Está definido que a previsão é de que sejam trocados 21641 rolos.

Tabela 3: Limites técnicos para troca de material de desgaste do TMPM.

| <b>Transportadores previstos (km)</b> | <b>Transportadores não previstos (km)</b> | <b>Rolos</b> | <b>Chapas de desgaste</b> | <b>Raspador</b> | <b>Revestimento de tambor</b> |
|---------------------------------------|---|--------------|---------------------------|-----------------|-------------------------------|
| 29,60                                 | 5,98                                      | 21641        | 2259                      | 5727            | 750                           |

Monitoramento de condições é o processo de coletar e avaliar parâmetros relevantes de condições de um determinado equipamento (vibração, som, temperatura, etc.), com objetivo de observar mudanças que sejam indicativas de anomalias indicativas de falhas iminentes ou falhas já ocorridas. Este processo é fundamental na detecção de falhas, que contribuem para redução de custos e diminuição de probabilidade de acidentes. Um processo de monitoramento inteligente deve ser capaz de aplicar técnicas de aprendizado de máquina e classificação de

padrões para tratar os dados captados pelos sensores do sistema de monitoramento, de forma a identificar antecipadamente possíveis falhas e diagnosticar dinamicamente o estado operacional do equipamento.

Técnicas de utilização de som para detectar falhas são aplicadas com sucesso na indústria (ver Seção 2). Estas técnicas com sensoriamento têm por objetivo emular a escuta especializada do inspetor humano, com a vantagem de ser um monitoramento contínuo com possibilidade de ser aliado com diagnóstico automático.

As técnicas de Emissão Acústica (do inglês *Acoustic Emission* - AE) são baseadas na captação direta de som, com resultados complementares às tradicionais medições de vibração e temperatura. A captação direta pode se dar através da instalação de sensores no equipamento a ser monitorado, com envio de dados para uma estação de trabalho fixa para tratamento das informações, ou através do deslocamento físico de pessoal de manutenção para gravação de som no local. Outra possibilidade é a utilização da faixa de ultrassons, com a perspectiva de diminuição de captação de ruído e posterior extração mais eficiente do sinal útil. A faixa de ultrassons tem se mostrado eficiente na previsão de falhas. Independentemente da forma de captação dos sinais sonoros, tem sido demonstrado em diversos estudos que é possível extrair informações destes sinais que ajudem no monitoramento de condições de equipamentos.

Com objetivo de auxiliar o atual sistema corretivo e de inspeção local, e utilizar análise de dados para detectar falhas mecânicas iminentes, o problema a ser tratado neste trabalho é a aplicação de monitoramento de condições com base em sinais sonoros para os transportadores. A literatura pesquisada define a aplicação do método para equipamentos com partes rolantes (motores elétricos, redutores, turbinas eólicas). O foco dos trabalhos pesquisados é o monitoramento de um componente específico do equipamento (rolamentos).

A proposta deste trabalho é a aplicação de uma metodologia de captação de ultrassom em componentes específicos que fazem parte dos transportadores de correia (roletes de carga), utilizando o sensor ultrassônico comercial *Ultraprobe 10000* para gravação em transportadores em funcionamento na planta do Terminal Marítimo de Ponta da Madeira (TMPM). As gravações foram feitas com o objetivo de captar sons de duas categorias gerais de roletes, chamadas no trabalho de “não-defeituosos” e “defeituosos”.

Roletes de carga recém instalados que não apresentavam nenhum ruído além do contato normal com a correia foram considerados como “não-defeituosos”. Para a categoria

“defeituoso”, foram buscados roletes que apresentavam aspecto deteriorado e que apresentavam características visuais e sonoras que costumam preceder falhas: aberturas perceptíveis e oxidação na vedação do rolete (indicativos de provável contaminação) acompanhados por ruídos perceptíveis de atrito interno do rolamento, rolete com extremidades oblongadas (que causam um ruído de batidas periódicas) e roletes com rotação irregular (rotação mais lenta e travamentos esporádicos).

O sensoriamento das emissões acústicas ultrassônicas ao longo do transportador, aliado a técnicas de inteligência artificial com reconhecimento de padrões de sinais para classificação de anomalias pode ser uma combinação importante na criação de um sistema de monitoramento dinâmico para diagnóstico e prevenção de falhas em rolos nos transportadores de correia.

## 1.2. Objetivos

O objetivo geral do trabalho é propor um modelo de classificação de rolos em iminente processo de falha. Serão modeladas técnicas de aprendizado de máquina tendo como entrada dados de Ultrassom (US). Os dados de ultrassom classificados devem permitir identificar sinais que sejam indicativos de falhas iminentes em partes rolantes dos transportadores de correia.

Estas falhas podem envolver defeitos nos rolamentos dos roletes, atrito entre partes metálicas e demais falhas de roletes que emitam sons característicos que, atualmente, são captados por processo de inspeção física local.

A Figura 8 apresenta um esquema da metodologia desenvolvida neste trabalho. Para concretização do objetivo geral, alguns objetivos específicos devem ser alcançados durante o desenvolvimento deste trabalho:

- *Realizar a coleta em campo de gravações de rolos utilizando o sensor ultrassônico Ultraprobe 10000.* As gravações foram realizadas com aproximação do sensor aos rolamentos (detalhado na Seção 4.1). Para a formação do banco de dados, foram feitas gravações em rolos sem ruídos perceptíveis e em rolos que apresentavam ruídos característicos de falhas já perceptíveis;
- *Propor, modelar e aplicar técnicas de pré-processamento de sinais necessárias para tratamento dos dados captados.* As gravações foram divididas em trechos,

sendo que cada trecho virou uma amostra de rolo “não-defeituoso” ou rolo “defeituoso”. Destas amostras foram calculados indicadores estatísticos para formação da base de dados;

- *Modelar e investigar técnicas de aprendizado de máquina e classificação que identifiquem padrões nos sinais que sejam indicativos de anormalidades no funcionamento do equipamento.* A base de dados foi utilizada para treinamento e teste de desempenho de algoritmos de aprendizado de máquina;

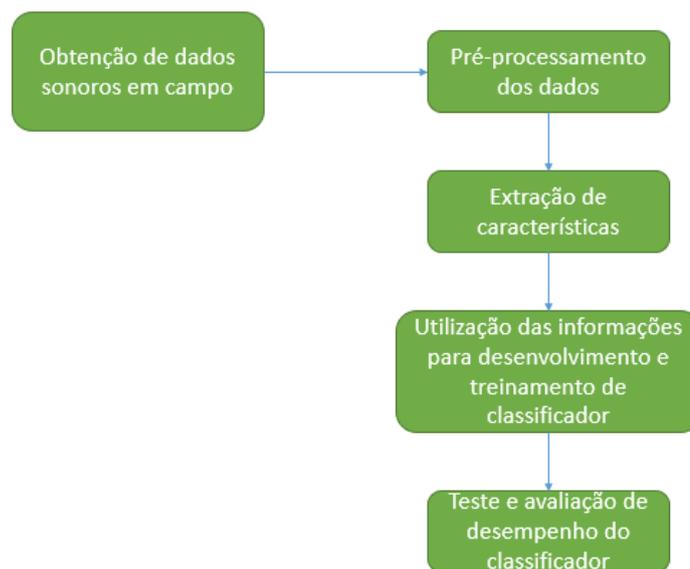


Figura 8: Metodologia desenvolvida no trabalho.

### 1.3. Organização do Trabalho

O trabalho descrito está organizado em mais seis capítulos. O Capítulo 2 faz um resumo da fundamentação teórica e dos trabalhos relacionados que serviram de base e respaldam esta tese. Apresenta conceitos sobre monitoramento inteligente e detecção de falhas em equipamentos utilizando, como sensoriamento para diagnóstico, dados de vibração, emissão acústica e técnicas de ultrassom.

O Capítulo 3 introduz conceitos de inteligência artificial e descreve as técnicas utilizadas de aprendizado de máquina.

O Capítulo 4 descreve a metodologia proposta, apresentada nas etapas: realização de gravações de ultrassom de rolos em campo, pré-processamento dos arquivos de áudio, separação dos dados em amostras e extração das características estatísticas desejadas e classificação das amostras, treinamento de um algoritmo de aprendizado de máquina com a base de dados, teste e avaliação de desempenho do classificador com uma nova base de dados para classificação final em rolo “não-defeituoso” ou rolo “defeituoso”.

Os Capítulos 4 e 5 apresentam e discutem os resultados obtidos com a metodologia proposta.

O Capítulo 6 apresenta a conclusão sobre o trabalho, mostrando a eficiência dos métodos utilizados, a relevância dos resultados obtidos e perspectivas de evolução para trabalhos futuros.

#### **1.4. Contribuições do trabalho**

O trabalho desenvolvido apresentou contribuições que podem ser classificadas das seguintes formas: técnicas, científicas, industriais e segurança.

##### **1.4.1. Contribuições técnicas**

Além da contribuição apresentada no desenvolvimento do trabalho, outras três contribuições técnicas foram construídas no desenvolvimento desta dissertação:

- 1) Programa de computador em Matlab, apresentado no Anexo 01, para tratamento e extração de dados dos arquivos de som;
- 2) *Dataset* (anexo físico) de sinais sonoros referentes a rolos com defeitos e sem defeitos para uso por demais pesquisadores que se interessem pelo tema;
- 3) *Dataset* (anexo físico) de dados tratados e classificados no domínio de tempo e domínio de frequência.

#### **1.4.2. Contribuições científicas**

A contribuição científica do trabalho é o desenvolvimento da metodologia para classificação de rolos. O trabalho contribui para demonstrar quais formas de pré-processamento e seleção de atributos são mais eficientes na captação de diferenças de padrões entre rolos. O trabalho mostra o potencial da aplicação de algoritmos de aprendizado de máquina para captação de defeitos em rolos sem necessidade de intervenção humana.

#### **1.4.3. Contribuições industriais**

A aplicação da metodologia em um processo de monitoramento de condições de transportadores de correia em uma planta industrial traria os seguintes ganhos:

- 1) Ganhos operacionais com o aumento da disponibilidade do equipamento e do OEE (*Overall Equipment Effectiveness*), devido à maior confiabilidade;
- 2) Redução de custos com danos em correias (rasgos) ou possíveis eventos de incêndios;
- 3) Monitoramento on-line das condições operacionais e diagnóstico das condições do equipamento.

No lugar da tradicional abordagem de “funcionar até falhar”, a análise dos dados capturados proporcionaria uma maneira de detectar falhas próximas, alimentando as equipes na planta com informações importantes.

#### **1.4.4. Contribuições de segurança**

A aplicação da metodologia em um processo automatizado com sensoriamento iria promover a redução de exposição de pessoas em processos de inspeção e manutenção, que são realizadas com proximidade ao equipamento em operação. A análise dos dados capturados utilizando a metodologia poderia proporcionar insumos para as equipes de manutenção trabalharem de forma mais direcionada, eliminando a necessidade das inspeções locais rotineiras.

## **2. Referencial Teórico e Fundamentação Científica**

Monitoramento inteligente e detecção de falhas são temas de análise em diversos trabalhos. Os tipos de dados mais utilizados são relativos a sensoriamento de vibração, emissão acústica, temperatura, técnicas de ultrassom e análise de esforços. Foram pesquisados trabalhos que utilizavam sensoriamento sonoro para monitoramento e diagnóstico de equipamentos com elementos rotativos.

Os trabalhos avaliados ou são trabalhos estruturantes, que propõem as bases teóricas para a utilização de monitoramento inteligente em equipamentos mecânicos de grande porte, como turbinas eólicas e viradores de vagões, ou são trabalhos que envolvem ensaios de bancada em partes específicas de equipamentos com partes rolantes, como os rolamentos de redutores ou motores elétricos. A contribuição deste trabalho foi aplicar a fundamentação teórica baseada nestes trabalhos para captação de defeitos em um equipamento com várias partes rolantes operando simultaneamente, como um transportador de correia, e utilizar sons gravados em equipamentos instalados e em operação em uma planta industrial, com todos os ruídos e interferências de outros equipamentos, para avaliar se o método tem viabilidade de aplicação em casos de indústrias operacionais.

### **2.1. Sensoriamento por vibração**

Um dos parâmetros mais citados na literatura de monitoramento de condições é o padrão dos sinais de vibração emitido pelo equipamento durante funcionamento.

O trabalho de Ali et al. (2015) propõe um modelo de extração de características de vibração emitidas por rolamentos. As características selecionadas foram usadas para treinamento de redes neurais artificiais para classificar tipos de defeitos em rolamentos. Os experimentos realizados indicam que os sinais captados na abordagem utilizada de acompanhar o funcionamento até a falha mostram indicações que podem categorizar defeitos. O trabalho mostra o potencial da aplicação de redes neurais para avaliação de degradação automática em rolamentos sem necessidade de intervenção humana.

Hoffman e der Merwe (2002) aplicam redes neurais para a tarefa de detecção e identificação de defeitos em rolamentos através da análise de espectro de vibração. O trabalho

cita que as características de espectro associadas a cada falha já são conhecidas, e que a utilização de redes neurais pode explorar ainda mais esse potencial. Diferentes formas de aprendizado supervisionado foram testadas para classificação e identificação de presença de defeitos, usando tanto o domínio do tempo quanto da frequência.

Diversos trabalhos focam na aplicação de monitoramento de condições feito de forma automática com inteligência artificial em turbinas eólicas. O trabalho de Li et al. (2009) traz uma coleção de trabalhos focados em extração de características do espectro de vibração para, através da utilização de redes neurais, diagnosticar componentes como redutores, rolamentos, rotores e lâminas dos geradores. Com abrangência para outros equipamentos mecânicos, Mjit et al. (2011) compara técnicas de análise de vibração para diagnosticar turbinas de correntes marítimas.

## **2.2. Sensoriamento acústico por fibra óptica**

A metodologia de sensoriamento de equipamentos mecânicos através de uma fibra óptica embarcada é tema de diversos artigos científicos, com bastante potencial em aplicações industriais. No trabalho de ZuDe et al. (2011) é proposto um modelo geral de diagnóstico e monitoramento inteligente de equipamentos baseado em sistemas embarcados e sensores de fibra óptica. Entre as vantagens do sensoriamento distribuído com fibra, destacam-se: capacidade de se aplicar em locais de acesso difícil e estreito, transmissão em longa distância e adaptação ao meio, resistência a influências eletromagnéticas e possibilidade de implementação de monitoramento e diagnóstico multi-parâmetros.

Martins et al. (2017) exploram as possíveis técnicas e aplicações de sensoriamento acústico distribuído, utilizando fibra óptica como elemento sensor. O foco do trabalho são os recentes avanços na tecnologia de fabricação das fibras, que criam novos desafios e possibilidades de aplicações industriais, militares e geofísicas.

Na página *web* da empresa Mining3 (2019) é descrita a pesquisa atual sendo desenvolvida na aplicação de fibra óptica para detecção de falhas em rolos ao longo de transportadores de correia. Um cabo de fibra é lançado ao longo do transportador, com um pulso de laser sendo emitido pela fibra. As perturbações do meio afetam o corpo do cabo e interferem na propagação do laser no interior da fibra. Este sinal, convertido em sinais elétricos, é utilizado

para fazer um perfil do transportador com pontos de falha. O processo está atualmente em teste em algumas plantas pelo mundo.

### **2.3. Aplicação de sensores de som e ultrassom**

No trabalho de Arredondo et al. (2016) é desenvolvida uma metodologia para detecção de falhas em motores elétricos através de sinais de som e vibração. Os tipos de falhas a serem observadas eram mecânicas (falhas de rolamento, desalinhamento de eixo e barras de rotor quebradas). O som do motor em operação foi captado por microfones. Os motores com falha apresentaram sons com alta intensidade em frequências características para cada tipo de falha que não eram apresentadas pelos motores sem falha, o que demonstra a aplicabilidade da detecção de falhas através do som.

O trabalho de Marquez et al. (2012) apresenta uma revisão geral de várias técnicas e métodos de processamento de sinais para monitoramento de condições de turbinas de vento utilizadas na geração de energia. Um dos sinais utilizados na literatura estudada para detecção de falhas em rolamentos, engrenagens, rotores e lâminas é a emissão acústica.

No trabalho de Kim et al. (2006), foi realizada uma comparação de monitoramento de condições de máquinas rotativas entre análise de vibração e análise de ultrassom. O trabalho destaca que a vantagem da utilização da faixa de ultrassons é a alta capacidade de detecção com respeito à relação sinal-ruído. Os ensaios realizados em laboratório mostraram uma comparação entre os sinais de rolamentos bons e com falhas internas, tanto no domínio do tempo quanto no domínio da frequência, com a utilização de diferentes métodos estatísticos. Os padrões apresentados mostraram que os rolamentos defeituosos apresentam características distintas na faixa de ultrassom em relação aos rolamentos sem falha, o que demonstra o potencial para a utilização de classificadores automáticos.

O trabalho de Tchakoua et al. (2013) apresenta uma revisão geral de métodos e técnicas utilizadas na indústria para monitoramento de condições em turbinas eólicas. Entre as técnicas pesquisadas, é citada a utilização de AE para detecção de falhas em elementos rolantes, com os principais parâmetros citados em literatura (*valor rms, kurtosis, crest fator, skewness*). Como vantagens para a utilização de emissões acústicas na faixa de ultrassons, são citadas a ampla faixa de frequências e o melhor desempenho em relação ao ruído ambiente. Como

desvantagens, estão o elevado custo de manter sensores próximos à fonte e o fato de que apenas certos tipos de falha apresentam características distintas em alta frequência.

Kedadouche, Thomas e Tahan (2014) aplicaram Empirical Mode Decomposition para decompor um sinal de emissão acústica no tempo, com objetivo de detectar falhas iminentes em rolamentos. Descritores estatísticos foram calculados para cada decomposição do sinal (*rms*, *crest factor*, *kurtosis* e *skewness*) e suas eficiências para detecção foram comparadas. Os autores descobriram que os sinais decompostos apresentaram maior sensibilidade às falhas de rolamento do que o sinal original não-decomposto. Os ensaios realizados em laboratório foram capazes de detectar uma falha de tamanho 40 micrômetros nos rolamentos, com proposta de ampliar o método para diversos tamanhos de danos internos.

Pela análise dos trabalhos citados nesta seção, a informação sonora captada de equipamentos mostra-se uma forma eficiente de percepção de anomalias e diagnóstico do estado de integridade de equipamentos. A forma como o som é captado pode variar em cada trabalho (fibra óptica, gravação direta por microfone ou gravação na faixa de ultrassom), mas fica evidenciado que os dados trazem informações relevantes que podem ser extraídas e utilizadas em um sistema antecipativo de detecção automática de anomalias operacionais e possíveis falhas, sendo a faixa de ultrassom a que apresenta maior capacidade antecipativa e menos suscetibilidade a ruídos. A contribuição deste trabalho, em comparação com os trabalhos pesquisados, é a utilização da fundamentação teórica aplicada com diferentes formas de extração de características e amostragem em equipamentos operacionais em uma planta industrial, em vez de ambientes controlados com defeitos inseridos nos equipamentos.

### **3. Inteligência Artificial e Métodos de Aprendizado de Máquina**

O desafio de resolver problemas cada vez mais complexos provocou o desenvolvimento de sistemas ditos “inteligentes” que modelam a inteligência biológica de seres vivos. De acordo com Engelbrecht (2007), estes sistemas, que formam o campo da Inteligência Artificial (IA), são compostos por algoritmos inteligentes que envolvem lógica, raciocínio dedutivo, sistemas especialistas e sistemas de aprendizado de máquina.

De acordo com conceitos de IA reunidos por Russell e Norvig (2009), as definições variam em duas dimensões principais: processos de pensamento e comportamento, medidos de duas formas: performance humana e racionalidade. Portanto, as definições de IA estão organizadas em quatro categorias: sistemas que pensam como humanos, sistemas que pensam de forma racional, sistemas que agem como humanos e sistemas que agem de forma racional. A abordagem humana é baseada em hipótese e experimentos (ciência empírica), enquanto a abordagem racional envolve matemática e engenharia.

#### **3.1. Aprendizado de Máquina**

Aprendizado de Máquina (do inglês *Machine Learning* - ML) é um ramo da IA que tem como principal característica o auto-aprendizado: aplicar modelagens estatísticas para detectar padrões e melhorar performance baseando-se em dados e informação empírica sem comandos diretos de programação. Mitchell (1997) definiu o aprendizado de um computador da seguinte forma: um programa aprende com a experiência E com relação a uma classe de tarefas T e medida de performance P se a performance nas tarefas T, medida por P, melhora com a experiência E.

Em aprendizado de máquina, os dados são divididos em conjunto de treino e conjunto de testes. O conjunto de treino é utilizado para desenvolver o modelo. Após o desenvolvimento do modelo, ele pode ser testado com um novo conjunto de dados (conjunto de testes). O modelo treinado e testado de forma satisfatória pode ser utilizado então para reconhecer padrões e realizar classificações, tomar decisões, etc.

Em Theobald (2017), conforme ilustrado na Figura 9, os modelos são construídos para agir com dados de entrada e não com comandos de programação diretos.

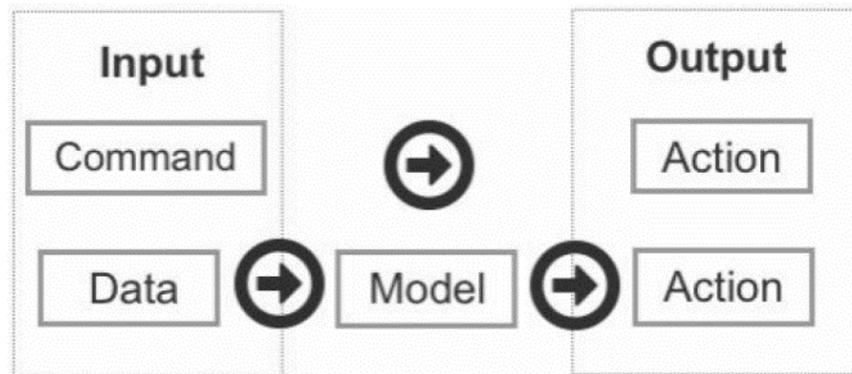


Figura 9: Comparação entre comandos de entrada e dados de entrada [20].

### 3.1.1. Aprendizado supervisionado

Neste trabalho foi utilizado um método de aprendizado supervisionado. Neste tipo de aprendizado, o modelo é alimentado com dados com vários atributos e o valor correto de saída para os dados. Como os dados já possuem sua correta saída, o conjunto de dados se diz já “classificado”. O algoritmo então decifra os padrões de atributos que existem no conjunto de dados e cria um modelo que pode reproduzir as mesmas regras com um novo conjunto de dados.

### 3.1.2. Aprendizado não-supervisionado

Nos casos de aprendizado não-supervisionado, nem todas as amostras são classificadas. Em vez disso, o algoritmo deve descobrir padrões escondidos e criar classificações. Algoritmos populares para estas aplicações são os métodos k-NN (*k Nearest Neighbors*), em que dados são agrupados com base em características similares. A vantagem nestes casos é a possibilidade de descoberta de padrões nos dados dos quais o usuário sequer possuía conhecimento.

### 3.1.3. Random Forests

O algoritmo Random Forests (RF) cria uma combinação de árvores de decisão, chamada de floresta. Estas árvores de decisão são combinadas para obter melhor desempenho. Uma árvore de decisão é um conjunto de “questões” ou tomadas de decisão que podem ter dois

caminhos a seguir: verdadeiro ou falso. À medida que as questões são respondidas, o algoritmo percorre um fluxograma (a árvore) rumo a uma previsão. Random Forests são conhecidas pela eficiência em evitar “overfitting”, facilidade de implementação, geração de informações relevantes além de classes, boas propriedades de generalização, eficiência em conjuntos numerosos de dados e boas estimativas de quais variáveis são mais importantes na tarefa de classificação (DOS SANTOS et al., 2019).

Durante a fase de treinamento, um modelo supervisionado aprende a relacionar os dados (características) com os resultados que queremos que o algoritmo aprenda a prever (alvo). As árvores de decisão aprendem a calcular as melhores perguntas a serem feitas utilizando as características que lhe foram alimentadas, de forma a realizar as melhores previsões possíveis. Os algoritmos RF apresentam uma combinação de decisões de diversas árvores, cada árvore da floresta considerando um conjunto aleatório de características e utilizando um conjunto aleatório de dados para treinamento, aumentando a diversidade com as melhores taxas de previsão de várias árvores de decisão montadas aleatoriamente (KOEHRSEN, 2019). A Figura 10 apresenta um diagrama esquemático de uma RF com um conjunto de três árvores de decisão distintas.

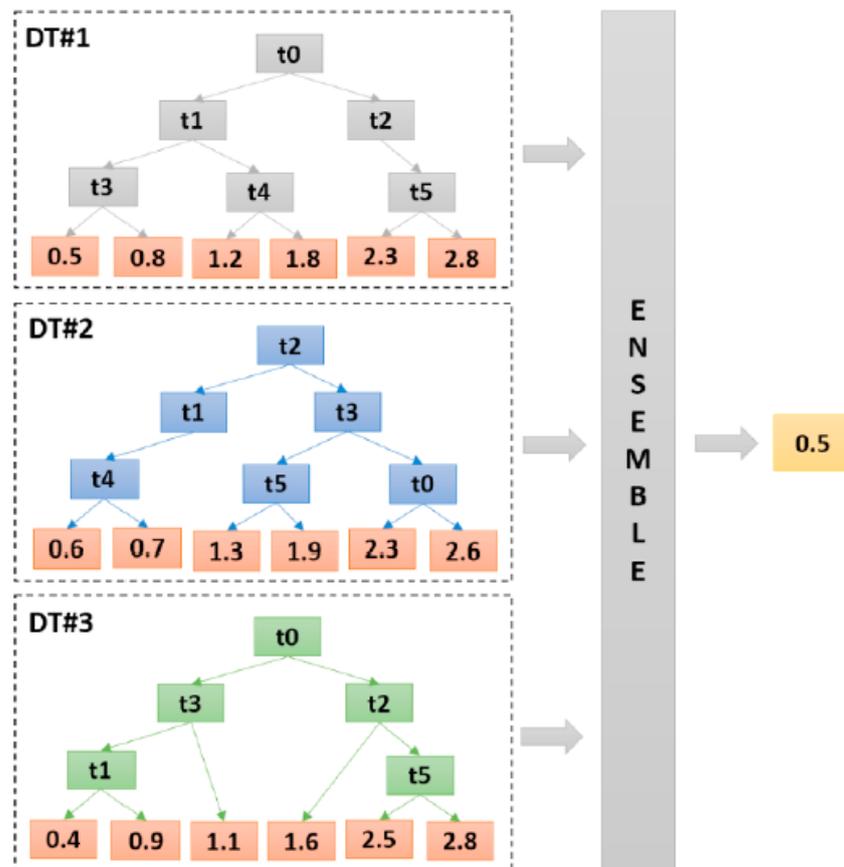


Figura 10: Diagrama esquemático de uma RF com 3 árvores de decisão. Fonte: DOS SANTOS et al., 2019.

### 3.1.4. Redes Neurais Artificiais

A forma mais simples e comum de Rede Neural Artificial (RNA) é a Multilayer Perceptron (MLP) (DOS SANTOS et al., 2019). É composta por elementos computacionais que são conectados em 3 camadas e possuem diferentes pesos: camada de entrada, camada escondida e camada de saída. As conexões dos elementos (chamados neurônios) são inspiradas no cérebro humano. As MLP possuem a habilidade de aprender padrões complexos e generalizar a informação aprendida (FERREIRA et al., 2017).

O processo de aprendizado da MLP é dividido em três partes: estimulação pela extração de exemplos, modificação de pesos em processos iterativos para minimizar o erro na saída e modificação nas respostas como resultado das mudanças que ocorreram anteriormente. A

Figura 11 exibe um esquemático de uma MLP com duas camadas escondidas de 6 neurônios cada.

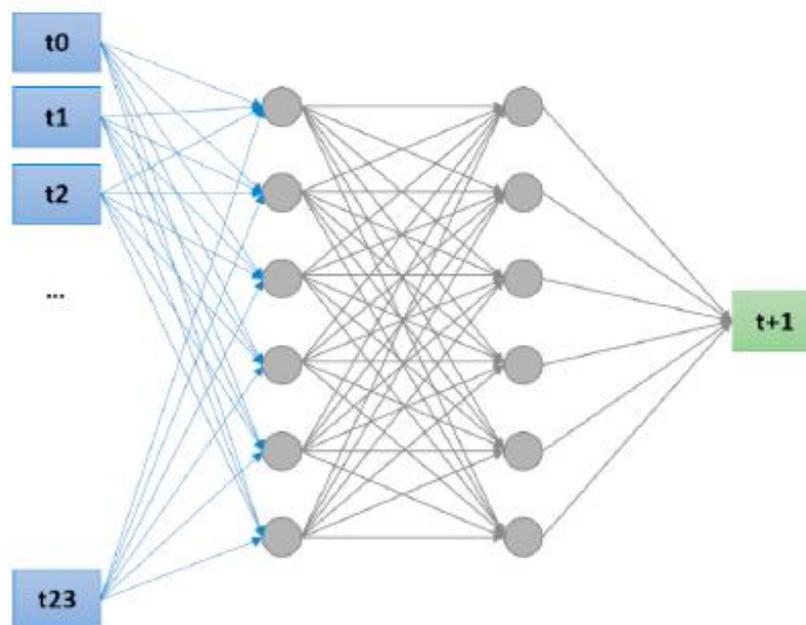


Figura 11: Diagrama esquemático de uma MLP com duas camadas. Fonte: DOS SANTOS et al., 2019.

Aprendizado de máquinas é um ramo da IA utilizado em tarefas de classificação, que são parte integral deste trabalho. Algoritmos de classificação foram utilizados para que o padrão observado nas gravações de ultrassom dos rolos não-defeituosos e defeituosos fossem aprendidos e, posteriormente, utilizados para classificar rolos desconhecidos.

Redes neurais, como MLP, são comumente utilizadas no aprendizado de padrões em dados complexos e generalização da informação aprendida, enquanto as RF também são bastante utilizadas com o mesmo objetivo em conjuntos numerosos de dados com boas propriedades de generalização e estimativa das variáveis mais significativas nas tarefas de classificação.

A escolha foi feita por estes dois métodos por efeitos de comparação: outras opções foram implementadas, porém os resultados obtidos não obtiveram melhores resultados em relação às RF e MLP. Com relação a árvores de decisão, testes foram feitos utilizando os algoritmos RepTree (WITTEN et al., 2016) e J48 (ROKACH e MAIMON, 2008). Em

comparação, as Random Forest apresentaram desempenho geral melhor. Redes bayesianas também foram testadas (GOLDSCHMIDT e PASSOS, 2005), porém não foram melhor investigadas por não apresentarem resultados melhores em relação às RF e MLP.

## 4. Materiais e Métodos

Nesta seção é descrita a metodologia desenvolvida pelo trabalho. As etapas seguidas são: realização de gravações de ultrassom de rolos em campo, pré-processamento dos arquivos de áudio com obtenção dos dados no domínio do tempo e da frequência através do algoritmo Transformada Rápida de Fourier (FFT), separação dos dados em amostras e extração das características estatísticas desejadas (média, mediana e desvio padrão), classificação das amostras, treinamento de um algoritmo de aprendizado de máquina (Random Forest e Multilayer Perceptron) com a base de dados, teste e avaliação de desempenho do classificador com uma nova base de dados para classificação final em rolos sem defeitos (rotulados como “não-defeituosos”) ou rolos com defeitos (rotulados como “defeituosos”).

Um sinal sonoro pode ser representado como uma forma de onda em função do tempo. Esta forma de onda representa a intensidade do som em cada instante de tempo. Qualquer forma de onda no tempo é uma soma de formas de onda senoidais de diferentes frequências. A Transformada de Fourier (TF) decompõe uma forma de onda no tempo em suas senoides constituintes. Na Figura 12, o sinal em cor vermelha representa uma forma de onda de amplitude variante no tempo, que pode ser decomposta em suas senoides constituintes que se somam para formar o sinal original. Estas senoides são as frequências fundamentais que formam o sinal original.

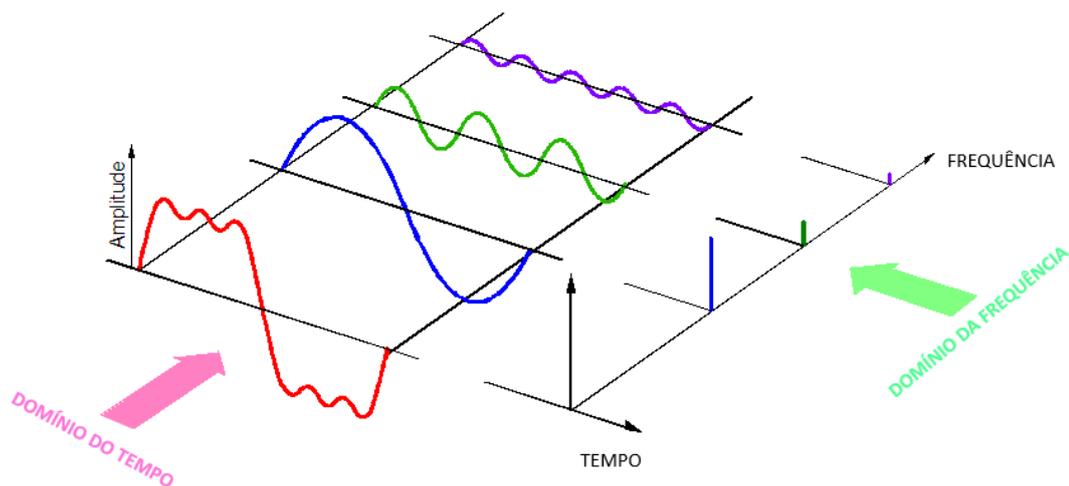


Figura 12: Metodologia aplicada a uma base de dados. Fonte: adaptado de STACK EXCHANGE, 2019.

A FFT é um algoritmo que aplica a Transformada de Fourier como uma função discreta. A computação numérica da TF requer dados de amostragem da função pois o computador pode apenas trabalhar com dados discretos. As Transformadas Discretas de Fourier (TDF) são úteis porque revelam periodicidades nos dados de entrada assim como as intensidades relativas dos componentes periódicos. A FFT é um algoritmo eficiente para realizar as TDF. O número de computações necessárias para realizar uma TDF é bastante reduzido pelo algoritmo FFT (o número de computações para  $N$  pontos é reduzida da ordem de  $N^2$  para  $N \log N$ ) (LATHI, 1998).

A Figura 13 exibe um diagrama com a metodologia aplicada a uma base de dados. Do banco de dados de gravações são extraídas as características desejadas para treinamento de um classificador. Este classificador treinado então é aplicado a uma nova gravação de um rolo desconhecido para realizar a classificação.

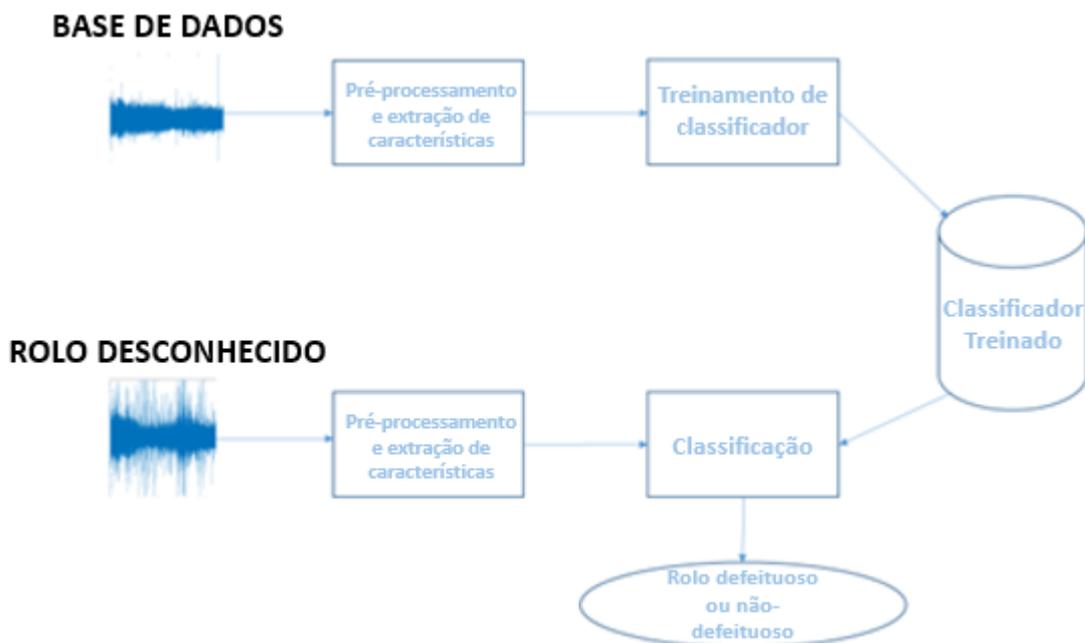


Figura 13: Metodologia aplicada a uma base de dados.

#### 4.1. Gravações de ultrassom de rolos realizadas em campo

A captação dos dados foi feita com um gravador ultrassônico, utilizado nas inspeções preditivas no TPM. Trata-se do sensor comercial *Ultraprobe 10000* (mostrado na Figura 14) da UE Systems Inc., que é utilizado no terminal para inspeção preditiva no monitoramento em redes elétricas. Este aparelho capta sons nas frequências de 20kHz a 100kHz, com software de gravação de dados para exportação em arquivos de extensão *.wav*.

O *Ultraprobe* utiliza uma técnica de processamento de sinais chamada “*Heterodyning*” para deslocamento dos ultrassons obtidos para frequências audíveis pelo ouvido humano (entre 20 e 20kHz). O objetivo deste processo é converter os sinais ultrassônicos em ruídos que podem ser ouvidos pelo operador humano no momento da inspeção do equipamento com auxílio de fones de ouvido. O operador experiente pode aprender os padrões do som convertido na detecção de rolos defeituosos.



Figura 14: Sensor ultrassônico Ultraprobe 10000 (UE Systems Inc.).

A gravação foi feita aproximando-se o sensor de rolos de transportadores (conforme orientação do fabricante) em funcionamento no TPM (conforme demonstrado na Figura 15), para posterior montagem de um banco de dados histórico a ser utilizado nas etapas seguintes.



Figura 15: Gravação de ultrassom de um rolo em um transportador do TMPM.

Um dos ajustes recomendados para realizar as gravações é o ajuste de sensibilidade. O objetivo deste ajuste é evitar que o som perca suas características e capte muitos ruídos externos caso fique elevado demais, ou que fique com ajuste baixo e não consiga captar o som completamente. A recomendação de ajuste pelo fabricante é a seguinte: iniciar a gravação com o valor de sensibilidade máximo (70), e observar o número de barras que aparecem no display. A partir do valor máximo, ajustar a sensibilidade até o nível de variação entre 4 e 5 barras durante a gravação (ver Figura 16)



Figura 16: Tela do usuário no Ultraprobe 10000.

A sensibilidade do instrumento foi ajustada de forma a minimizar a captação ruídos externos dos demais rolos em funcionamento próximos ao rolo medido, assim como ruídos em geral de uma planta em operação. Nas gravações realizadas em campo para este trabalho, o ajuste feito para a gravação foi entre 60 e 65. Na Figura 17 é exibido um dos transportadores do TPM utilizado para a captação de dados.



Figura 17: Transportador TR-315K-09, um dos transportadores utilizado para captação de dados.

O objetivo foi montar bancos de dados de gravações de rolos sem anomalias perceptíveis pelos sentidos humanos (chamados de rolos “não-defeituosos”), e bancos de dados de gravações de rolos com anomalias perceptíveis (através de inspeção visual e auditiva, chamados de rolos “defeituoso”). Como anomalias perceptíveis foram considerados: barulhos periódicos de falha em rolamentos, atrito entre metais e ruídos uniformes de grande intensidade. Os padrões observados nestas duas condições foram utilizados para treinamento de um classificador, que teve seu desempenho posteriormente testado. A precisão dos classificadores é apresentada ao final do trabalho.

#### **4.2. Pré-processamento dos arquivos de áudio obtidos e criação do banco de dados**

Foram feitas 10 gravações em rolos não-defeituosos, e 10 gravações em rolos defeituosos, conforme definição da Seção 4.1. As gravações duraram aproximadamente 20 segundos cada. Para cada gravação, foram realizados dois pré-processamentos básicos: obtenção de dados no domínio do tempo e no domínio da frequência.

Para aquisição dos dados no domínio do tempo, cada arquivo foi transformado em um vetor de dados de amplitude sonora normalizados entre -1 e 1. A taxa amostral foi de 16000 dados para cada segundo de áudio. Posteriormente, para cada teste de desempenho estes vetores de dados seriam separados em intervalos diferentes para ampliação do número de amostras.

A Figura 18 mostra gráficos típicos de valores no domínio do tempo de rolos não-defeituosos e defeituosos. Cada segundo de áudio apresenta 16000 valores de amplitude normalizados de -1 a 1.

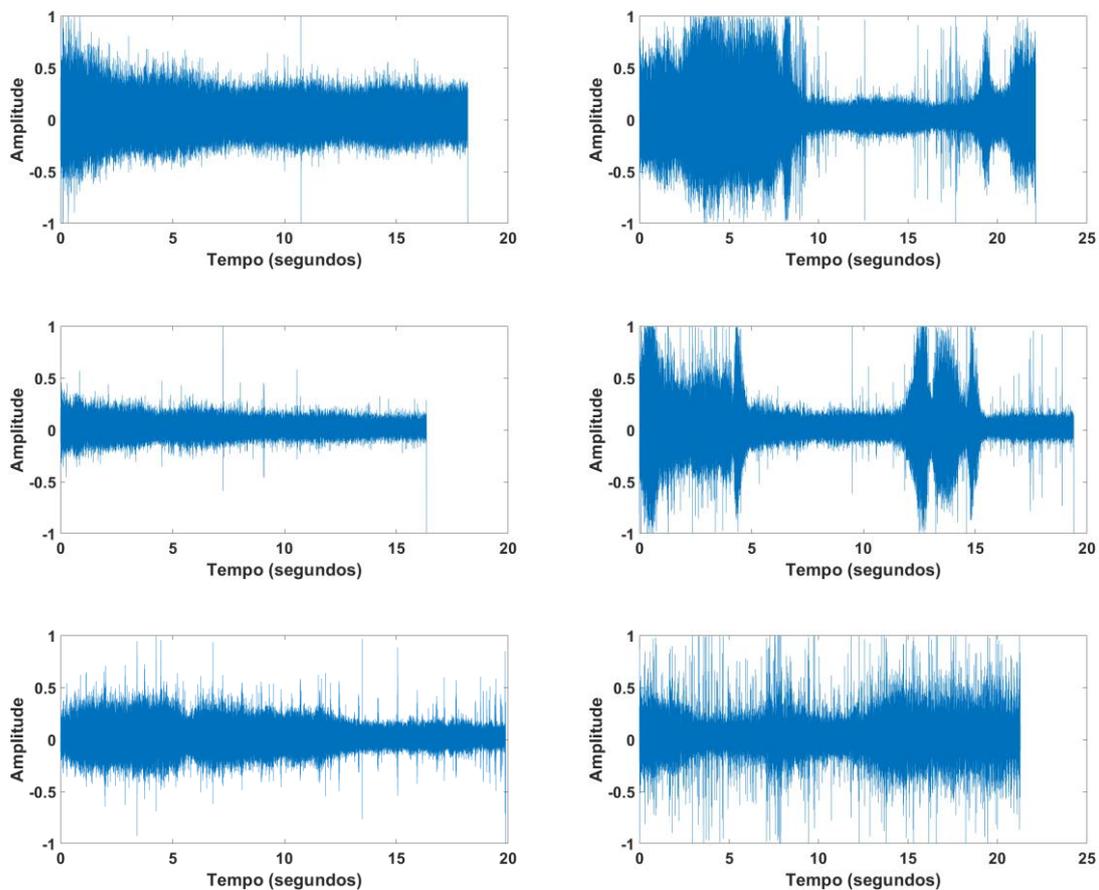


Figura 18: Gráficos amplitude x tempo: à esquerda, rolos sem defeitos; à direita, rolos defeituosos.

Para transformação dos dados no domínio da frequência, em cada amostra de dados foi aplicado o algoritmo FFT, com a distribuição de valor absoluto (magnitude) de 4096 valores de frequência. Posteriormente, estas distribuições de frequências seriam divididas em intervalos para construção das amostras de testes.

A Figura 19 mostra gráficos típicos de valores no domínio da frequência de rolos não-defeituosos e defeituosos. Os valores de frequência vão de 0 a 8000Hz, pois os valores na faixa de ultrassom são transformados para este intervalo pelo sensor para escuta local durante o processo de inspeção. Estes gráficos correspondem aos gráficos dos mesmos rolos exibidos na Figura 18.

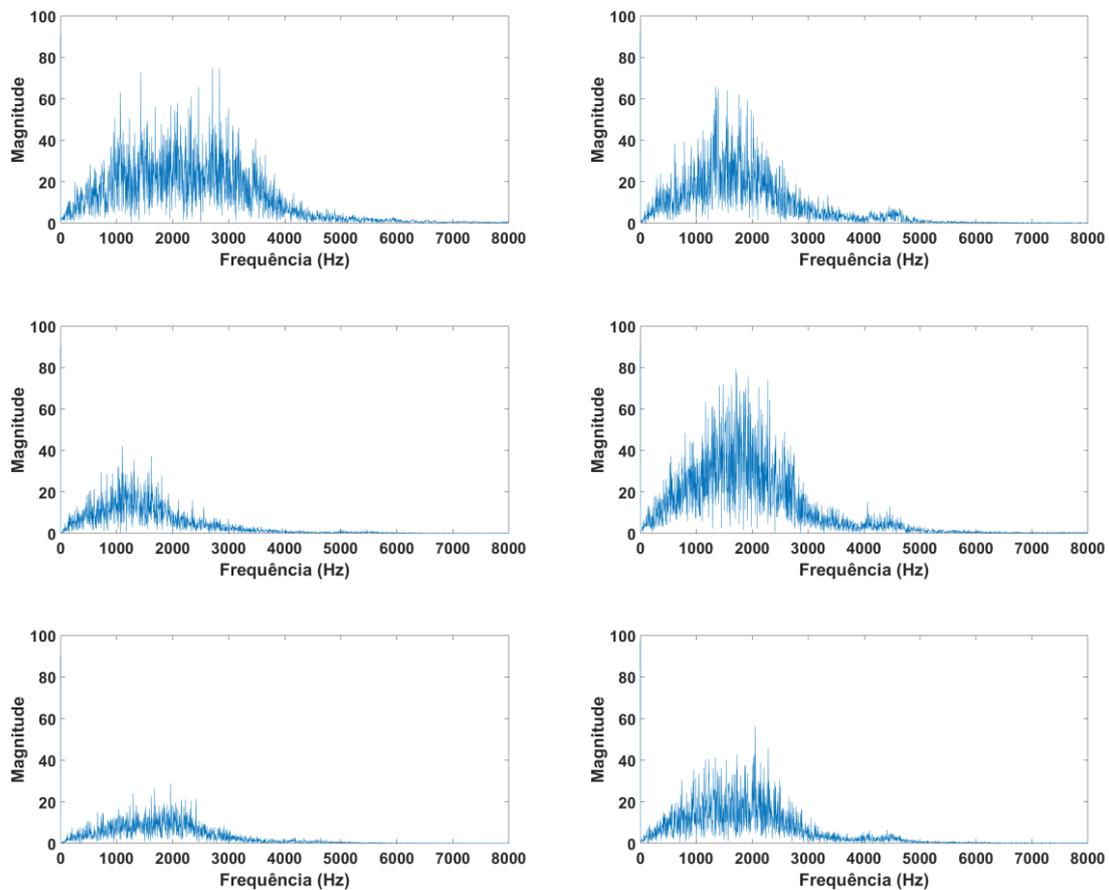


Figura 19: Gráficos magnitude x frequência: à esquerda, rolos sem defeitos; à direita, rolos defeituosos.

### 4.3. Métodos estatísticos e classificação das amostras

Para comparação dos desempenhos dos classificadores, foram elaboradas quatro formas diferentes de experimentos, com diferentes amostragens e extração de características.

#### 4.3.1. Experimento 1 – Vetor de atributos de meio segundo no domínio do tempo

No Experimento 1, para cada amostra de meio segundo de um arquivo de áudio, foram calculadas as seguintes medidas estatísticas: média, mediana e desvio padrão de 8000 valores de amplitude de som ao longo do tempo. Juntamente com essas três medidas, foi colocada a classificação do rolo: não-defeituoso ou defeituoso. Assim, foi formada uma base de dados com

378 amostras de meio segundo classificadas como rolo não-defeituoso, e 390 amostras de meio segundo classificadas como rolo defeituoso.

A forma de montagem dos vetores de atributos está ilustrada na Figura 20. Cada meio segundo de um arquivo de áudio gerou um vetor com três atributos: valor da média, mediana, desvio padrão e classe “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

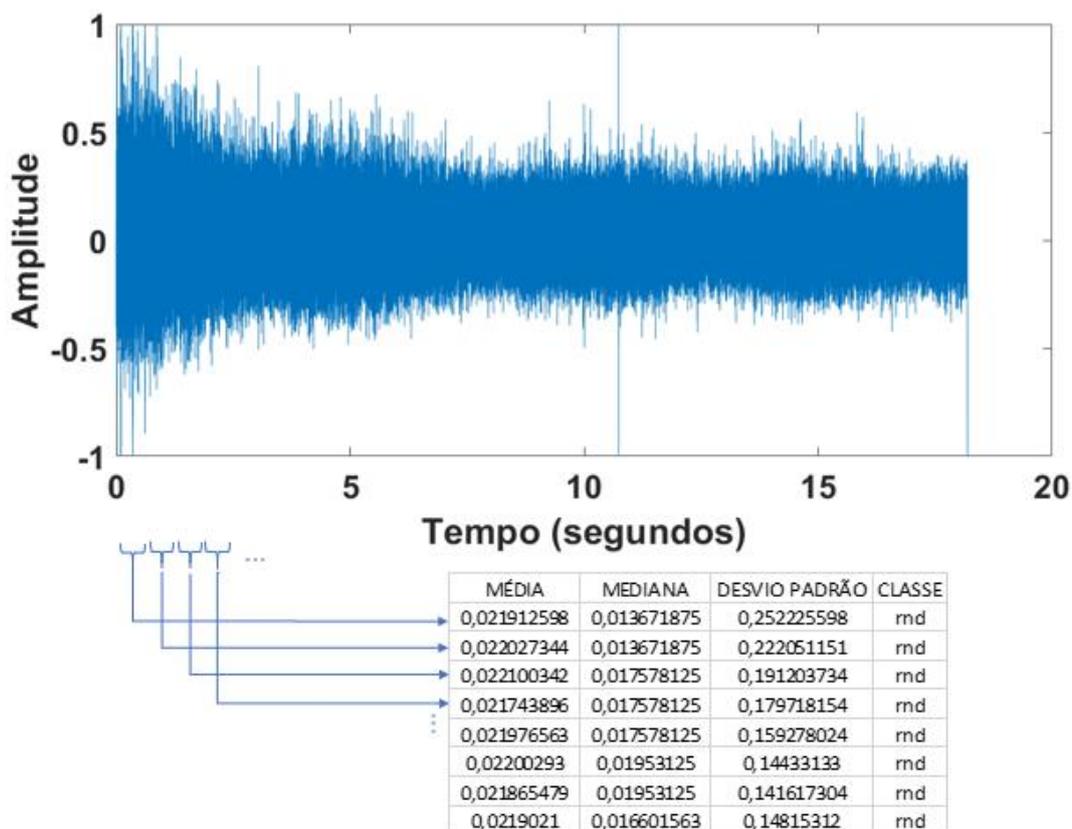


Figura 20: Cada amostra de meio segundo de áudio gerou um vetor de média, mediana, desvio padrão e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

#### 4.3.2. Experimento 2 – Vetor de atributos de 1 segundo no domínio do tempo

No Experimento 2, para cada amostra de um segundo de um arquivo de áudio, foram calculadas as seguintes medidas estatísticas: média, mediana e desvio padrão de cada quarto de segundo. Foi assim montado um vetor com 12 valores de amplitude de som (quatro valores de média, quatro valores de mediana e quatro valores de desvio padrão) ao longo do tempo. Juntamente com essas doze medidas, foi colocada a classificação do rolo: não-defeituoso ou

defeituoso. Assim, foi formada uma base de dados com 192 amostras de um segundo classificadas como rolo não-defeituoso, e 206 amostras de um segundo classificadas como rolo defeituoso.

A forma de montagem dos vetores está ilustrada na Figura 21. Cada segundo de um arquivo de áudio gerou um vetor com doze atributos: quatro valores de média, mediana, desvio padrão e classe “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

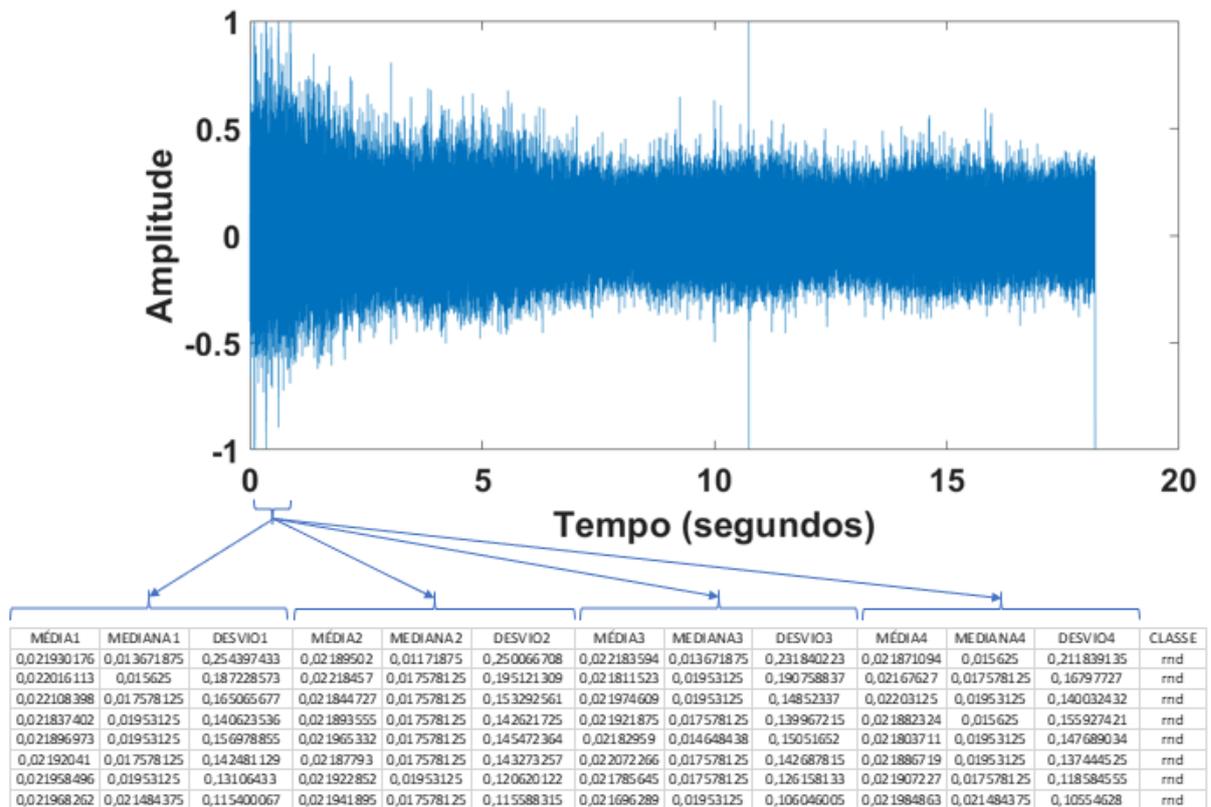


Figura 21: Cada amostra de um segundo de áudio gerou um vetor de média, mediana, desvio padrão para cada quarto de segundo, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

#### 4.3.3. Experimento 3 – FFT do arquivo de áudio completo dividida em 40 partes

No Experimento 3, para cada arquivo de áudio foi realizada a Transformada de Fourier, com o algoritmo FFT, com a distribuição dos valores de amplitude em 4096 valores de frequências diferentes (de 0 a 8000Hz). Cada FFT foi dividida em 40 partes iguais, compostas por 102 valores de frequência. De cada parte foi extraída: média, mediana e desvio padrão.

Assim, foi formada uma base de dados com 400 amostras classificadas como rolo não-defeituoso, e 400 amostras de um segundo classificadas como rolo defeituoso.

A forma de montagem dos vetores está ilustrada na Figura 22. Cada uma das 40 partes da FFT de um arquivo de áudio gerou um vetor com quatro atributos: valor de média, mediana, desvio padrão e classe “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

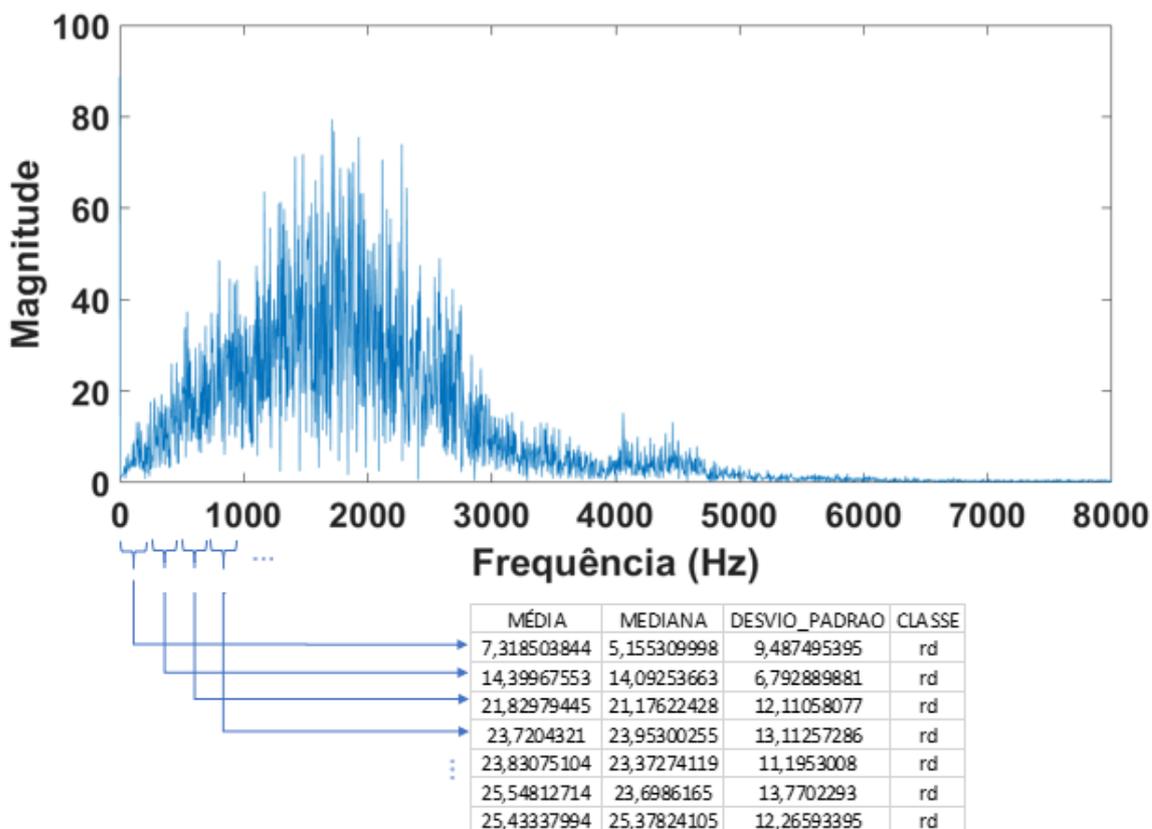


Figura 22: Cada FFT de um arquivo de áudio foi dividida em 40 partes, gerando para cada parte um vetor com valores de média, mediana, desvio padrão, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

#### 4.3.4. Experimento 4 – média móvel da FFT de cada 5 segundos de áudio

No Experimento 4, foi realizada a Transformada de Fourier de cada 5 segundos de gravação, com o algoritmo FFT, com a distribuição dos valores de amplitude em 4096 valores de frequências diferentes (de 0 a 8000Hz). Cada FFT foi dividida em 40 partes iguais, compostas por 102 valores de frequência. De cada parte foi extraída a média, caracterizando a

média móvel da FFT. As amostras foram então vetores de 40 valores de média. Assim, foi formada uma base de dados com 37 amostras classificadas como rolo não-defeituoso, e 40 amostras de um segundo classificadas como rolo defeituoso.

A forma de montagem dos vetores está ilustrada na Figura 23. Cada FFT de 5 segundos de áudio gerou um vetor com 41 atributos: valor de média de cada uma das 40 partes e classe “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

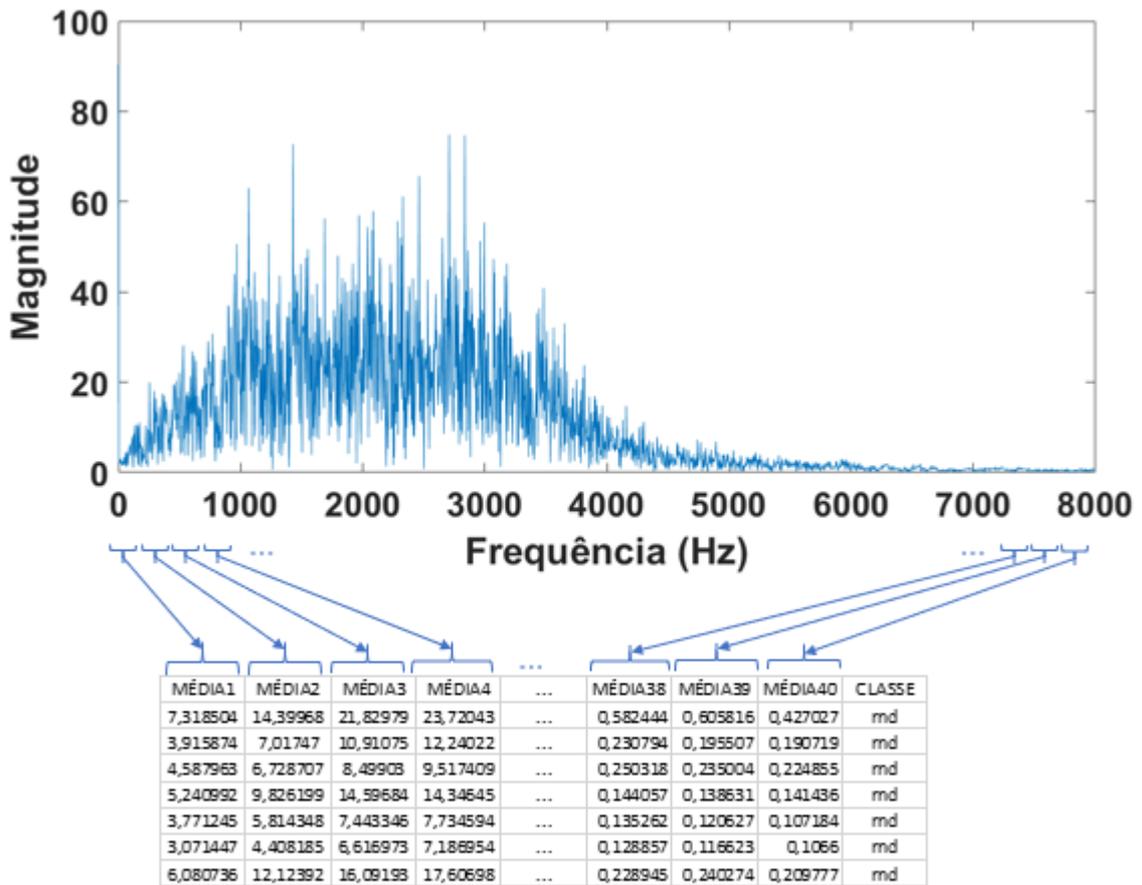


Figura 23: Para cada 5 segundos de áudio foi aplicada a FFT e esta foi dividida em 40 partes, gerando um vetor de 40 valores médios para cada FFT, e uma classificação em “rnd” (rolo não-defeituoso) ou “rd” (rolo defeituoso).

#### 4.4. Treinamento e teste de desempenho do classificador

O total de amostras classificadas em “não-defeituoso” e “defeituoso” foram utilizadas para treinamento de algoritmos de aprendizado de máquina tipo RF e MLP, para diversidade na tarefa de classificação e comparação de desempenho.

A base de dados total foi utilizada para rodar os classificadores RF e MLP utilizando o pacote de softwares Waikato Environment for Knowledge Analysis (HALL et al., 2019). Os experimentos foram realizados utilizando validação cruzada repetida 10 vezes para avaliar o desempenho. Neste caso, os dados são divididos em 10 conjuntos de mesmo número de dados. Em uma dada repetição, um conjunto é utilizado para teste e os outros 9 são utilizados para treinamento. Nas demais 9 repetições, um conjunto diferente é usado para teste e os restantes para treinamento. Desta forma, cada dado do conjunto é utilizado 9 vezes para treinamento e 1 vez para teste. O resultado final exibido é a performance média obtida nas 10 repetições. Cada teste foi iniciado com 10 condições iniciais (valores de “seed”) diferentes para efeitos de comparação da simulação.

Cada experimento foi realizado com as seguintes configurações de RF e MLP: RF com 10 árvores (RF10), RF com 50 árvores (RF50), RF com 100 árvores (RF100), MLP com uma camada de 5 neurônios (MLP5), MLP com uma camada de 10 neurônios (MLP10), MLP com duas camadas de 5 neurônios cada (MLP5x5) e MLP com duas camadas de 10 neurônios cada (MLP10x10). Foram montados gráficos com *boxplots* com a porcentagem de classificações corretas no teste de cada configuração.

Diversas configurações das implementações de RF e MLP foram testadas. No caso das RF, configurações acima de 100 árvores não apresentaram melhoras nos resultados de classificação. No caso das MLP, números maiores de camadas ou neurônios prejudicavam o desempenho do classificador, mostrando características de “*overfitting*” (excesso de especialização em relação ao conjunto de treinamento).

## 5. Resultados e Discussão

Este Capítulo apresenta os resultados dos experimentos realizados com objetivo de avaliar a metodologia proposta no Capítulo 1.

Os Experimentos 1 e 2 foram realizados utilizando-se o domínio do tempo. Os Experimentos 3 e 4 foram realizados utilizando-se o domínio da frequência, através do algoritmo FFT aplicado nos valores extraídos do domínio do tempo. Foram realizados experimentos com algoritmo RF com diferentes configurações: 10, 50 e 100 árvores, e MLP com diferentes configurações: uma camada com 5 e 10 neurônios e duas camadas com 5 e 10 neurônios cada. Cada experimento foi repetido 10 vezes com diferentes valores de *seed*. Foi aplicada validação cruzada repetida 10 vezes em cada experimento. Estas configurações foram repetidas para os Experimentos 1, 2, 3 e 4.

Primeiramente foi realizada uma análise de dispersão dos atributos de entrada, para avaliação de sobreposição e como isso deveria se refletir na qualidade da classificação. Depois, foi montado um diagrama de caixas (*boxplot*) para efeito de comparação de desempenho com as taxas de acerto de classificação de cada experimento.

### 5.1. Análise de dispersão dos atributos de entrada

Através da análise de dispersão dos atributos de entrada utilizados nos algoritmos de ML, podemos inferir quais atributos e métodos estatísticos são mais característicos na distinção entre rolos não-defeituosos e defeituosos. Nos experimentos realizados foram utilizadas diferentes formas de organizar os atributos, e os gráficos de dispersão mostram quais formas apresentam maior sobreposição de dados, que caracterizam maior dificuldade de classificação para o algoritmo.

A Figura 24 exibe a comparação entre os dados utilizados no Experimento 1. Foram colocadas lado a lado as distribuições dos valores dos atributos para comparação de sobreposição entre os valores correspondentes aos rolos não-defeituosos (em cor amarela) e os rolos defeituosos (em cor azul). A figura mostra a distribuição dos valores de média, mediana e desvio padrão.

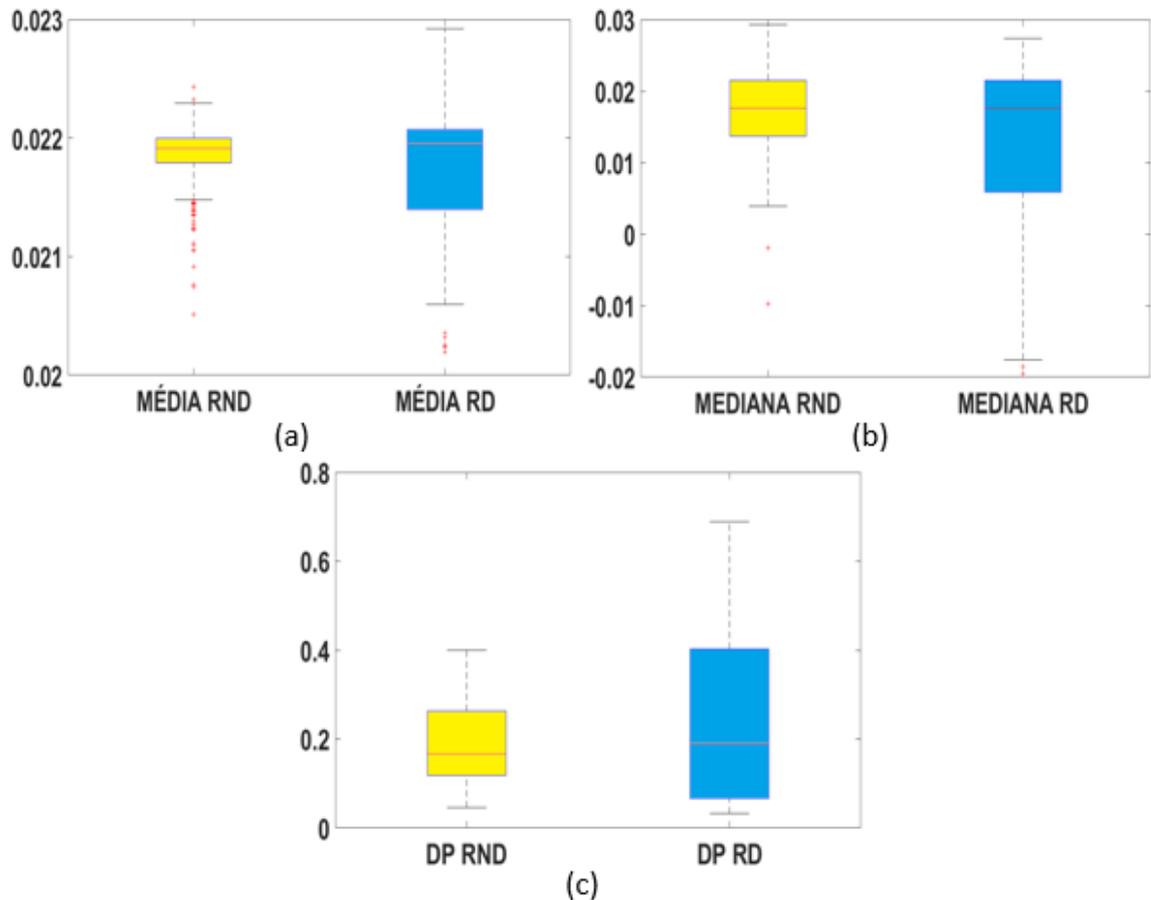
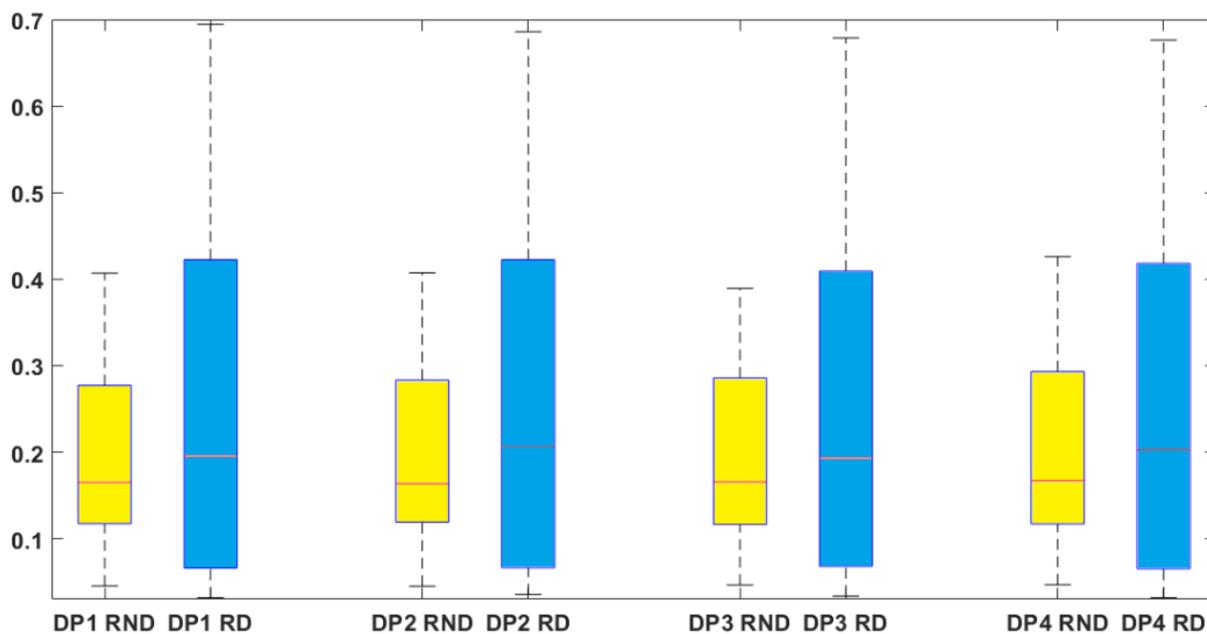


Figura 24: Comparação de gráficos de dispersão para os atributos do Experimento 1. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP).

A Figura 25 exibe a comparação entre os dados utilizados no Experimento 2, para comparação de sobreposição de valores. A Figura 25 (a) mostra a comparação entre rolo não-defeituoso e defeituoso da distribuição das 4 médias. A Figura 25 (b) mostra a comparação na distribuição das medianas e a Figura 25 (c) mostra a comparação na distribuição dos desvios padrão.





(c)

Figura 25: Comparação de gráficos de dispersão para os atributos do Experimento 2. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP).

A Figura 26 exibe a comparação entre os dados utilizados no Experimento 3. Pode-se perceber a maior sobreposição entre os valores correspondentes aos rolos não-defeituosos e os rolos defeituosos neste caso em relação aos demais experimentos. A figura mostra a distribuição dos valores de média, mediana e desvio padrão.

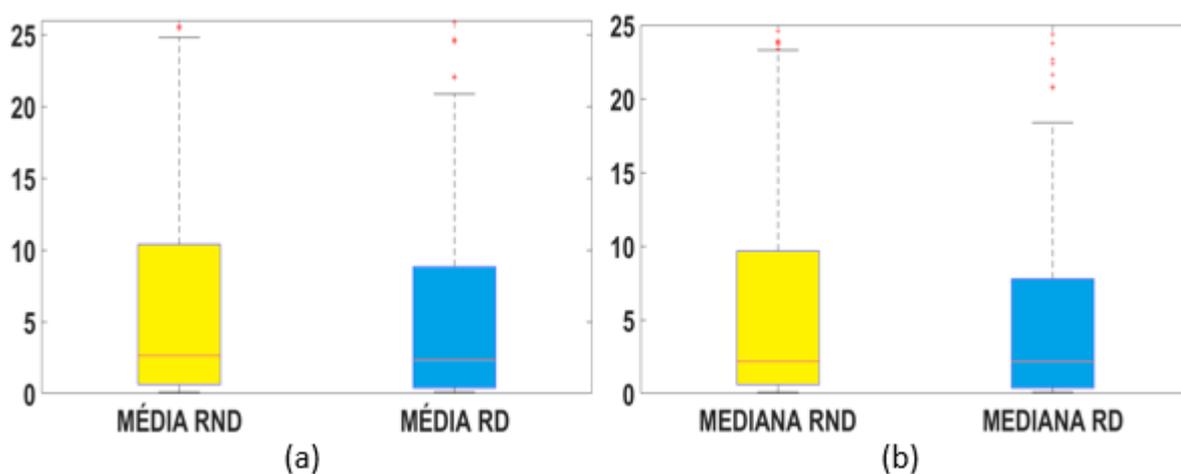


Figura 26: cont.

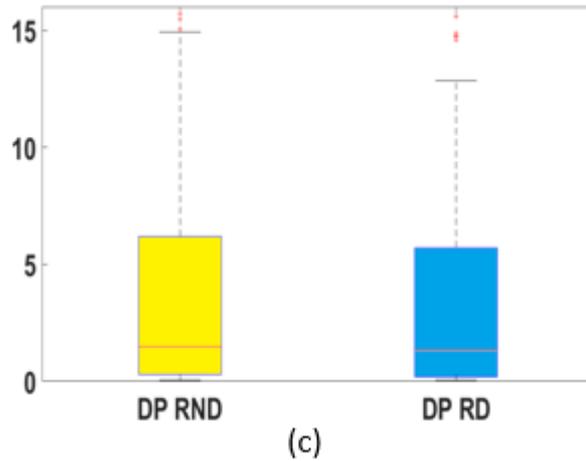


Figura 26: Comparação de gráficos de dispersão para os atributos do Experimento 3. “RND” significa rolo não-defeituoso e “RD” significa rolo defeituoso. (a) Distribuição dos valores de média. (b) Distribuição dos valores de mediana. (c) Distribuição dos valores de desvio padrão (DP).

A Figura 27 exhibe a comparação entre os dados utilizados no Experimento 4. A figura mostra os pares de média lado a lado para comparação entre rolos não-defeituosos (cor amarela) e defeituosos (cor azul). O primeiro par (1 e 2), por exemplo, exhibe a primeira média dos rolos: 1 é a média dos rolos não-defeituosos e 2 é a média dos rolos defeituosos. O segundo par (3 e 4) exhibe a segunda média dos rolos: 3 é a média dos rolos não-defeituosos e 4 é a média dos rolos defeituosos; e assim sucessivamente. Pode-se perceber pela distribuição que estes valores apresentam menos sobreposição com relação aos experimentos anteriores, e seria esperado que este experimento apresentasse melhores resultados pois os valores utilizados neste caso são mais distintos para cada tipo de rolo.

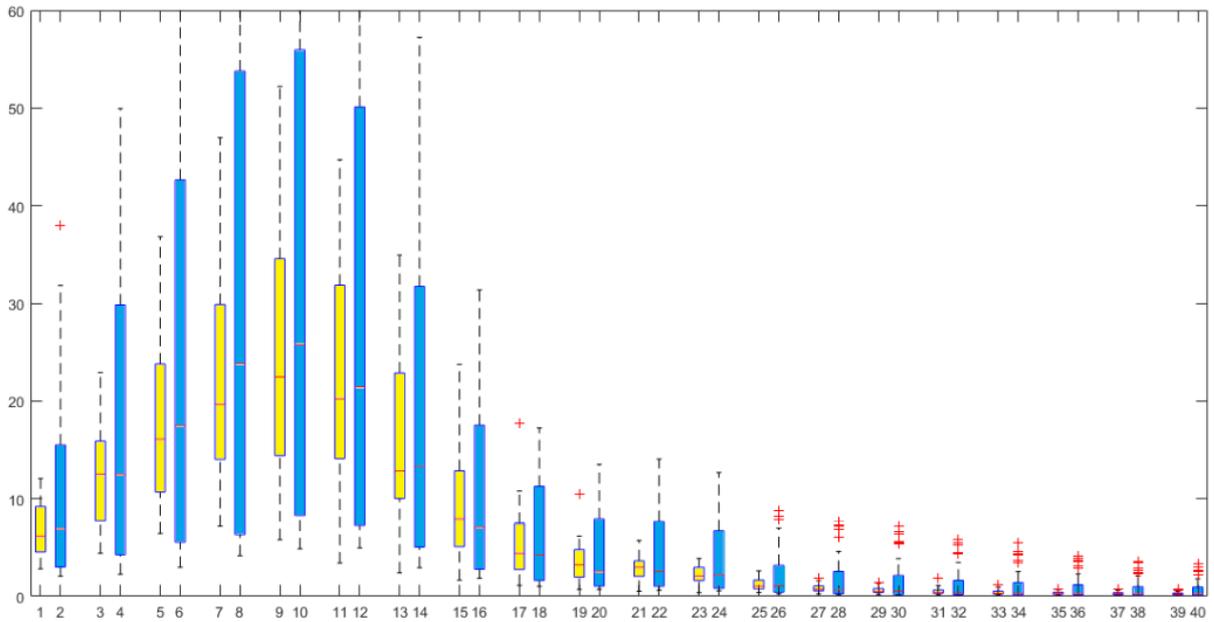


Figura 27: Comparação de gráficos de dispersão para os atributos do Experimento 4. Cada par de *boxplots* apresenta a comparação entre a dispersão de rolos não-defeituosos (amarelo) e defeituosos (azul).

## 5.2. Avaliação de desempenho com matriz de confusão

Esta seção explica a forma de avaliação do desempenho dos experimentos, considerando as seguintes medidas de desempenho: sensibilidade, precisão, especificidade, valor preditivo negativo e acurácia. Estas medidas são exibidas em uma matriz de confusão, conforme mostra a Figura 28.

Dada uma amostra com casos positivos e negativos para uma determinada detecção, os resultados dos testes de classificação dos casos analisados podem ser divididos em quatro grupos: VP (Verdadeiros Positivos): número de casos corretamente classificados como positivos; FP (Falsos Positivos): número de casos erroneamente classificados como positivos; VN (Verdadeiros Negativos): número de casos corretamente classificados como negativos; e FN (Falsos Negativos): número de casos erroneamente classificados como negativos. Estes valores são utilizados para gerar as medidas adotadas por este trabalho.

O objetivo é demonstrar de que forma os erros aparecem em cada experimento, dado que, dos dois tipos de erros possíveis (falsos negativos e falsos positivos), o falso negativo

representa maior prejuízo para a aplicação desta metodologia. O falso negativo representaria um caso em que um rolo defeituoso deixou de ser detectado pela metodologia, o que poderia causar danos ao transportador, especialmente se o sistema de detecção for totalmente dependente da classificação automática. O falso positivo seria apenas um falso alarme, causando inconvenientes e custos para a equipe de manutenção que iria realizar a troca do componente sem real necessidade; entretanto, não haveria danos ao equipamento neste caso.

|             |          | CLASSE PREVISTA          |                          |                |
|-------------|----------|--------------------------|--------------------------|----------------|
|             |          | Positivo                 | Negativo                 |                |
| CLASSE REAL | Positivo | Verdadeiro positivo (VP) | Falso negativo (FN)      | Sensibilidade  |
|             | Negativo | Falso positivo (FP)      | Verdadeiro negativo (VN) | Especificidade |
|             |          | Precisão                 | Valor preditivo negativo | Acurácia       |

Figura 28: Matriz de confusão com as medidas de desempenho.

### 5.2.1. Sensibilidade

Sensibilidade mede o desempenho da classificação em relação aos casos positivos. É dada pela razão entre o número de previsões positivas corretas e o número total de casos positivos. No caso deste trabalho, seria a razão entre o número de rolos classificados como não-defeituosos pelo algoritmo em relação ao total real de rolos não-defeituosos. Os melhores desempenhos são números mais próximos de 1.

$$Sensibilidade = \frac{VP}{VP + FN}$$

### 5.2.2. Especificidade

Especificidade mede o desempenho da classificação em relação aos casos negativos. É dada pela razão entre o número de previsões negativas corretas e o número total de casos negativos. No caso deste trabalho, seria a razão entre o número de rolos classificados como defeituosos pelo algoritmo em relação ao total real de rolos defeituosos. Os melhores desempenhos são números mais próximos de 1.

$$\text{Especificidade} = \frac{VN}{VN + FP}$$

### 5.2.3. Precisão

Precisão é dada pela razão entre o número de previsões positivas corretas e o número total de previsões positivas. No caso deste trabalho, seria a razão entre o número de rolos classificados corretamente como não-defeituosos pelo algoritmo em relação ao total de casos previstos como não-defeituosos. Os melhores desempenhos são números mais próximos de 1.

$$\text{Precisão} = \frac{VP}{VP + FP}$$

### 5.2.4. Valor preditivo negativo

Valor preditivo negativo é dado pela razão entre o número de previsões negativas corretas e o número total de previsões negativas. No caso deste trabalho, seria a razão entre o número de rolos classificados corretamente como defeituosos pelo algoritmo em relação ao total de casos previstos como defeituosos. Os melhores desempenhos são números mais próximos de 1.

$$\text{Valor preditivo negativo} = \frac{VN}{VN + FN}$$

### 5.2.5. Acurácia

Acurácia mede a porcentagem total de casos classificados corretamente. É dada pela razão entre o número total de previsões corretas e o número total de amostras. Os melhores desempenhos são números mais próximos de 1.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}$$

### 5.3. Investigação com Aprendizado de Máquina no domínio do tempo

Esta seção demonstra os resultados obtidos nos Experimentos 1 e 2, que utilizam os dados no domínio do tempo, através de percentuais de classificações corretas.

#### 5.3.1. Experimento 1 – Utilizando amostras de meio segundo com média, mediana e desvio padrão

O primeiro experimento foi realizado nas amostras de meio segundo extraídas das 20 gravações de rolos descritas na Seção 4.3.1.

O melhor resultado para as RF foi obtido com a configuração de 100 árvores e seed 5. Neste caso o teste com o algoritmo RF com 100 árvores apresentou 74,2188% de instâncias classificadas corretamente. A Figura 29 apresenta a matriz de confusão com as relações entre verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos para este caso.

|             |     | CLASSE PREVISTA  |                                  |                        |
|-------------|-----|------------------|----------------------------------|------------------------|
|             |     | RND              | RD                               |                        |
| CLASSE REAL | RND | 298              | 80                               | Sensibilidade = 0,788  |
|             | RD  | 114              | 276                              | Especificidade = 0,707 |
|             |     | Precisão = 0,723 | Valor preditivo negativo = 0,775 | Acurácia = 0,742       |

Figura 29: Matriz de confusão para o melhor caso do Experimento 1.

Foram realizados também testes com a rede neural MLP com as seguintes configurações de camadas: 5, 10, 5x5 e 10x10. No Experimento 1, a MLP apresentou resultados de classificação inferiores em relação às RF. Os piores desempenhos foram os das MLP com duas camadas, enquanto a melhor porcentagem de acerto foi obtida pela MLP com 1 camada apenas de 5 neurônios.

A Figura 30 exibe um diagrama de caixas com os resultados do Experimento 1. Pode-se perceber o melhor desempenho dos algoritmos RF em relação às redes MLP, com menor variabilidade e melhores taxas de acerto.

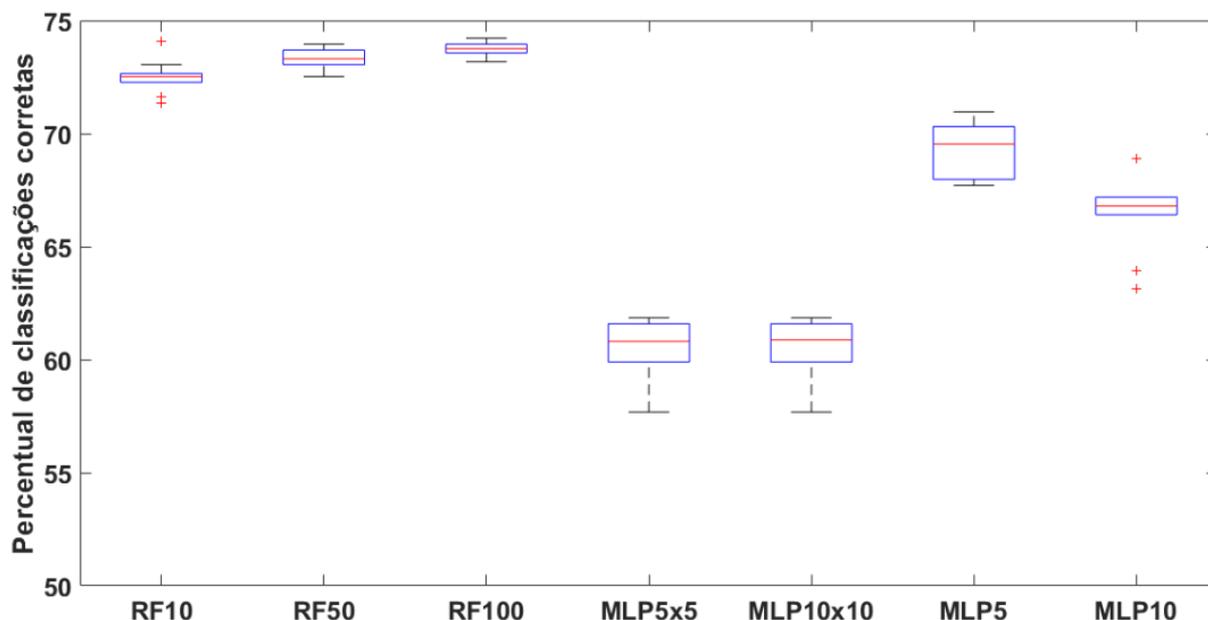


Figura 30: Resultados obtidos no Experimento 1.

### 5.3.2. Experimento 2 – Utilizando amostras um segundo com vetor de 12 valores de média, mediana e desvio padrão

O segundo experimento foi realizado nas amostras de um segundo extraídas das 20 gravações de rolos descritas na Seção 4.3.2.

O melhor resultado foi obtido com o algoritmo RF com 100 árvores, que apresentou 75,1256% de instâncias classificadas corretamente quando aplicado com validação cruzada. A

Figura 31 apresenta a matriz de confusão com as relações entre verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos para este caso.

|             |     | CLASSE PREVISTA  |                                  |                        |
|-------------|-----|------------------|----------------------------------|------------------------|
|             |     | RND              | RD                               |                        |
| CLASSE REAL | RND | 150              | 42                               | Sensibilidade = 0,781  |
|             | RD  | 57               | 149                              | Especificidade = 0,723 |
|             |     | Precisão = 0,724 | Valor preditivo negativo = 0,780 | Acurácia = 0,751       |

Figura 31: Matriz de confusão para o melhor caso do Experimento 2.

Pela Figura 32 percebe-se, assim como no Experimento 1, que a rede neural MLP apresentou desempenho consideravelmente inferior. As MLP com 2 camadas de neurônios cada (MLP5x5 e MLP10x10) apresentaram, no pior caso, 59,2965% de instâncias classificadas corretamente quando aplicado com validação cruzada repetida 10 vezes.

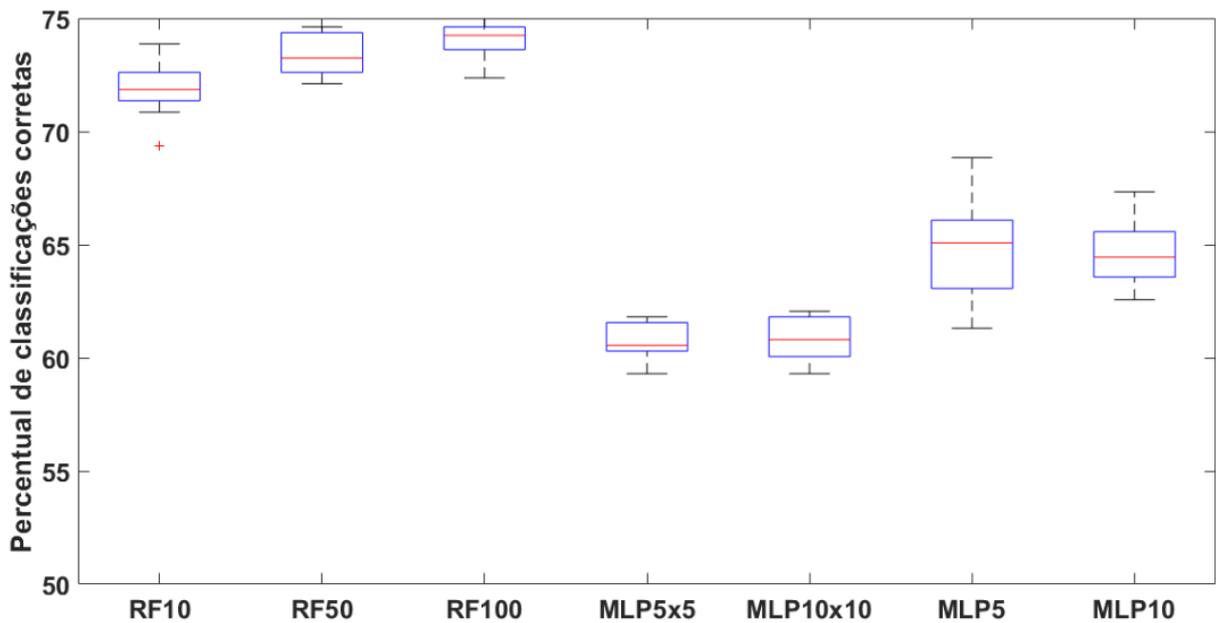


Figura 32: Resultados obtidos no Experimento 2.

## 5.4. Investigação com Aprendizado de Máquina no domínio da frequência

Esta seção demonstra os resultados obtidos nos Experimentos 3 e 4, que utilizam os dados no domínio da frequência, através de percentuais de classificações corretas.

### 5.4.1. Experimento 3 – Utilizando vetor de 3 valores para FFT

O terceiro experimento foi realizado nas amostras 40 partes de uma FFT aplicada em cada uma das 20 gravações de rolos descritas na Seção 4.3.3. A Figura 33 apresenta a matriz de confusão com as relações entre verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos para este caso.

Esta forma de experimento foi a que apresentou os piores resultados, como pode ser observado na Figura 34. O algoritmo RF apresentou desempenho inferior aos Experimentos 1 e 2, com melhor taxa de acerto de 66,5% nas configurações com 50 e 100 árvores, enquanto as MLP apresentou como pior resultado uma taxa de 47,5% nas configurações com duas camadas.

|             |     | CLASSE PREVISTA  |                                  |                       |
|-------------|-----|------------------|----------------------------------|-----------------------|
|             |     | RND              | RD                               |                       |
| CLASSE REAL | RND | 256              | 144                              | Sensibilidade = 0,64  |
|             | RD  | 124              | 276                              | Especificidade = 0,69 |
|             |     | Precisão = 0,673 | Valor preditivo negativo = 0,657 | Acurácia = 0,665      |

Figura 33: Matriz de confusão para o melhor caso do Experimento 3.

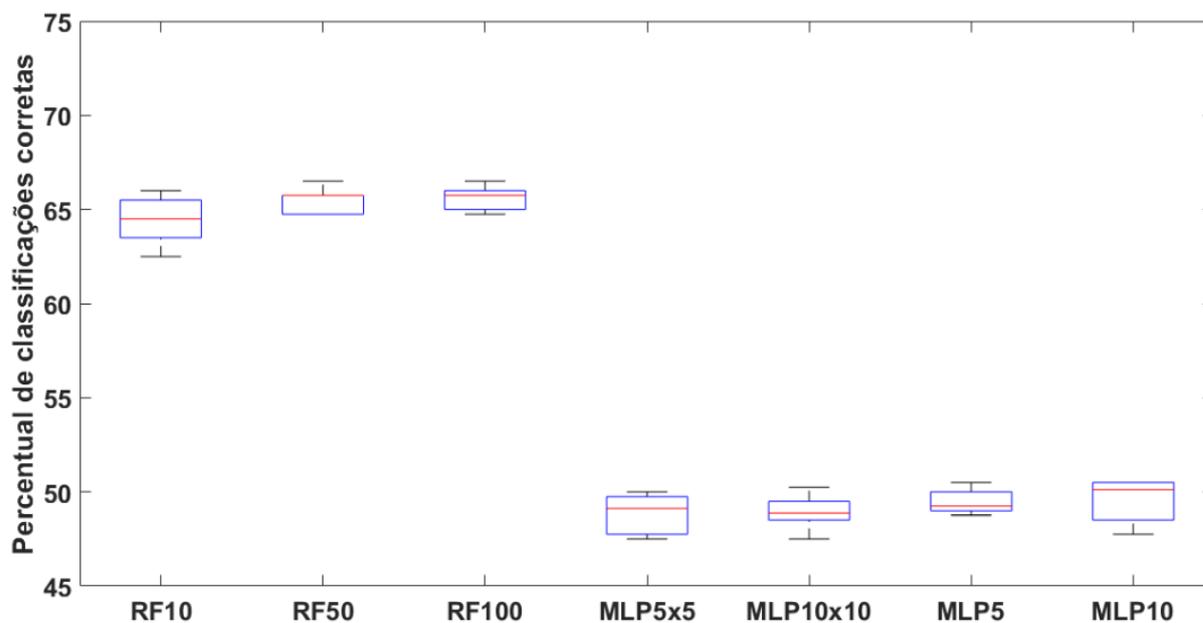


Figura 34: Resultados obtidos no Experimento 3.

#### 5.4.2. Experimento 4 – Utilizando média móvel das FFT de amostras de 5 segundos

O quarto experimento foi realizado nas amostras 40 partes de uma FFT aplicada em cada 5 segundos das 20 gravações de rolos descritas na Seção 4.3.4. A Figura 35 apresenta a matriz de confusão com as relações entre verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos para este caso.

Esta forma de experimento foi a que apresentou os melhores resultados, como pode ser observado na Figura 36. O melhor resultado foi obtido pela MLP com 10 neurônios, que obteve em um dos testes porcentagem de classificação correta de 90% dos casos.

|             |     | CLASSE PREVISTA  |                                 |                        |
|-------------|-----|------------------|---------------------------------|------------------------|
|             |     | RND              | RD                              |                        |
| CLASSE REAL | RND | 35               | 5                               | Sensibilidade = 0,875  |
|             | RD  | 3                | 37                              | Especificidade = 0,925 |
|             |     | Precisão = 0,921 | Valor preditivo negativo = 0,88 | Acurácia = 0,9         |

Figura 35: Matriz de confusão para o melhor caso do Experimento 4.

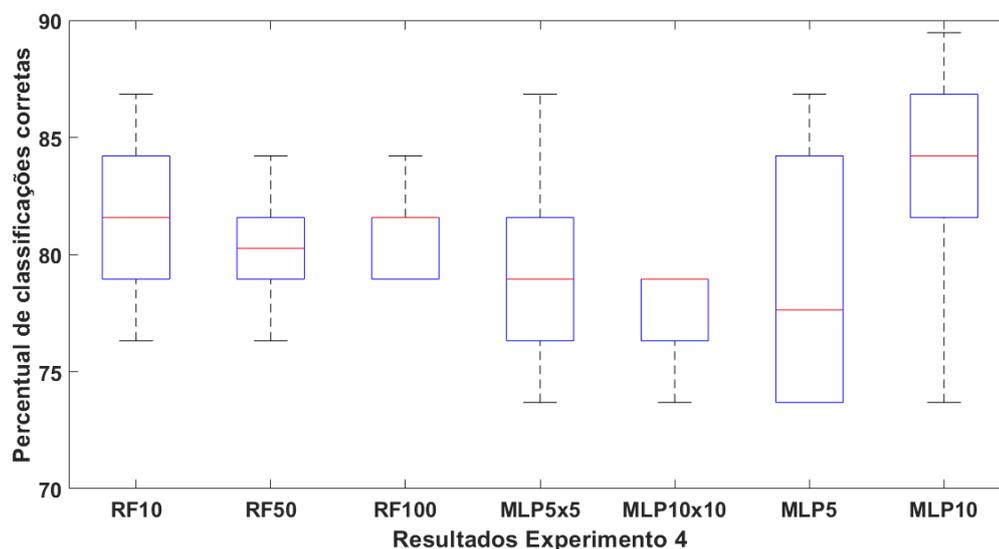


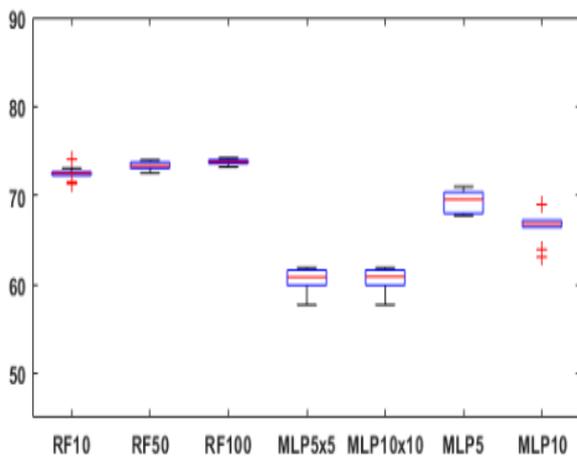
Figura 36: Resultados obtidos no Experimento 4.

## 5.5. Resultados gerais dos experimentos

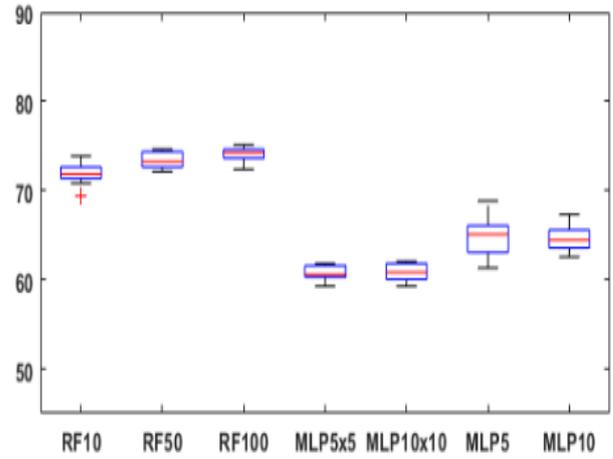
A Figura 37 apresenta, para os 4 experimentos propostos, os resultados de taxa de acerto colocados juntos para efeito comparativo. Para os Experimentos 1 e 2, utilizando domínio do tempo, podemos ver que o RF100 foi o método que atingiu os melhores resultados tanto para média geral como para melhor caso. O melhor caso para E1 foi 74,22% e o melhor caso para E2 foi 75,13%, as médias foram respectivamente 73,74% e 73,94%. Pela similaridade entre os resultados de E1 e E2, verifica-se que o vetor de atributos mais detalhado não foi capaz de trazer taxas de acerto mais altas. Os piores resultados gerais foram obtidos no Experimento 3, que

emprega o mesmo vetor de atributos do Experimento 1, mas utilizando a FFT. A taxa de acerto para o melhor caso foi de 66,5%, com média de 65,63% para o RF100.

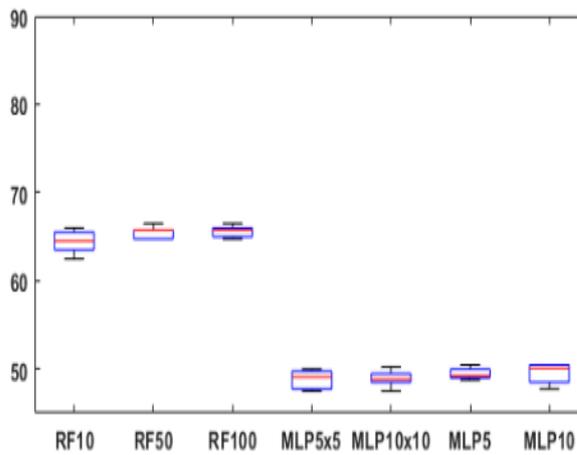
Os melhores resultados gerais foram obtidos no Experimento 4, que emprega FFT em um vetor de atributos mais detalhado, composto por 40 valores de média. Podemos ver para o Experimento 4 que diversos métodos de aprendizado superam os 80% da taxa de acerto. O melhor modelo neste caso foi a MLP10, com média de acerto de 83,68%, sendo o melhor caso com taxa de acerto de 90%. O que E4 ressalta é que, para os dados observados, tanto o modelo como o pré-processamento e definição dos atributos são fundamentais.



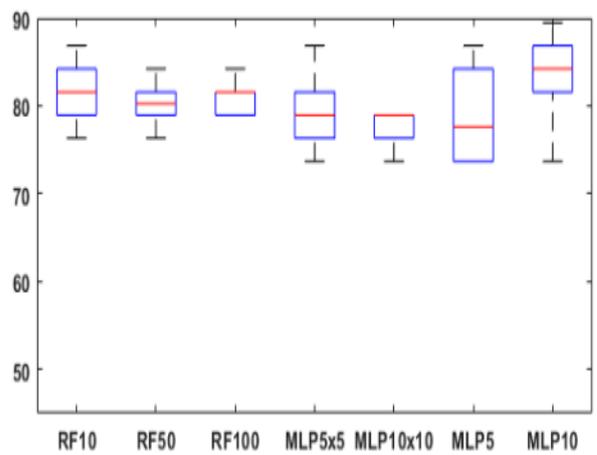
(a) Resultado E1



(b) Resultado E2



(c) Resultado E3



(d) Resultado E4

Figura 37: Taxas de acerto de todos os 4 experimentos.

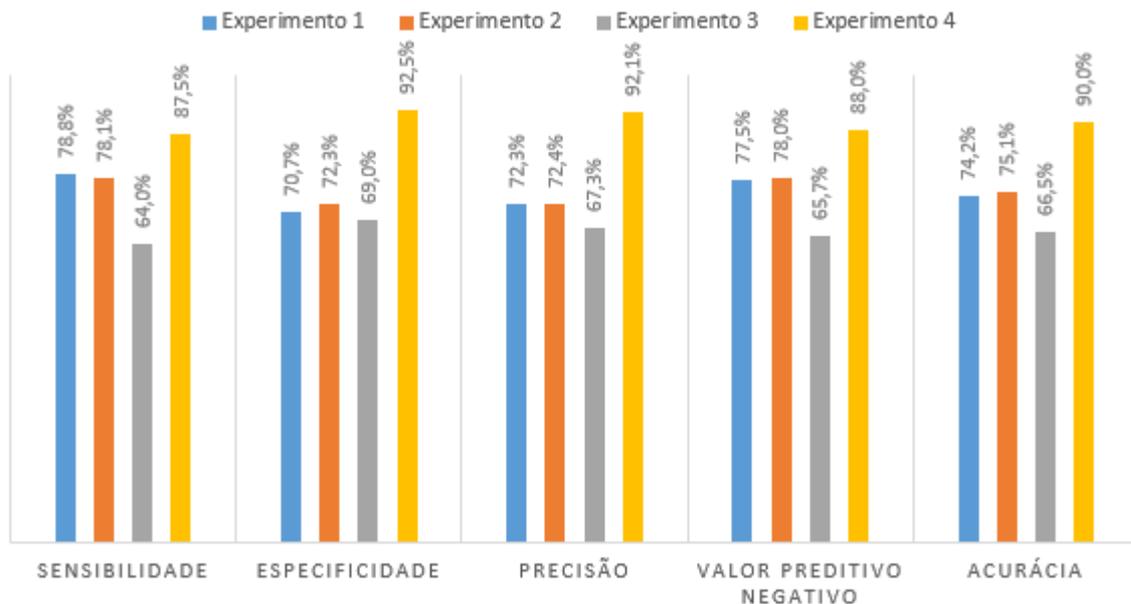


Figura 38: Análise de desempenho de todos os 4 experimentos.

A Figura 38 apresenta, para os 4 experimentos propostos, os indicadores de desempenho para o melhor teste. Percebe-se que o Experimento 4 apresentou os melhores desempenhos em todos os indicadores. Isto é particularmente importante pois os casos de falsos negativos são especialmente preocupantes para esta aplicação. Um falso negativo seria um rolo defeituoso que não foi detectado pela metodologia. Para os indicadores de casos negativos (especificidade e valor preditivo negativo), observa-se o bom desempenho do Experimento 4 com relação aos demais.

Os resultados obtidos por este trabalho foram satisfatórios na detecção de anomalias percebidas por inspeção humana visual e sonora. Os trabalhos pesquisados na área de detecção de falhas por sensoriamento acústico apresentam aplicações em condições controladas em laboratório com defeitos conhecidos, logo este trabalho contribui na aplicação desta metodologia em equipamentos operacionais em uma planta industrial.

Em comparação com os trabalhos relacionados, neste caso os defeitos não eram previamente conhecidos, pois foram feitas gravações em rolos já instalados e funcionais nos transportadores do TPM. A forma de detecção de defeitos para rolos “não-defeituosos” foi a percepção dos sentidos humanos. Existe a possibilidade de que alguns desses rolos

considerados não-defeituosos já apresentassem sinais indicativos na faixa de ultrassom que evidentemente não puderam ser percebidos. Já os rolos “defeituosos” já apresentavam efeitos perceptíveis na faixa de som audível, portanto a capacidade previsão com maior antecedência fica apenas inferida pelo trabalho. Fica comprovado, porém, que o ruído causado pelos demais rolos em funcionamento próximo e demais ruídos da planta não atrapalharam na captação do padrão de defeitos.

## 6. Conclusão e Trabalhos Futuros

O trabalho desenvolvido apresentou a criação de uma metodologia para detecção de defeitos operacionais em rolos de transportadores de correia e posterior classificação em rolo não-defeituoso ou defeituoso. Os dados de entrada para esta classificação foram obtidos através da gravação acústica com um sensor que capta frequências na faixa de ultrassom. Após gravação feita em rolos que não apresentavam defeito aparente (não-defeituosos) e em rolos que apresentavam defeitos perceptíveis (defeituosos), a base de dados foi formada e os dados foram tratados de forma a extrair características que determinavam o padrão de cada caso. A base de dados devidamente rotulada foi usada então para treino e teste de classificadores automáticos usando duas técnicas de aprendizado de máquina, para efeitos de comparação: *Random Forests* e *Multilayer Perceptron*. Foram criados quatro tipos de experimentos (com os dados no domínio do tempo e da frequência organizados com diferentes vetores de atributos) para avaliar quais formas de pré-processamento e definição de atributos são mais eficientes na obtenção de resultados.

Os resultados obtidos no Capítulo 4 demonstram que o bom desempenho da metodologia desenvolvida depende bastante da forma como cada experimento organizou a base de dados. Os Experimentos 1 e 2 usam os dados no domínio do tempo: o melhor caso para o Experimento 1 foi 74,22% e o melhor caso para o Experimento 2 foi 75,13%, e as médias foram respectivamente 73,74% e 73,94%. Os piores resultados gerais foram obtidos no Experimento 3, que emprega o mesmo vetor de atributos do Experimento 1, mas utilizando o domínio da frequência. A taxa de acerto para o melhor caso foi de 66,5%, com média de 65,63%. Os melhores resultados gerais foram obtidos no Experimento 4, que emprega no domínio da frequência um vetor de atributos mais detalhado com a média móvel de 40 valores. Para o Experimento 4 diversos métodos de aprendizado superaram os 80% da taxa de acerto. O melhor modelo neste caso apresentou média de acerto de 83,68%, sendo o melhor caso com taxa de acerto de 90%.

O presente trabalho apresenta o potencial da detecção de padrões indicativos de falhas iminentes em rolos de transportadores de correia. A forma de detecção escolhida, ultrassom, mostrou que pode apresentar bom poder de predição dos padrões típicos que um rolo apresenta antes da falha. A tecnologia tem possibilidade de aplicação em um sistema automatizado que iria monitorar dinamicamente toda a extensão da correia e indicar com antecedência quais rolos

estão apresentando os sintomas característicos de falha. A forma de captação dos sinais pode se dar por sensores fixos ou através de equipamentos móveis (robôs, por exemplo) que percorrem toda a extensão da correia capturando continuamente as gravações. As informações poderiam ser enviadas então para um sistema centralizado que iria ter toda a informação de condições de todos os transportadores da planta para gerenciamento.

Embora esta metodologia tenha apresentado resultados satisfatórios, trabalhos futuros podem explorar ainda mais a forma de pré-processamento dos dados, com outros métodos estatísticos, outras formas de organizar os atributos e outras técnicas de reconhecimento de padrões com inteligência artificial, visto que essas etapas mostraram diferentes resultados neste trabalho. A formação de um banco de dados com mais gravações também pode trazer melhores resultados no aprendizado dos algoritmos. Outra possibilidade é a utilização de métodos não-supervisionados no treinamento do classificador; nestes casos, haveria a possibilidade de detecção de situações não percebidas pelo especialista humano.

Alguns trabalhos futuros foram vislumbrados durante o desenvolvimento desta pesquisa, entre eles: (1) uso do sistema inteligente para identificar falhas específicas em rolos de correia transportadora; (2) monitoramento por longo período de tempo para estudo de métodos de previsão de vida útil; (3) estudo de eletrônica e mecânica para acoplamento do sistema em um robô de inspeção de correias.

## Referências Bibliográficas

AKT KUPPER. **Carrying Idlers Compact**. Disponível em <<https://www.kuepper.eu/en/carrying-idlers/compact.html>>. Acesso em: 25 jun. 2019. Citado na página 19.

ALI, Jaouher Bem; FNAIECH, Nader; SAIDA, Lotfi; CHEBEL-MORELLO, Brigitte; FNAIECH, Farhat. **Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals**. *Applied Acoustics* 89 (2015) 16–27. Citado na página 28.

CONDITION MONITORING. **Condition Monitoring of Critical Mining Conveyors**. Disponível em <<https://www.emerson.com/documents/automation/article-condition-monitoring-of-critical-mining-conveyors-en-us-21248.pdf>>. Acesso em: 22 jun. 2019. Citado na página 17.

DELGADO-ARREDONDO, P.A.; MORINIGO-SOTELO, Daniel; OSORNIO-RIOS, R.A.; AVINA-CERVANTES, J.G.; ROSTRO-GONZALEZ, Horacio; ROMERO-TRONCOSO, R.J. **Methodology for fault detection in induction motors via sound and vibration signals**. *Mechanical Systems and Signal Processing*, 2016. Citado na página 30.

DOS SANTOS, U.J.L.; PESSIN, Gustavo; DA COSTA, C.A.; RIGHI, R.R. **AgriPrediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops**, *Computers and Electronics in Agriculture* 161 (2019) 202–213. Citado na página 34.

ENGELBRETCH, A.P. **Computational Intelligence: An Introduction**, Wiley Publishing, 2007. Citado na página 32.

FERREIRA, J.J.; CARVALHO, Eduardo; FERREIRA, B.V.; DE SOUZA, Cleidson.; SUHARA, Yoshihiko; PENTLAND, Alex; PESSIN, Gustavo. **Driver behavior profiling: An investigation with different smartphone sensors and machine learning**. *PLoS ONE* 12(4): e0174959, 2017. Citado na página 35.

FLEXCO INSIGHTS. **Roller Performance - Technical Solutions For Belt Conveyor Productivity**. Disponível em <[http://documentlibrary.flexco.com/X2640\\_enAU\\_2525\\_INSCCTlife\\_0813.pdf](http://documentlibrary.flexco.com/X2640_enAU_2525_INSCCTlife_0813.pdf)>. Acesso em: 24 jun. 2018. Citado na página 18.

GARCIA-MARQUEZ, F.P.; TOBIAS, A.M.; PEREZ, J.M.P.; PAPAELIAS, Mayorkinos. **Condition monitoring of wind turbines: techniques and methods**. Renewable Energy, v. 46, p. 169-178, 2012. Citado na página 30.

GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel. **Data mining: um guia prático**. Gulf Professional Publishing, 2005.

HALL M; FRANK E; HOLMES G; PFAHRINGER B, REUTEMANN P, WITTEN IH (2009) **The WEKA data mining software: an update**. SIGKDD Explor 11. Citado na página 49.

HOFFMANN, A.J.; van der MERWE, N.T. **The application of neural networks to vibrational diagnostics for multiple fault conditions**. Computer Standards & Interfaces 24 (2002) 139–149. Citado na página 28.

INGENIUM DESIGN. **Conveyor Belt Design Engineering**. Disponível em <<http://www.ingeniumdesign.us/services/conveyor-belt-design-engineering/>>. Acesso em: 22 jun. 2019. Citado na página 18.

KEDADOUCHE, Mourad; THOMAS, Marc; TAHAN, Antoine. **Empirical Mode Decomposition of Acoustic Emission for Early Detection of Bearing Defects. Advances in Condition Monitoring of Machinery in Non-Stationary Operations**. Lecture Notes in Mechanical Engineering. Springer, Berlin, Heidelberg, p. 367-377, 2014. Citado na página 31.

KIM, Y.H.; TAN, Andy; MATHEW, Joseph; YANG, B.S. **Condition monitoring of low speed bearings: a comparative study of the ultrasound technique versus vibration measurements**. WCEAM, v. 29, 2006. Citado na página 30.

KOEHRSEN, Will. **Random Forest Simple Explanation**. Disponível em <<https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d>>. Acesso em: 14 mai. 2019. Citado na página 34.

KRÓL, Robert; ZIMROZ, Radoslaw. **Failure analysis of belt conveyor systems**, Politechnika Wroclaw, Instytut Górnictwa, pl. Teatralny 2, 50-051 Wroclaw. Citado na página 18.

LATHI, B.P. **Signal Processing and Linear Systems**, Berkeley Cambridge Press, 1998. Citado na página 38.

LI, Yaoyu; YANG, Zhongzhou; WU, Xin. **A Review of Recent Advances in Wind Turbine Condition Monitoring and Fault Diagnosis**. Conference Paper · July 2009. DOI: 10.1109/PEMWA.2009.5208325 · Source: IEEE Xplore. Citado na página 29.

LODEWIJKS, Gabriel. **Strategies for automated maintenance of belt conveyor systems**, in: Proceedings of the International Materials Handling Conference (Beltcon) 12, Johannesburg, South Africa, 2003. Citado na página 18.

MARTINS, W.A.; CAMPOS, M.L.R.; CHAVES, R. S.; LORDELO, C. P. V.; ELLMAUTHALER, Andreas; NUNES, L.O.; BARFOOT, D.A. **Communication Models for Distributed Acoustic Sensing for Telemetry**. IEEE Sensors Journal, VOL. 17, NO. 15, AUGUST 1, 2017. Citado na página 29.

MINING3. **Conveyor belt monitoring for wear detection - Pre-empting roller failure using fibre optics**. Disponível em <<https://www.mining3.com/research/das-conveyor-belt/>>. Acesso em: 14 jun. 2018. Citado na página 28.

MITCHELL, T.M. **Machine Learning**, McGraw-Hill, 1997. Citado na página 32.

MJIT, Mustapha; BEAUJEAN, Pierre-Philippe J.; VENDITTIS, David J. **Comparison of Fault Detection Techniques for an Ocean Turbine**. Annual Conference of the Prognostics and Health Management Society, 2011. Citado na página 28.

ROKACH, Lior; MAIMON, Oded Z. **Data mining with decision trees: theory and applications**. World scientific, 2008.

RUSSELL, S.J.; NORVIG, Peter. **Artificial Intelligence: A Modern Approach**, Prentice Hall, 2009. Citado na página 32.

STACK EXCHANGE. **Intuition Behind Fourier Coefficients**. Disponível em <<https://math.stackexchange.com/questions/340429/intuition-behind-fourier-coefficients>>. Acesso em 22 jun. 2019. Citado na página 37.

TCHAKOUA, Pierre; WAMKEUE, Rene; TAMEGHE, T.A.; EKEMB, Gabriel. **A review of concepts and methods for wind turbines condition monitoring**. Computer and Information Technology, World Congress, 2013. Citado na página 30.

THEOBALD, Oliver. **Machine Learning For Absolute Beginners**, 2017. Citado na página 33.

WHEELER, Craig; AUSLING, Daniel. **Evolutionary belt conveyor design**, in: Proceedings of the International Materials Handling Conference (Beltcon) 14, Johannesburg, South Africa, 2007. Citado na página 18.

WITTEN, Ian H. et al. **Data Mining: Practical machine learning tools and techniques**. Morgan Kaufmann, 2016.

ZUDE, Zhou; QUAN, Liu; QINGSONG, Ai; CHENG, Xu. **Intelligent monitoring and diagnosis for modern mechanical equipments based on the integration of embedded technology and FBGS technology**. Measurement, v. 44, p. 1499-1511, 2011. Citado na página 29.

## Anexo A: Código fonte para extração de dados no domínio do tempo

```
% Lê o arquivo wav
% y = m x n matriz normalizada entre -1 e 1 com os dados
%     m = numero de amostras lidas
%     n = numero de canais de audio
% fs = numero de amostras por segundo
[y,fs] = audioread('arquivo de som');

% t = vetor do tempo
% linspace(tempo inicial,tempo final,numero de amostras em y)
t = linspace(0,length(y)/fs,length(y));

plot(t,y); xlabel('Tempo (segundos)'); ylabel('Amplitude');
ax = gca; % current axes
ax.FontWeight = 'bold';
ax.FontSize = 28;

% comprimento da fft
Nfft = 4096;
% f = vetor da frequencia
% linspace(frequencia inicial, frequencia final, tamanho do vetor
% frequencia)
f = linspace(0,fs,Nfft);
% G = FFT das amostras y em Nfft pontos
G = abs(fft(y,Nfft));

figure
plot(f(1:Nfft/2),G(1:Nfft/2)); xlabel('Frequência (Hz)');
ylabel('Magnitude');
ax = gca; % current axes
ax.FontWeight = 'bold';
ax.FontSize = 28;

% Valores iniciais para o laço
i = 1;
rep_fft = 1;
incremento = 25; % define o número de frequências que irá conter cada uma
das 40 partes
incremento_rep = 25;

% Laço que divide a FFT em 40 partes
for n = 1:40

    vetor_media_fft = G(i:incremento_rep);
    media_fft(rep_fft) = mean(vetor_media_fft);
    i = i + incremento;
    rep_fft = rep_fft + 1;
    incremento_rep = rep_fft*incremento;

end
```

```

% Valores iniciais para o laço
size_y = numel(y); % número de dados no tempo
i = 1;
tempo = 8000; % define o intervalo de tempo, sendo que 16000 corresponde a
1 segundo
tempo_rep = 8000;
rep = 1;

while i < size_y-tempo

    vetor_media = y(i:tempo_rep);
    media(rep) = mean(vetor_media);
    mediana(rep) = median(vetor_media);
    desvio(rep) = std(vetor_media);
    i = i + tempo;
    rep = rep + 1;
    tempo_rep = rep*tempo;

end

media = media';
mediana = mediana';
desvio = desvio';

```

## **Anexo B: Cabeçalho do dataset dos experimentos**

Número de gravações:

- Rolos não defeituosos: 10 gravações
- Rolos defeituosos: 10 gravações

Sensibilidade ajustada do instrumento:  $S = 60$  a  $65$

Frequência central: 40kHz

Datas das gravações:

- Dia 20/02/19, entre 14 e 16hrs: 4 rolos defeituosos e 3 rolos não-defeituosos;
- Dia 08/03/19, entre 10 e 12hrs: 3 rolos defeituosos e 7 rolos não-defeituosos;
- Dia 27/03/19, entre 14 e 16hrs: 3 rolos defeituosos.

## Anexo C: Dataset do Experimento 1 (20 primeiras linhas)

Número de amostras:

- 378 amostras de meio segundo de rolos não defeituosos
- 390 amostras de meio segundo de rolos defeituosos

media,mediana,desvio\_padrao,classe

0.021912598,0.013671875,0.252225598,rnd

0.022027344,0.013671875,0.222051151,rnd

0.022100342,0.017578125,0.191203734,rnd

0.021743896,0.017578125,0.179718154,rnd

0.021976563,0.017578125,0.159278024,rnd

0.02200293,0.01953125,0.14433133,rnd

0.021865479,0.01953125,0.141617304,rnd

0.0219021,0.016601563,0.14815312,rnd

0.021931152,0.017578125,0.151325552,rnd

0.02181665,0.017578125,0.149100159,rnd

0.02189917,0.017578125,0.142868812,rnd

0.021979492,0.017578125,0.140081976,rnd

0.021940674,0.01953125,0.125942659,rnd

0.021846436,0.017578125,0.122422283,rnd

0.021955078,0.01953125,0.11548701,rnd

0.021840576,0.01953125,0.105789922,rnd

0.021938232,0.01953125,0.104226249,rnd

0.02195874,0.01953125,0.111736952,rnd

0.021909912,0.01953125,0.113219376,rnd

## Anexo D: Dataset do Experimento 2 (20 primeiras linhas)

Número de amostras:

- 192 amostras de meio segundo de rolos não defeituosos
- 206 amostras de meio segundo de rolos defeituosos

media1,mediana1,desvio1,media2,mediana2,desvio2,media3,mediana3,desvio3,media4,mediana4,desvio4,classe

0.021930176,0.013671875,0.254397433,0.02189502,0.01171875,0.250066708,0.022183594,0.013671875,0.231840223,0.021871094,0.015625,0.211839135,amd

0.022016113,0.015625,0.187228573,0.02218457,0.017578125,0.195121309,0.021811523,0.01953125,0.190758837,0.02167627,0.017578125,0.16797727,amd

0.022108398,0.017578125,0.165065677,0.021844727,0.017578125,0.153292561,0.021974609,0.01953125,0.14852337,0.02203125,0.01953125,0.140032432,amd

0.021837402,0.01953125,0.140623536,0.021893555,0.017578125,0.142621725,0.021921875,0.017578125,0.139967215,0.021882324,0.015625,0.155927421,amd

0.021896973,0.01953125,0.156978855,0.021965332,0.017578125,0.145472364,0.02182959,0.014648438,0.15051652,0.021803711,0.01953125,0.147689034,amd

0.02192041,0.017578125,0.142481129,0.02187793,0.017578125,0.143273257,0.022072266,0.017578125,0.142687815,0.021886719,0.01953125,0.137444525,amd

0.021958496,0.01953125,0.13106433,0.021922852,0.01953125,0.120620122,0.021785645,0.017578125,0.126158133,0.021907227,0.017578125,0.118584555,amd

0.021968262,0.021484375,0.115400067,0.021941895,0.017578125,0.115588315,0.021696289,0.01953125,0.106046005,0.021984863,0.021484375,0.10554628,amd

0.021896973,0.021484375,0.104536019,0.021979492,0.01953125,0.103928608,0.02203418,0.018554688,0.110799316,0.021883301,0.01953125,0.112680589,amd

0.021908691,0.018554688,0.112839234,0.021911133,0.01953125,0.113612354,0.021915527,0.017578125,0.119149436,0.02165332,0.017578125,0.117012399,amd

0.022137695,0.017578125,0.115369283,0.021951172,0.01953125,0.120528995,0.02184082,0.021484375,0.130650024,0.02185498,0.01953125,0.118225429,amd

0.02192627,0.021484375,0.118243813,0.021967285,0.01953125,0.119583733,0.021830566,0.01953125,0.115779923,0.021948242,0.01953125,0.108107895,amd

0.021914551,0.021484375,0.104011081,0.021825195,0.021484375,0.106509522,0.022058594,0.017578125,0.097490708,0.021894043,0.021484375,0.098951671,amd

0.02174707,0.017578125,0.102414662,0.021996094,0.01953125,0.101794234,0.021885742,0.01953125,0.105887462,0.021910156,0.01953125,0.109746993,amd

0.021884277,0.021484375,0.1127506,0.021831055,0.021484375,0.121699878,0.02207373,0.01953125,0.120466405,0.021897949,0.01953125,0.117249395,amd

0.021936523,0.020507813,0.112528221,0.021947754,0.01953125,0.110101788,0.021835938,0.021484375,0.114498299,0.021975586,0.021484375,0.113237028,rnd

0.021907715,0.021484375,0.105804883,0.021773926,0.017578125,0.107757568,0.021845215,0.021484375,0.109592857,0.022074219,0.021484375,0.102234318,rnd

0.021880371,0.021484375,0.099914989,0.021960938,0.017578125,0.105994778,0.021853027,0.021484375,0.110664142,0.021896973,0.01953125,0.105858204,rnd

0.022087891,0.021484375,0.106495988,0.02205957,0.01953125,0.093012177,0.021802246,0.021484375,0.099301435,0.021888672,0.01953125,0.104450345,rnd

## Anexo E: Dataset do Experimento 3 (20 primeiras linhas)

Número de amostras:

- 400 amostras de frequências
- 400 amostras de frequências

média,mediana,desvio\_padrao,classe

7.318503844,5.155309998,9.487495395,rnd

14.39967553,14.09253663,6.792889881,rnd

21.82979445,21.17622428,12.11058077,rnd

23.7204321,23.95300255,13.11257286,rnd

23.83075104,23.37274119,11.1953008,rnd

25.54812714,23.6986165,13.7702293,rnd

25.43337994,25.37824105,12.26593395,rnd

23.74345653,21.45483519,13.0499992,rnd

17.75668144,17.37976977,8.374123462,rnd

10.46626225,9.533033621,6.105008783,rnd

5.726527155,5.302668485,2.748286568,rnd

3.126676897,3.04589519,1.638170372,rnd

2.242219429,2.123241907,1.286575551,rnd

1.888694421,1.822915634,0.881556605,rnd

1.433479002,1.412207409,0.66998114,rnd

1.020626451,0.95728537,0.470917432,rnd

0.73452112,0.683271287,0.397175551,rnd

0.582444085,0.582376984,0.23636605,rnd

0.605815925,0.576166158,0.27302221,rnd

## Anexo F: Dataset do Experimento 4 (20 primeiras linhas)

Número de amostras:

- 40 amostras de frequências
- 40 amostras de frequências

7.318503844,14.39967553,21.82979445,23.7204321,23.83075104,25.54812714,25.43337994,23.743  
45653,17.75668144,10.46626225,5.726527155,3.126676897,2.242219429,1.888694421,1.433479002  
,1.020626451,0.73452112,0.582444085,0.605815925,0.427026792,0.454183861,0.541894515,0.6373  
44107,0.723842701,0.908065314,1.409419128,1.702053601,2.296057845,2.892156287,5.040859792  
,9.437503575,16.68471285,23.12903311,25.62083422,24.71238843,24.81672746,23.45424614,23.85  
09476,14.50095425,7.853615308,rb

3.915874417,7.017469914,10.9107546,12.24022289,13.35758462,14.27337073,15.08712483,14.204  
22792,10.7799938,5.840816158,3.277332874,1.818257636,1.081131242,0.707706505,0.460749012,  
0.392303615,0.289335015,0.230794047,0.195506684,0.190719469,0.177217916,0.194414009,0.221  
185257,0.277352127,0.373990598,0.42930505,0.658465703,1.053306767,1.627341141,2.872789416  
,5.533600376,10.23041316,13.95877232,14.85206284,13.69901234,13.76404833,12.67216484,11.28  
701339,7.578571512,3.838268994,rb

4.587962657,6.728706881,8.499029827,9.51740852,12.01848221,10.25450258,11.68095194,11.112  
86694,8.784604784,5.424020129,2.800590758,1.508044204,0.929310772,0.648410813,0.453892411  
,0.348665871,0.29776294,0.250317734,0.235003979,0.224855101,0.227597311,0.244259424,0.2372  
43031,0.281042924,0.338477794,0.447293437,0.603878188,0.870970937,1.324003192,2.52123197,  
4.792654798,8.540684093,10.73065841,11.50198011,10.77276701,11.4268559,9.961545234,8.6913  
90784,7.273405025,4.485006471,rb

5.240992396,9.826199276,14.59683685,14.34644672,11.45239017,5.856516521,4.087656921,2.670  
340153,1.639379026,1.111350959,0.730494873,0.516129737,0.61327463,0.617574227,0.398687086  
,0.233398622,0.176211359,0.144056986,0.1386312,0.141436276,0.150657089,0.139100132,0.14170  
9789,0.16078935,0.232490557,0.34206054,0.613561349,0.632852605,0.501916501,0.683239037,1.0  
55107384,1.503952832,2.343864075,4.147240359,5.385473447,9.895702561,14.36581305,15.35683  
338,10.19590125,5.691527986,rb

3.771244672,5.814348316,7.443345818,7.734593745,6.218909946,4.794008381,3.271340651,2.239  
836285,1.460580374,0.796756385,0.614324765,0.438091317,0.53325411,0.599056645,0.35133009,  
0.219937346,0.154766105,0.135262015,0.120626801,0.107183608,0.106169317,0.121355,0.132813

985,0.149314829,0.205616563,0.31449721,0.600333532,0.526850161,0.453877399,0.586677465,0.6  
9877236,1.414551812,2.038288673,3.046721515,4.523042517,6.12457742,7.497263169,7.46525102  
3,6.143696093,3.698496897,rb

3.071446969,4.408184908,6.616973059,7.186954248,5.801639731,3.403610327,2.400493895,1.659  
00314,1.118726756,0.70532054,0.504260582,0.3651624,0.452454905,0.46472294,0.236906473,0.15  
4799874,0.145925402,0.128857026,0.11662269,0.106600477,0.107154587,0.110886005,0.12923454  
6,0.137388582,0.155292541,0.211735395,0.451639764,0.462165344,0.357597111,0.491794774,0.65  
0350661,0.959523504,1.618154213,2.273306545,3.32867148,5.157590197,7.181744271,7.19298928  
7,4.361411348,2.757535028,rb

6.08073649,12.1239161,16.09193491,17.60697819,21.47955796,18.71961242,10.26831497,6.76403  
0968,4.295303502,2.610553314,2.720028111,1.908031995,0.937608346,0.624853381,0.465700414,  
0.391426908,0.290865883,0.228944907,0.240274439,0.209777012,0.215703716,0.229826704,0.242  
980012,0.283537084,0.349575295,0.439713094,0.619121303,0.854410525,1.677684433,2.55889925  
4,2.749798753,3.964292622,6.203253541,9.510719306,17.22134211,21.67566176,18.24100938,16.8  
4284783,12.50330828,6.654631402,rb

6.237630557,11.83947839,18.07664405,19.27743551,20.26504969,17.9357044,12.34313615,6.4504  
44733,4.138269791,2.620481608,2.902566468,2.193539335,0.905367087,0.65966321,0.444493347,  
0.331393893,0.283467743,0.237680343,0.205625089,0.154718825,0.163861374,0.200060747,0.226  
402781,0.273852639,0.316930365,0.434572704,0.588023398,0.841027278,1.873523063,2.94174971  
8,2.638571517,3.880854663,5.75306114,11.13889008,17.87150214,19.25981535,20.29811551,18.30  
02028,12.69163,6.705675138,rb

3.335681098,5.097699651,7.596394109,8.516851486,9.948852334,9.619180953,5.196275751,3.352  
444873,1.695940741,0.957139911,1.069816419,0.84102449,0.517734063,0.304674077,0.225631608  
,0.175740018,0.154677257,0.124935663,0.101247506,0.094255677,0.099103978,0.099918071,0.120  
496974,0.148536804,0.167160753,0.213008555,0.289833381,0.489371677,0.827539518,1.03064033  
,0.941001016,1.514497078,2.854805827,4.787179956,9.018288583,10.33184528,8.63603515,7.9129  
70356,5.524035092,2.976893657,rb

4.302623793,7.170105065,9.663192518,14.42808051,17.88652713,17.11444291,9.248677515,4.473  
180961,2.573820579,1.445229196,1.256556453,0.976568325,0.613111632,0.373754513,0.29408994  
6,0.240940489,0.195862889,0.170626992,0.154303445,0.160820075,0.163853885,0.16443762,0.165  
09616,0.190063921,0.226882845,0.277159581,0.375698694,0.542035421,0.969498006,1.189575396  
,1.404534752,2.273705432,4.157098448,8.288080511,15.73937834,18.52044021,15.03268763,10.17  
988648,7.573207598,4.339180502,rb